Modélisation compositionnelle et cinétique des bio-huiles de pyrolyse rapide issues de la biomasse lignocellulosique

Ana Rita Costa da Cruz

N°d'ordre NNT : 2019LYSE1006

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

École Doctorale 206
(Chimie, Procédés, Environnement)

Spécialité de doctorat : Procédés

Soutenue publiquement le 25/01/2019, par :
Ana Rita Costa da Cruz

Modélisation compositionnelle et cinétique des bio-huiles de pyrolyse rapide issues de la biomasse lignocellulosique

---

Devant le jury composé de :

Mme TAYAKOUT-FAYOLLE, Mélaz	Professeure, LAGEP	Présidente
M. DUFOUR, Anthony	Chargé de Recherche CNRS, LRPG	Rapporteur
M. THYBAUT, Joris	Professeur, Ghent University	Rapporteur
Mme BRUNET, Sylvette	Directrice de Recherche CNRS, Université de Poitiers	Examinatrice
Mme ESSAYEM, Nadine	Directrice de Recherche CNRS, IRCELYON	Examinatrice
M. KLEIN, Michael	Professeur, University of Delaware	Examinateur
M. JOLY, Jean-François	Ingénieur, IFPEN	Directeur de thèse
Mme CHARON, Nadège	Ingénieure, IFPEN	Invitée
M. VERSTRAETE, Jan	Ingénieur, IFPEN	Invité

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. le Professeur Frédéric FLEURY**
Président du Conseil Académique	M. le Professeur Hamda BEN HADID
Vice-président du Conseil d'Administration	M. le Professeur Didier REVEL
Vice-président du Conseil Formation et Vie Universitaire	M. le Professeur Philippe CHEVALIER
Vice-président de la Commission Recherche	M. Fabrice VALLÉE
Directrice Générale des Services	Mme Dominique MARCHAND

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur G.RODE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. X. PERROT
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. THEVENARD
Département Chimie Biochimie	Directeur : Mme C. FELIX
Département GEP	Directeur : M. Hassan HAMMOURI
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur G. TOMANOV
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : M. le Professeur J-C PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Y.VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. le Professeur E.PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

"[...] "Hallo!" said Piglet, "what are you doing?"

"Hunting," said Pooh.

"Hunting what?"

"Tracking something," said Winnie-the-Pooh very mysteriously.

"Tracking what?" said Piglet, coming closer.

"That's just what I ask myself. I ask myself, What?"

"What do you think you'll answer?"

"I shall have to wait until I catch up with it," said Winnie-the-Pooh. [...]"

A. A. Milne, "Pooh and Piglet Go Hunting and Nearly Catch a Woozle" in *Winnie-the-Pooh*

"Sob céus e climas estranhos, rodeada de cangurus, ameaçada certamente pelas tribos primitivas do interior, ao alcance do terrível boomerang, a couve portuguesa dá uma lição de constância e de fidelidade às origens, ao mesmo tempo que mostra ao mundo as nossas raras qualidades de adaptação, o nosso universalismo, a nossa vocação de viajantes. E continua a crescer."

José Saramago, "Elogio da Couve Portuguesa" in *A Bagagem do Viajante*

Acknowledgements

This thesis was carried out at *IFP Energies nouvelles* (*IFPEN*) in Solaize within the division *Conception Modélisation Procédés*, here I would like to thank to Pierre Porot, Luc Nougier and Christophe Boyer for all trust placed in me and as well for providing me with the opportunity to do this thesis.

In the following context, I'm also thankful to Anthony Dufour and Joris Thybaut, who accepted to review the manuscript and also to Sylvette Brunet, Nadine Essayem, Michael Klein and Mélaz Tayakout-Fayolle, who accepted to be a part of the jury. Thank you all for your suggestions and corrections, which help improve this work. To Michael, I would like to extend my appreciation for all the suggestions and remarks he made when visiting *IFPEN*.

I would also like to thank to Jean-François Joly, director of the thesis, for all his support and guidance during these three years. His perspective over this project lead me to the correct strategy and taught me to go beyond my summarized explanations into a more pedagogic approach.

To Jan Verstraete, the promotor of this thesis, I also leave a (very big) work of appreciation. So far you are the best "teacher" I ever had and it was a pleasure to learn from you. I wish you existed in a pocket version! Thank you for all your help, advice, patience for my stubbornness and above all, for teaching me that I can do much more than what I actually think I can. I still have difficulty in believing in myself, but thanks to you, at least now I try before I quit.

To Nadège Charon, the co-promotor of this thesis, I would like to thank you for all the enthusiasm, support and patience for my French. I could not have asked for a better person to teach me all about analytical techniques and the challenges behind biomass and bio-oil. Thank you once again, in particular to your kind smile that always made our meetings better.

I'm also thankful to Luis Oliveira, the debug expert, and to Damien Hudebine, the backup of the debug expert, for all the help in the programing and for all the advices not always work related. It was a pleasure to be a part of what you created (with Jan) and be able to extend it further. I hope I made you all proud!

Regarding the experimental part, I could not have made it without Elodie Dusson, Manuel Menard and Yohann Mouillet. To Elodie, I owe all the teachings related to the experimental setup and the push to improve my French. To Manuel Menard and Yohann Mouillet, I thank you for your last-minute help with my many, weird and not easily analyzable samples.

These acknowledgements could not continue without a big thanks to Tiago Sozinho and Alberto Servia. Thank you both for the support and encouragement, special in the last moments. I would also like to thank the entire R123 department and Claire, Isabelle and Martine for the occasional team building

activities. I had a lot of fun, although I never led any of my teams to victory. Thank you all as well, for the endless help, patience and kindness.

A word of appreciation goes also to Ana Pinheiro, who always helped and supported me since the first moment I arrived to Lyon until the every last.

To my office mates, Sonia, Caroline, Charlotte, Damien, Vincenzo, Oliver, Hanane, Stefan, Phuong, Giulia (for 1 month) and Maxime (for two weeks), I thank you for all the moments, laughs and carpooling. To Gabriel, who appeared after I left, thank you for all your kindness and last-minute help. To my partners in crime, Sonia, Caroline, Charlotte and Damien, I thank you for everything, especially the support, patience and help with my French. I'm also thankful to the Portuguese crew, Leonor, Ruben, Mariza, Ana Teresa, Diogo, Pedro and Ana Sofia, to my Brazilian sunshines, Larissa and Izabel, and also to Elsy, Julie, July, Marco, Geoffrey and Rick. Thank you all and everyone who helped me in a way or another and always tried to make me feel at home.

A vast thanks also goes to Ester Gutiérrez, who even though she caught the good and bad days, always believed and stood there for me. In Lyon, you were my family and to you, I owe all the support words, the happy laughs and all the other great moments, who forced me out of my bubble. More is yet to come, but until there, *muchas gracias Estrelita*!

Although not always physically present, I cannot go by without thanking my friends who always support and stand by me. To Bárbara, Diogo, Filipa, André, Mariana and Rita, I leave a huge thanks. We started all this together and I hope we can also end up this together. *Siga o próximo!*

À minha família, em especial aos meus pais, não há nenhuma palavra no mundo para vos agradecer. Eu sou quem sou graças a todos vós e as oportunidades que me propocionaram e desafiaram a aceitar. Se eu tenho força, é porque nós somos fortes. Se eu tenho coragem, é porque nós somos corajosos. E se eu tenho saudade, é porque vocês são o meu mundo. Obrigada, obrigada e para sempre obrigada.

For last but not least, this work is dedicated to my aunt, Maria Helena Costa. Wherever you are, I know you are very proud of me. *Para ti, da tua Titinha, a carochinha muito rosadinha…*

Thank you all!

Merci à tous!

Muito Obrigada!

Résumé

La pyrolyse rapide est une des voies de conversion thermochimique qui permet la transformation de biomasse lignocellulosique en bio-huiles. Ces bio-huiles, différentes des coupes lourdes du pétrole, ne peuvent pas être directement incorporées dans des unités existantes de raffinage. En effet, en raison de leur forte teneur en oxygénés, les bio-huiles nécessitent une étape de pré-raffinage, telle que l'hydrotraitement, afin d'éliminer ces composés.

L'objectif de ce travail est de comprendre la structure, la composition et la réactivité de bio-huiles grâce à la modélisation de données expérimentales. Pour comprendre leur structure et leur composition, des techniques de reconstruction moléculaire basées sur des données analytiques ont été appliquées générant ainsi un mélange synthétique, dont les propriétés correspondent à celles de la bio-huile. Pour comprendre leur réactivité, l'hydrotraitement de molécules modèles a été étudié : gaïacol et furfural. Pour cela, un modèle déterministe et un modèle stochastique ont été créés pour chacun d'eux. L'approche déterministe visait à déterminer une gamme de paramètres cinétiques, qui ont ensuite été affinés par l'approche stochastique créant un nouveau modèle. Cette approche a permis de générer un réseau de réactions en définissant et en utilisant un nombre limité de familles et règles des réactions. Finalement, le mélange synthétique a été utilisé dans la simulation stochastique de l'hydrotraitement de la bio-huile, étayée par la cinétique des molécules modèles.

En conclusion, ce travail a permis de recréer la fraction légère de la bio-huile et de simuler leur l'hydrotraitement, via les paramètres cinétiques des composés modèles, qui prédisent de manière raisonnable les effluents de l'hydrotraitement de celles-ci.

Mots-Clés : Bio-huile, Biomasse lignocellulosique, Pyrolyse rapide, Composés oxygénés, Modélisation Compositionnelle, Modélisation Cinétique, Composés Modèles, Reconstruction Moléculaire, Simulation Stochastique

Abstract

Fast pyrolysis is one of the thermochemical conversion routes that enable the transformation of solid lignocellulosic biomass into liquid bio-oils. These complex mixtures are different from oil fractions and cannot be directly integrated into existing petroleum upgrading facilities. Indeed, because of their high levels of oxygen compounds, bio-oils require a dedicated pre-refining step, such as hydrotreating, to remove these components.

The aim of the present work is to understand the structure, composition and reactivity of bio-oil compounds through modeling of experimental data. To understand the structure and composition, molecular reconstruction techniques, based on analytical data, were applied to generate a synthetic mixture, whose properties are consistent with those of the bio-oil. To understand the reactivity, the hydrotreating of two model molecules was studied: Guaiacol and Furfural. A deterministic and a stochastic model were created for each compound. The deterministic approach intended to retrieve a range of kinetic parameters, later on refined by the stochastic simulation approach into a new model. This approach generates a reaction network by defining and using a limited number of reaction classes and reaction rules. To consolidate the work, the synthetic mixture was used in the stochastic simulation of the hydrotreating of bio-oils, supported by the kinetics of the model compounds.

In summary, the present work was able to recreate the light fraction of bio-oils and simulate the hydrotreating of bio-oils, via the kinetic parameters of model compounds, which can reasonably well predict the effluents of the hydrotreating of these.

Key words: Bio-oil, Lignocellulosic biomass, Fast pyrolysis, Oxygenated compounds, Composition Modeling, Kinetic Modeling, Model Compounds, Molecular Reconstruction, Stochastic Simulation.

Title: Composition and kinetic modeling of bio-oil from fast pyrolysis from lignocellulosic biomass.

Résumé Substantiel

La biomasse est une matière organique avec un fort potentiel de production de carburants et de molécules biosourcées. Pour obtenir ces composés, la biomasse doit d'abord être prétraitée et convertie en une matière facilement gérable, sous forme liquide. Un de ces procédés est la pyrolyse rapide, qui génère un solide, un gaz et un liquide, appelé bio-huile. Ce mélange est une matrice oxygénée complexe, avec un grand nombre de molécules, et par conséquent, une grande diversité de caractéristiques. Dans la bio-huile, il est possible de trouver des composés tels que les alcools, les aldéhydes, les sucres, les acides carboxyliques, les furanes, les cétones et aussi les phénols et méthoxy-phénols. Les caractéristiques moléculaires d'une bio-huile sont très compliquées à acquérir expérimentalement. La quantité d'oxygène dans la bio-huile ne permet pas de la mélanger avec des carburants classiques composés d'hydrocarbures.

Grâce à la conversion de bio-huiles, il est possible d'obtenir des carburants et d'autres composants chimiques. L'hydrotraitement est le procédé le plus courant pour valoriser les bio-huiles. En raison de la complexité associée au mélange et au procédé, la communauté scientifique a commencé par l'analyse l'hydrotraitement des composés modèles de bio-huile. En bref, ce mélange et ses molécules modèles, lorsqu'elles sont dans les conditions d'hydrotraitement, peuvent être converties en hydrocarbures et en eau par des réactions d'hydrogénation, d'hydrogénolyse, d'hydrodésoxygénation, de décarboxylation, de décarbonylation, de craquage / hydrocraquage et de polymérisation (Wildschut 2009). Il est important de noter que ces réactions sont extrêmement influencées par le solvant, le catalyseur, la température et la structure de la molécule.

Même si ces modèles sont extrêmement utiles pour étudier, identifier et quantifier des informations intrinsèques sur la cinétique des bio-huiles, le réseau réactionnel global de la bio-huile ne peut être totalement investigué. Des techniques appropriées pour les mélanges complexes doivent être mises en œuvre afin d'obtenir une description moléculaire de ces liquides, du réseau réactionnel et du modèle cinétique, ainsi qu'une simulation correcte du mélange complexe sans grande difficulté. Pour la présente thèse, il a été considéré que coupler la reconstruction stochastique et la reconstruction par maximisation d'entropie serait une bonne approche pour la description moléculaire des bio-huiles, tandis que l'algorithme de simulation stochastique serait la meilleure méthode pour la simulation de l'hydrotraitement.

Le but du présent travail était de comprendre la structure, la composition et la réactivité des composés de bio-huiles à travers la modélisation de données expérimentales. Pour y parvenir, la présente thèse a utilisé une méthode de reconstruction moléculaire stochastique et un algorithme de simulation stochastique pour, in silico, gérer la complexité du système, depuis les molécules modèles simples jusqu'à la matrice complexe, la bio-huile.

Afin de suivre la stratégie proposée, le processus d'hydrotraitement de deux molécules modèles (gaïacol et furfural) a été analysé. Les tests expérimentaux ont été effectués à 13 MPa sur un catalyseur NiMo / Al_2O_3 réduit, entre 100 °C à 300 °C et avec un intervalle de temps de réaction entre 1 h à 5 h. Trois tests « blancs » ont également été effectués pour les deux molécules modèles (un sur le guaiacol et deux sur le furfural). À la fin, le bilan matière global a été bouclé pour tous les tests, avec une valeur moyenne de 101%. Les effluents liquides et gazeux de ces tests ont été analysés par GC-FID/MS et GC-FID/TCD. En résumé, les résultats des tests expérimentaux représentent plus de 80% de carbone quantifié (60% pour l'essai à 250 °C). Pour cette raison, il a été considéré que les résultats étaient appropriés pour l'application de la stratégie de cette thèse.

Pour tous les tests expérimentaux, la consommation d'hydrogène était presque constante. Cependant, celle-ci augmente avec le temps de réaction, suggérant une augmentation des réactions d'hydrotraitement, et diminue avec la température, ce qui signifie qu'à haute température, l'hydrogène commence à être produit. En ce qui concerne la conversion du gaïacol et du furfural, les deux augmentent avec le temps de réaction et la température comme prévu. Cependant, une ligne de base dans la conversion du guaiacol entre 1 h et 3 h de réaction peut être observée. Cela peut être justifié par la compétition entre le gaïacol et ses dérivés pour les sites catalytiques (Wang et al. 2013). En ce qui concerne la distribution des produits, il est assez difficile d'identifier les voies réactionnelles de chaque composé, car celles-ci peuvent provenir de plusieurs réactions et réactifs différents. Cependant, et d'accord avec les produits analysés, il a été considéré que le gaïacol se convertit principalement par déméthylation, déméthoxylation, hydrodésoxygénation, saturation et transalkylation, tandis que le furfural subirait des réactions d'hydrogénation du carbonyle, hydrodésoxygénation, saturation, ouverture de cycle et décarbonylation. Ces réactions et composés ont été considérés pour la création d'un modèle cinétique d'hydrotraitement du gaïacol et du furfural.

Un modèle déterministe a été créé pour chaque composé modèle, afin de déterminer une gamme appropriée de paramètres cinétiques à utiliser ultérieurement dans le modèle stochastique SSA (de l'anglais « Stochastic Simulation Algoritm ». Cette gamme a été basée sur des tests expérimentaux des deux composés modèles et dans les bilans matière de chaque composé dans un réacteur batch. Ces bilans ont été comptabilisés avec les périodes dynamiques du réacteur, le réchauffement et le refroidissement. Le réseau réactionnel de chaque composé modèle a été créé à partir du réactif et des composés à plus grand sélectivité. Comme déjà mentionné, le réseau réactionnel du gaïacol est associé aux réactions de déméthoxylation, déméthylation, transalkylation, hydrodésoxygénation et saturation, tandis que le réseau réactionnel du furfural correspond non seulement aux réactions d'hydrodésoxygénation et saturation, mais également à la décarbonylation, à l'hydrogénation du carbonyle et à l'ouverture du cycle. En première approche, on a supposé que toutes les réactions étaient de pseudo-premier ordre et suivaient la loi d'Arrhenius. Pour les réactions réversibles, comme l'aromatisation et l'oxydation du groupement hydroxyle, les constantes d'équilibre ont été extraites de la littérature et ont été supposées égales pour

toutes les réactions de la même famille. Pour le gaïacol, le modèle prédit assez bien la concentration de certains composés, tandis que d'autres sont surestimées ou sous-estimées. Les écarts constatés dans le modèle peuvent être dus à des paramètres cinétiques inadaptés, à l'absence de certaines réactions et à la non-prise en compte de l'impact de la structure de la molécule et des sites catalytiques sur la réactivité. Pour le furfural, le modèle prédit assez bien la concentration de tous les composés pour les basses températures (100 ºC et 150 ºC), mais celle-ci est surestimée ou sous-estimée à 250 ºC. Comme dans le cas du gaïacol, les écarts constatés dans le modèle peuvent être dus à des paramètres cinétiques inadaptés, à l'absence de certaines réactions et à la non-prise en compte de l'impact de la structure de la molécule et des sites catalytiques sur la réactivité. De plus, l'absence de données expérimentales joue un rôle important dans l'estimation des paramètres. Bien que les deux modèles présentent certains défauts, ils prédisent raisonnablement bien (avec un écart de 10 %) les résultats de la plupart des tests expérimentaux.

Après la création du modèle déterministe, les paramètres estimés ont été testés avec le SSA. Cette approche est capable de décrire un réseau de réactions, molécule par molécule, au travers une évolution discrète dans le temps, tout en conservant les détails moléculaires, évitant ainsi la pré-définition des voies de réaction. Pour le présent travail, le SSA a été adapté aux périodes dynamiques du réacteur et les réactions subies par le gaïacol et le furfural dans le procédé de l'hydrotraitement. Au total, pour l'hydrotraitement du gaïacol et du furfural, vingt-six réactions ont été mises en œuvre dans l'algorithme : décomposition du méthanol et vaporeformage, méthanation du monoxyde et du dioxyde de carbone (à l'équilibre), réaction du gaz à l'eau (à l'équilibre), hydrogénation du carbonyle (en équilibre), décarbonylation, déshydratation, déméthoxylation, déméthylation, hydrodésoxygénation, méthylation, transalkylation, hydrogénation des insaturés, hydratation et ouverture de cycles. Pour les simulations avec les paramètres déterministes, les deux composés du modèle prédisent des tendances similaires pour les composés majoritaires. Les déviations observées peuvent provenir de la constante d'équilibre de l'aromatisation et l'oxydation du groupement hydroxyle, ainsi que de la présence d'autres molécules, qui ne sont pas présentes dans le modèle déterministe. Pour la génération des modèles cinétiques via le SSA, les paramètres déterministes ont été utilisés comme un point de départ. Pour le gaïacol, le SSA a été capable de prédire les concentrations des composés majoritaires, alors que les molécules restantes sont surestimées ou sous-estimées. Les écarts constatés dans le modèle peuvent être dus à des paramètres cinétiques inadaptés, à l'absence de réactions et à la non-prise en compte de l'impact de la structure de la molécule et des sites catalytiques sur la réactivité. Pour le furfural, le SSA prédit les concentrations de la plupart des composés à basse température. À des températures plus élevées, le modèle ne peut prédire la composition, probablement en raison de la présence de réactions qui ne sont pas prises en compte ni dans le modèle déterministe ni dans le modèle stochastique (par exemple, hydratation). Pour les deux composés modèles, il a été supposé que les valeurs d'énergie d'activation similaires à celles

utilisées dans les modèles déterministes, tandis que les facteurs pré-exponentiels ont été réajustés via le « Stochastic Simulation Algoritm ».

Les paramètres cinétiques ajustés via le SSA ont été appliqués pour la simulation de l'hydrotraitement de bio-huiles. Cependant, l'approche stochastique nécessitait une bibliothèque de molécules représentatives des fractions de bio-huiles. Cette bibliothèque a été obtenue par une technique de reconstruction moléculaire. Cette méthode consiste en une première étape, la reconstruction stochastique ou SR (pour « Stochastic Reconstruction »), qui génère une bibliothèque équimolaire de molécules conforme aux données expérimentales utilisées pour la reconstruction. La fraction molaire des molécules dans la bibliothèque est ensuite ajustée aux propriétés du mélange via une deuxième étape, la maximisation d'entropie ou REM (pour « Reconstruction by Entropy Maximisation »). Avant l'application de la méthode, de nouvelles structures moléculaires, de nouvelles familles de ^{13}C RMN et une nouvelle corrélation pour l'analyse SEC ont été implémentées dans le code source, parallèlement à la création d'une fonctionnalité conçue pour reconnaitre les familles des groupes fonctionnels oxygénés. La reconstruction a été appliquée à deux échantillons de bio-huile (A et B) avec deux diagrammes de construction différents. Les modèles compositionnels ont été générés grâce à l'analyse élémentaire, à l'analyse ^{13}C RMN et à l'analyse SEC. Pour la première approche de reconstruction (BD#1), le modèle a estimé de manière correcte l'analyse élémentaire, l'analyse ^{13}C RMN et la fraction initiale de l'analyse SEC. Les points expérimentaux de cette analyse ont présenté des déviations, principalement associées à l'absence de composés de poids moléculaire élevé dans les bibliothèques. Pour la deuxième approche de reconstruction (BD#2), les résultats étaient similaires à ceux du BD#1. En ce qui concerne la validation des modèles, les deux prédisaient des densités proches des valeurs attendues, mais contenaient peu d'oxygène par molécule. Pour cette raison, la REM n'a pas été appliquée, car la méthode surestimerait les fractions molaires de certaines molécules et en sous-estimerait d'autres, ce qui n'est pas souhaitable. En conclusion, bien que les deux BD présentent certains manquements, les modèles compositionnels créés ici peuvent reconstruire des bibliothèques capables de représenter la fraction légère et connue de bio-huiles. Entre les deux BD, il a été constaté que le BD#2 présentait moins d'écarts et pouvait mieux estimer les propriétés que le BD#1. Pour cette raison, les résultats issus de la BD#2 ont été appliqués à la simulation de l'hydrotraitement de bio-huiles.

Finalement, le mélange synthétique a été utilisé dans la simulation stochastique de l'hydrotraitement de bio-huiles, étayée par la cinétique des molécules modèles. Le simulateur a prédit correctement les composés élémentaires et les fonctions chimiques du carbone, sauf les carbones associés aux fonctions oxygénées. En plus, le SSA a prédit une performance HDO extrêmement faible, qui ne correspond pas aux données expérimentales. Les raisons de ces écarts peuvent être dues à une cinétique inadaptée, à l'absence de certaines réactions et également à des écarts expérimentaux.

En résumé, le présent travail a démontré que, grâce à une technique de reconstruction moléculaire et à une base cinétique sur la réactivité des composés oxygénés, il est possible de simuler au niveau moléculaire l'hydrotraitement de bio-huiles. Cependant, pour obtenir une prévision correcte du comportement d'une bio-huile au cours du procédé, une bibliothèque de molécules appropriée et un modèle cinétique correct doivent être disponibles. Pour les deux, il est nécessaire de s'appuyer sur la littérature et sur une solide banque d'informations expérimentales. Dans le cas présent, les deux contraintes étaient extrêmement fortes. Cependant, il a été possible de poser les bases d'une approche générique, de développer les outils informatiques et de générer des résultats significatifs pour la composition et la réactivité d'une bio-huile et de ses molécules modèles.

Summary

To face the dilemma caused by the increasing energy consumption, alongside with the decline of the oil resources and its harmful emissions, several eco-friendly technologies have been proposed including the transformation of lignocellulosic biomass.

One of the many thermal routes of transforming biomass into more valuable products is pyrolysis. This process is responsible for the production of a three-phase effluent, where the quantities of each phase vary according to the parameters of the procedure. When placed under a maximum of heat-rate and gas velocity extraction, the process, now designated as fast pyrolysis, generates a greater quantity of liquid phase. Termed bio-oils, fast pyrolysis liquids are complex chemical mixtures containing large amounts of organic compounds. This substance is extremely diverse in composition and has a great potential for the production of fuels and chemicals that could eventually displace the products obtained from fossil origin. However, this liquid fraction is extremely rich in oxygen, which, due to its high reactivity, makes it difficult to manage, as well as to unveil its molecular identity. Up to now, the scientific community has identified the main bio-oil families (alcohols, ketones, aldehydes, furans, phenols, carbohydrates and carboxylic acids). Moreover, fast pyrolysis liquids are also totally immiscible with hydrocarbon fractions, requiring, at least in their initial unit operations, a dedicated upgrading scheme, which is currently still at the development level.

To improve their fuel properties, bio-oils must undergo a catalytic upgrading process, which includes hydrotreating and hydroconversion, under hydrogen pressure. The process involves a first stabilization phase at a mild temperature followed by one or two hydrotreating stages at a higher thermal level. The first stage converts the more reactive components at a temperature below 300 °C, thereby avoiding the polymerization reactions, while the second and/or third stages convert the remaining molecules at a temperature above 300 °C (Venderbosch et al. 2010). In the last two steps, a phase separation can take place, depending of the amount of water present in the medium and the amount of oxygen in the partially upgraded bio-oil.

In line with the scientific community, the present thesis intends to develop a kinetic model for the upgrading process which may then be used in the maximization of the targeted products and of the oxygen removal. Furthermore, it also plans to not only expand the knowledge regarding bio-oils, but also to understand their behavior under hydrotreating conditions. To achieve these aims, a strategy is defined through the information available in the open literature concerning this subject. Besides the experimental studies of one single molecule (single model components) or mixtures of known molecules of bio-oil (multi model components), little information can be found on full-range bio-oils. Nevertheless, it is possible to extract from the material available that the reactivity of each bio-oil family is maintained through the process.

From the retrieved information, a kinetic model for the hydrotreating process may be constructed. However, since the classic kinetic modeling approaches are not suitable for complex and not well-known systems, this thesis proposes the use of an alternative technique, designated as the stochastic simulation algorithm. This procedure is based on a Monte Carlo method where, through random events, the reactions are selected and the temporal space evolves. Developed in IFPEN and based on the work of Gillespie (Gillespie 1976) (Gillespie 1992), this technique, called Stochastic Simulation Algorithm (SSA), was already validated for various light and heavy petroleum fractions (Oliveira 2013). However, the current approach misses its extension to the bio-based oxygenated mixtures. This adaptation advanced gradually, starting with single model compounds, as guaiacol and furfural in the present case, and ending with full-range bio-oils.

Be that as it may, due to the little information concerning the kinetic parameters over reduced NiMo/γ-Al_2O_3 and the absence of an optimizer in the SSA, the model compounds were firstly modeled by means of a deterministic approach. The deterministic kinetic model was generated only with the molecules quantified by GC-FID in the experimental tests. The aim was to acquire an idea of the range for the kinetic parameters, in order to accelerate the parameters estimation through the SSA for a larger reaction network. Both deterministic models were able to predict reasonably well the effluents of the hydrotreating of guaiacol and furfural. The deterministic kinetic parameters were implemented in the stochastic approach, together with the typical hydrotreating reactions of both model compounds. From here, two stochastic models were created for the hydrotreating of guaiacol and furfural. The guaiacol model had a limited prediction capacity, while the furfural model had a good prediction performance at low temperatures.

Once all the main reaction pathways and their kinetics were covered by the SSA, samples of bio-oils were used for the validation of the kinetic parameters. However, the SSA requires a description of the bio-oil at the molecular level. As mentioned before, this information is not available from current analytical techniques. Hence, the validation with bio-oil samples needed an auxiliary tool, named molecular reconstruction, to help identify the unknown molecular species in the bio-oil. Also developed in IFPEN and based on the work of Klein's group at the University of Delaware (Neurock 1992), this technique was already validated for petroleum fractions (Oliveira 2013), and now also requires the extension to bio-based mixtures.

After implementing the necessary features, as for instance the ^{13}C NMR oxygenated fractions, in the molecular reconstruction technique, a synthetic mixture of bio-oil molecules was generated thanks to the pre-existing available experimental data. Although the created synthetic mixture lacks macromolecules, the library of molecules can represent reasonably well the light fraction of bio-oil. This detailed library, together with the kinetic parameters of the model compounds, was then used in the stochastic approach to finally simulate the hydrotreating of full-range bio-oils. Through this approach,

it was possible to simulate the hydrotreating of bio-oil, while keeping a molecular description of the process.

To conclude, the present work was divided into four parts: ***Part I – Literature review, Objectives and Strategy***; ***Part II – Experimental approach***, ***Part III – Kinetic modeling of model compounds***; ***Part IV – Hydrotreating of bio-oils***.

- ***Part I*** contains an **Introduction**, which highlights the context and importance of the thesis subject, together with the main scientific questions behind the subject. This part also includes chapter **I. State of the Art** and chapter **II. Objectives and Strategy**. The first chapter, **State of the Art**, briefly summarizes the context and the scientific work found in the open literature concerning the subject. The aim of this study is to answer the questions posed in the introduction and select a suitable approach to reply to the main objective of the present work. The following chapter, **Objectives and Strategy**, summarizes not only the main objective, but also the strategy employed to fulfill this objective, according to the selected methodology.
- ***Part II*** contains chapter **III. Experimental methodology** and chapter **IV. Experimental results.** The third chapter, **Experimental Approach**, briefly summarizes the methodology and the experimental background of the present work, while the fourth chapter, **Experimental results,** presents all the experimental results which will be used during the present work.
- ***Part III*** contains chapter **V. Deterministic Kinetic Model**, chapter **VI. Deterministic Results**, chapter **VII. Stochastic Kinetic Model** and chapter **VIII. Stochastic results**. The fifth and seventh chapters, **Deterministic Kinetic Model** and **Stochastic Kinetic Model**, present, respectively, the methodology behind the deterministic and stochastic approaches and the creation of both models for guaiacol and furfural. The sixth and eighth chapters, **Deterministic Results** and **Stochastic Results**, report the results of both modeling techniques.
- ***Part VI*** contains chapter **IX. Molecular Reconstruction Methodology**, chapter **X. Molecular Reconstruction Results** and chapter **XI. Stochastic Simulation of Bio-oil Hydrotreating**. The ninth chapter, **Molecular Reconstruction Methodology**, describes the methodology behind the molecular reconstruction approach, while the tenth, **Molecular Reconstruction Results**, summarizes the results obtained for bio-oil through this methodology. The eleventh and last chapter, **Stochastic Simulation of Bio-oil HDT**, contains the application of the kinetic parameters and the library of molecules to the simulation of the hydrotreating of bio-oil.

To finish, the manuscript includes as well a ***General Conclusion***, which summarizes the main conclusions of the thesis and the ***Perspectives***, where ideas to improve the present work are proposed. The ***References*** may be found in the end of the document, alongside with the ***Glossary***.

Table of Contents

PART I: Literature review, objectives and strategy

"Biomass is considered the renewable energy source with the highest potential to contribute to the energy needs of modern society for both the developed and developing economies world-wide"
(Bridgwater 2003)

Table of Contents

Introduction

To face the actual energetic dilemma, eco-friendly technologies have been studied, as renewable energies and treatment of fossil fuels emissions. From the array of renewable energies, biomass takes the first role, as it presents a high energetic potential. This source, like petroleum fractions, is a carbon-based reserve. It can therefore be transformed into chemical building blocks and fuels that are important for mankind. In order to achieve this goal, biomass must undergo several processes, from pretreatment to conversion into intermediate compounds such as bio-char, gases and bio-oil.

Bio-oil is an extremely complex organic mixture that comes from the fast pyrolysis of biomass. Besides the high amount of carbon, this mixture possesses a great quantity of oxygenated species, which cause the decrease of the quality of the bio-oil upgraded products: low thermal stability above 100°C, high acidity, lower heating value, immiscibility with hydrocarbons, etc. In order to incorporate these liquids into transportation fuels, bio-oils must be converted and purified into chemical structures that comply with the strict fuel specifications based on hydrocarbons.

The upgrading of bio-oil can be done through a dedicated process scheme that includes a series of hydrotreating and hydroconversion steps. The performance of these techniques, in particular of hydrotreating, and the quality of the effluent, is directly related to the oxygen removal. However, to achieve a proper elimination of the undesired oxygen atoms, several studies around this process must be performed. Since experimental assays in pilot units are expensive, most of the producers found ground in kinetic models, which not only express the reaction networks, but also allow a profound study of chemical system under question, all in an economic way.

Motivated by the ability of bio-oils to generate fuels and valuable chemical building blocks, the present thesis focuses on the construction of a modeling tool for the conversion of this mixture. However, since bio-oil possesses a high complexity, both in molecular and reaction terms, to unlock these shackles, three main questions regarding the subject must be clarified:

1st – How to describe a bio-oil mixture in molecular terms?

2nd – How does a bio-oil mixture reacts under hydrotreating conditions?

3th – How to create an *in silico* tool to represent the hydrotreating of bio-oil?

Through these answers, not only an appropriate methodology but also a supportive layer of knowledge regarding the complexity and reactivity of bio-oils will hopefully be obtained.

Chapter I. State of the Art

The following section will introduce the context and the subject of the present work. This section will present a brief perspective of biomass, in particular lignocellulosic biomass, and its conversion into pyrolysis oils. Afterwards, the highlights of the bio-oil characterization alongside their upgrading into valuable fuels will be discussed. Furthermore, a critical analysis of the pyrolysis oils models in the open literature will be presented. Finally, this section will also contain the background on complex modeling methodologies, from the molecular description all the way to the simulation of reactions.

I.1 Biomass to bio-oil

As highlighted in the introduction, biomass and its derivatives may represent not only a new energetic hope for the future, but also a strong base for the renewable production of products that are essential to our society. As a promise of balance between the development of the society and the protection of the earth's ecosystem, this feedstock has become more and more a topic of research. The present work is focused on the use of biomass as an energy source, and more precisely on the upgrading of fast pyrolysis bio-oils from lignocellulosic biomass in order to produce fuels. Therefore, biomass will be the first topic of discussion presented in this chapter, together with a brief description of its characterization and technological development.

I.1.1 Biomass

According to the International Energy Association, 'biomass […] is any organic, i.e. decomposable, matter derived from plants or animals available on a renewable basis.' (IEA). In the energy landscape, it is a 'green' energy source that is organic and renewable. This type of energy has both advantages and disadvantages. Overall, biomass and its derivatives, besides being renewable sources, are easier to access and store when compared to the other renewable energies that, once transformed into energy, need to be use immediately. Nevertheless, the searches for biomass resources promote deforestation and above all increase the competition between its production fields and those used for food production. Another concern is the relatively low energy density of the resource when compared to fossil fuels. Furthermore, liquid fossil fuels with huge underground reserves can easily be transported, compared to a BtL (Biomass to Liquids) refinery that will only use solid biomass collected within a radius of approximately 200 km.

I.1.1.1 Biomass composition

Concerning its composition, biomass is constituted mainly by carbon, hydrogen and oxygen. However, it can also present minor amounts of inorganic compounds, such as nitrogen, sulfur and metals (Jong and Ommen 2015). In addition, the properties of biomass can also be affected by several factors as the

origin type, the external conditions, the age and harvesting season, the amount of fertilizers (if used), the blending strategies and, finally, the technology, transport, handling and storage (Jong and Ommen 2015). When taking into account this description, the biomass concept can be extended to several types of material. However, in the case of the present work, only the lignocellulosic biomass will be focused on.

Lignocellulosic biomass can be considered as an agricultural residue or forests resources. In fact, lignocellulosic material can encompass trees, barks, bushes, grasses, corn stover/cobs, sugarcane bagasse and even paper mill discards (Dahiya 2015) or lignins. Since these materials are considered as low-value products, the conversion of lignocellulosic biomass into biofuels is therefore an excellent market opportunity.

Looking at lignocellulosic material from a different perspective, this type of source is a complex mixture of carbohydrates and non-carbohydrates found in the leaves and stems of plants, which are more precisely, cellulose, hemicellulose, and lignin, as depicted in Figure I-1. It is important to mention that the fraction of these three constituents in the biomass varies significantly according to the biomass type. In Table I-1, there is a brief description of the main lignocellulose biomass components.

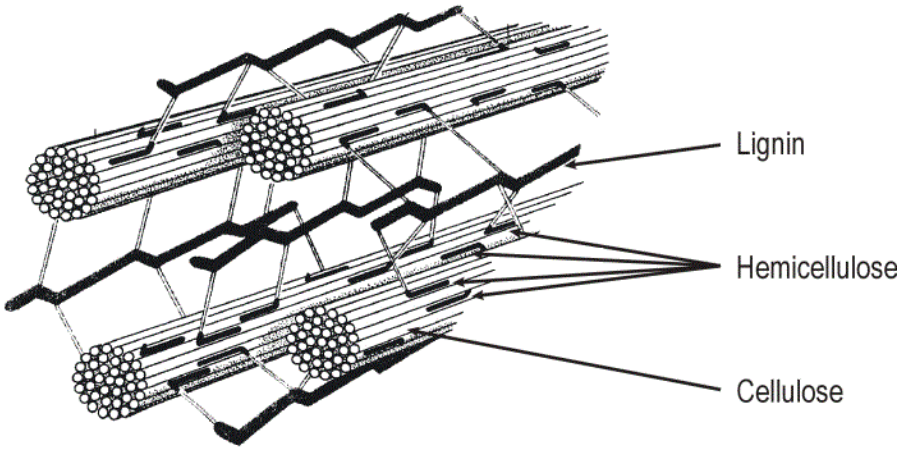
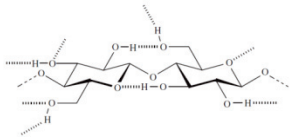
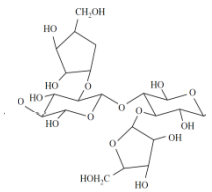
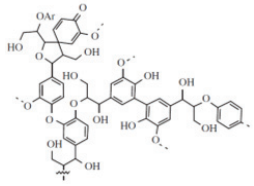


Figure I-1 - Lignocellulosic biomass structure (Gnansounou and Dauriat).

Table I-1- Lignocellulosic biomass composition (Jong and Ommen 2015b) (Jong and Ommen 2015a) (Dahiya 2015) (Dahiya 2015).

Biomass fraction	Molecular unit	Molecular structure	Observations
Cellulose 40-50 % wt dry basis	Cellobiose		Supramolecular structure due to the hydrogen bonds and the various glucose polymer chains that in a parallel configuration get twisted forming microfibrils. These characteristics strengthen the cellulose by increasing its stability and therefore reduce the probability of degradation.
Hemicellulose 25-35 % wt dry basis	Hexosans: Six carbon sugars		Due to a branched structure, this polymer has less stable intermolecular linkages. Despite this, this structure is the responsible for the protection and fixation of the cellulose microfibrils to each other and to the plant cell wall.
	Pentosans: Five carbon sugars		
Lignin 20-30 % wt dry basis	Heterogeneous compounds with different aromatic structures		Vast structure supported by ether bonds, carbon-to-carbon linkages and covalent bonds with cellulose and hemicellulose. These linkages are established by side groups, such as hydroxyl, methoxyl and carbonyl. This compact structure provides mechanical support to the plant wall, increasing the difficulty of its degradation and leading to the production of char.
Minor compounds	Organic compounds	Starch, free sugars, residual proteins, and oils and fats	Their presence reduces the quality of the bio-oil. Proteins contain nitrogen which can release NO_x/N_2O during the degradation processes. Hence, the biomass usually demands a pre-treatment procedure.
	Inorganic compounds	Nutrients: nitrogen, phosphorus, potassium, calcium, magnesium, sulfur, iron, zinc, chlorine, copper, etc.	Besides being responsible for growth and maintenance of the plants, they modify the lignin product properties and cause a decrease in the yield of sugars from cellulose, producing more char. Once again, this implies the need for a pre-treatment procedure of the biomass.

I.1.2 Conversion processes

Thanks to the organic composition of biomass, production of valuable resources to society, such as energy, fuels and chemicals, is possible. To reach these products, the raw material needs to undergo a conversion, which may have a biochemical, chemical or thermochemical nature, or any combination of the above. Since the conversion of biomass to fuels and chemicals is not straightforward, this usually requires a complex multi-step pathway. However, whatever the chosen pathway, biomass first undergoes a pre-treatment step in order to prepare the material in the most suitable way for the downstream steps of the conversion process.

I.1.2.1 Pretreatment

The pretreatment step aims to remove undesirable compounds that affect the product qualities. In general, these procedures degrade the macromolecules into smaller compounds, like sugars, and they also separate the lignin and other plant constituents from fermentable materials. However, the molecular structure of the biomass being pretreated needs to be taken into consideration, as the pretreatment procedure varies according to its characteristics.

I.1.2.2 Biochemical processes

The biochemical process is based on enzymes, on bacteria or on other microorganisms. These are the responsible for the conversion of biomass into more valuable products. Each conversion process has its methodology, enzymes or microorganisms, and its specific products. In the case of aerobic digestion, the effluents are composed of carbon dioxide and water, while, for anaerobic digestion, the products consist in a fertilizer and carbon dioxide or methane gas. In fermentation, biomass is converted into ethanol, carbon dioxide and waste. Afterwards, these intermediate compounds are used for the production of gaseous or liquid fuels or even to generate energy.

I.1.2.3 Chemical processes

The chemical processes imply the use of chemical agents to transform biomass into fuels. Through the addition of hydrogen, acids or alcohols, biomass is converted into hydrocarbons or other compounds, equally valuable for mankind. As in the previous process, these intermediate compounds are then used in the society as a source of energy.

I.1.2.4 Thermochemical processes

Thermochemical processes consist in the conversion the organic matter by inflecting heat to the system. These processes encompass gasification, combustion, and pyrolysis. Gasification is the conversion of solid biomass into various gases using heat and varying amounts of oxygen. Combustion consists in

decomposition the biomass in the absence of oxygen and at low temperatures. Pyrolysis is the decomposition of biomass at high temperatures in the absence of oxygen.

Combustion

As mentioned above, combustion consists in decomposition of biomass in the absence of oxygen at low temperatures. It is a widely commercially process that occurs in a 450 °C – 2000 °C (Dahiya 2015) temperature range and which can provide heat and power while burning biomass. In here, all compounds are transformed into carbon dioxide and monoxide, water, ash and energy.

When the produced wastes are recycled into the process, combustion emerges in the energetic market as a feasible possibility. However, once converted into energy it is impossible to store, implying an immediate use. Besides, this process not only has a low efficiency varying in a range of 15 – 30 % depending strongly on the plant size and age, but also is the responsible for emissions and ash production, as mentioned in the first paragraph (Bridgwater 2003).

Gasification

The gasification process transforms all the feedstock into carbon dioxide, water, carbon monoxide, methane and hydrogen in a temperature range of 700 °C – 850 °C (Dahiya 2015) and in the presence of oxygen. This gaseous effluent is also known as syngas.

Unlike combustion, this method proceeds in several steps: first, the moisture present in the feedstock evaporates, after which the matter degrades through pyrolysis producing gas, oils and char, which will all be *a posteriori* reduced to gas by partial oxidation. The composition of this gas depends on the feed composition, water content, reaction temperature, and extent of oxidation of the pyrolysis. Nevertheless, not all the products after pyrolysis are transformed into gas. Contaminant tars in the final gas are difficult to remove, reducing therefore its quality. Furthermore, the gas generated is extremely difficult to transform and, as well as combustion, impossible to store implying its immediate use. In summary, if this process is not well adjusted in terms of quality and does not possess a straight utilization, it will lose field in the energy market.

Pyrolysis

In the pyrolysis process, with temperatures above the 300°C and less than 650°C (Dahiya 2015), the lignin, cellulose, hemicellulose and other residual organic molecules segregate into smaller structures in the absence of oxygen. The process effluent is characterized by a solid, liquid and gas phase. The solid product is known as bio-char and is composed by the macromolecules undestroyed during the process, degradation products and recombination of molecules pre-existing in biomass. The liquid and gas phase are in fact a mixture of gases that are released during the pyrolysis of biomass. After condensation, the product molecules are segregated into a liquid and a gas phase, as the non-volatile vapors form an oil,

termed bio-oil, while the remaining volatile vapors form the designated gas phase. According to the operating conditions and technology strategy, the pyrolysis process can be considered slow, intermediate or fast, as shown in Table I-2. Furthermore, pyrolysis can be performed in the presence of a catalyst (catalytic pyrolysis) or reactive gases, as hydrogen (hydropyrolysis). Both processes have a higher operational complexity than traditional pyrolysis. However, both can also produce high quality bio-oil (less reactivity), or even hydrocarbons, from biomass (Wang et al. 2013).

According to the operatory conditions and technology strategy, the pyrolysis process can be considered slow, intermediate or fast. The first process possesses the lower temperature range and the higher vapor residence time. By increasing the temperature and reducing the residence time, the mode passes to the intermediate pyrolysis. When the process is operated at high temperatures and very low residence times, it becomes known as fast pyrolysis. These two characteristics are responsible for the variation in the composition of the effluents, which give the process the flexibility to produce of the most desired product, as shown in Table I-2.

Table I-2 – Typical product yields from different modes of pyrolysis (Bridgwater 2007).

Mode	Solid	Liquid	Gas
Fast	12 %	75 %	13 %
Intermediate	20 %	50 %	30 %
Slow	35 %	30 %	35 %

As highlighted in the previous paragraph, the intermediate pyrolysis is a mode which possesses characteristics that are in between those of the fast mode and the slow mode. Therefore, only the latter two will be focused on in this section. Nonetheless, it is important to always remember that the characteristics of intermediate pyrolysis are placed in between those of the fast and the slow mode. Regarding these two modes of pyrolysis, they vary not only in terms of heat transfer, as for the temperature, heating rate and particle size, but also in terms of mass transfer, as for the residence time of the vapors. These differences are shown in Table I-3.

When bio-char is the desirable product, the operating conditions are adjusted to match the slow pyrolysis process. As the name states, this process has a higher residence time of vapor in the main reactor together with a slower heating rate. Under these conditions, biomass suffers a slow burning procedure that leads to an increase of the char. Although this product possesses several applications, its low density enhances the cost of its transport, which may have a negative economic impact on the overall slow pyrolysis process. For the fast pyrolysis method, the vapor residence time is much lower, but the heating rate is typically six hundred times higher than under slow pyrolysis conditions (Dahiya 2015). Therefore, the biomass burns more rapidly improving the bio-oil production.

Table I-3 - Comparison between slow and fast pyrolysis (Dahiya 2015).

Characteristics	Slow Pyrolysis	Fast Pyrolysis
Particles size	> 2 mm	< 2 mm
Temperature	400 °C – 600 °C	400 °C – 650 °C
Vapor residence time	5 min – 30 min	≈ 2 s
Heating rate	< 10 °C/min	≈ 100 °C/s

To finalize this comparison, it is important to highlight the size of the particles of the heat transfer medium, which is usually sand. As mentioned before, the heating rate is the main key for the fast pyrolysis process, which implies an efficient heat transfer. To achieve this aim, the sand particles require a small diameter to ensure that the entire particle, even its center, is hot during the process. If large particles were used, they might be cold in the interior leading to a decrease in the heat transfer efficiency between the biomass and the sand. In the case of the slow pyrolysis process, the particle diameter does not have a wide impact as the heating rate is extremely low compared to fast pyrolysis.

In present work, the bio-oils under question come from the fast pyrolysis technology, as this is the mode that has a higher production of liquid. This mixture is extremely interesting for the energy market, especially because it can generate power and fuels with a high efficiency.

I.1.3 Bio-oil

Bio-oil is normally a homogeneous mixture, but when the water content is above 35-40wt%, a multi-phase system may appear (Bridgwater 2007). Excluding the water, bio-oil is composed mostly by oxygen and carbon, approximately in similar amounts. Additionally, it possesses less than 10 % of hydrogen, with some traces of nitrogen and other elements (Bridgwater 2003). It should be mentioned, however, that the composition and characteristics of bio-oils also depend on the origin of the biomass and on the pyrolysis conditions.

As a result of its characteristics, this complex substance is an unstable mixture constituted by a vast number of different molecules with diversified characteristics. Consequently, it is a product that is extremely difficult to analyze, even with the use of powerful analytical techniques. In this section, more details will be given regarding bio-oils characteristics and analytical techniques. More information can be found in the articles of Staš et al. (Staš et al. 2014) and Hao et al. (Hao et al. 2016).

I.1.3.1 Bio-oil properties

Bio-oils are extremely complex mixtures composed of widely varying molecule types. These are highly different from each other and are associated to different structures from which they originate, such as cellulose, hemicellulose and lignin. Therefore, bio-oils possess unique characteristics and properties, which turn it into an incredible flexible material when treated properly. Bio-oils are considered as a micro-emulsion, meaning that have a continuous phase, which is an aqueous solution of mainly cellulose and hemicellulose decomposition products, and a discontinuous phase, constituted of mainly of pyrolytic lignin macromolecules. Apart from these characteristics, bio-oils include a huge amount of oxygen, not only incorporated in its molecules (phenols, alcohols, etc.), but also due to the amount of water present in the medium.

Table I-4 - Bio-oils physicochemical properties (Bridgwater 2003).

Physical Property	Moisture content	pH	Elemental analysis				
			C	H	O	N	Ash
Typical Value	15 – 30 %	2.5	55 – 58 %	5.5 – 7.0 %	35 – 40 %	0 – 0.2 %	0 – 0.2 %

Table I-4 - Bio-oils physicochemical properties (Bridgwater 2003). [Continuation]

Physical Property	Specific gravity	Higher heating value	Dynamic Viscosity at 40 °C (with 25 % water)	Solids
Typical Value	1.20	16 – 19 MJ/kg	40 – 100 cP	< 1 %

Regarding the physical properties of bio-oils, these liquids present an acidic character and a higher density than crude oils (about 1.2). Viscosity also varies strongly with the presence of water, with temperature and with storage time. Furthermore, it is important to highlight that the variations in pyrolysis temperature also promote the production of undesirable solids and high molecular weights compounds. In terms of color, the appearance of the bio-oils changes with the presence of char and the chemical composition, but it is usually a very dark (brown to black) liquid. A brief summary of the bio-oil properties is given in Table I-4.

I.1.3.2 Bio-oil composition

Regarding its derivation, it is impossible to assemble the molecules of the pyrolytic oil and track them back to their molecule of origin. However, it is possible to establish a source path for some compounds, when they share characteristics with cellulose, hemicellulose or lignin. According to the work of Patwardhan (Patwardhan 2010), cellulose is mostly responsible for the carbohydrates, but also for the alcohols, carboxylic acids, furans, aldehydes and ketones. Furthermore, the author noted that the latter

five components are also products of hemicellulose. Regarding the lignin, the author established that aromatic compounds, such as phenolic and methoxy-phenolic compounds, are its main products.

As a simple assessment of bio-oil composition, the compounds were regrouped into bio-families according to their characteristics, in particular according to their chemical functions, as presented in Table I-5 (Staš et al. 2014). It should be stressed that the shown composition is based on what can be analyzed in the bio-oil, not accounting for the (unanalyzed) heavy and polar compounds. Therefore, Table I-5 is only a partial picture of the analyzed fraction of bio-oil, and does not represent the entire composition.

Table I-5 - Typical distribution of the bio-oils major compound classes (Staš et al. 2014).

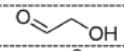
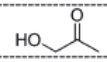
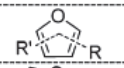
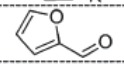
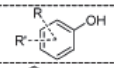
Bio-oils Families	Aldehydes	Alcohols	Carbohydrates	Carboxylic acids	Furans
Fraction *w*%	10 – 20	2 – 5	20 – 35	4 -15	1 – 4

Table I-5 - Typical distribution of the bio-oils major compound classes (Staš et al. 2014). [Continuation]

Bio-oils Families	Ketones	Phenolic monomers	Phenolic oligomers	Water
Fraction *w*%	1 – 5	2 – 5	15 – 30	20 – 30

In molecular terms, bio-oil contains a vast number of distinct oxygenated molecules, as shown in Table I-6. Among the low molecular weight oxygen species there are acids, alcohols, aldehydes, ketones and some esters, mostly originated from reaction between phenolic and alcohols compounds (González et al. 2016). Also, higher molecular weight oxygenated molecules, such as sugars, furans, phenolics, can be found in bio-oil in addition to miscellaneous molecules, which represent somehow the diversity of bio-oil by groups several functional groups in the same compound. Finally, high molecular species also exist in bio-oil. These have different origins since they are not totally degraded molecules of cellulose, hemicellulose and lignin to the new molecules formed during the pyrolysis process, as seen by Bayerbach and Meier (Bayerbach and Meier 2009). These authors proposed that pyrolytic lignin "(…) emerges as results of both recombination reactions [, as Diels-Alder,] and thermal ejection of intact lignin fragments.".

Table I-6 – Major compounds in typical wood pyrolysis oils (Wang et al. 2013).

Catalog		Typical Compounds	Structure[b]
Water		Water	H_2O
Simple oxygenates	Acids	acetic acid, formic acid, propanoic acid, methyl-propanoic acid, butanoic acid, methyl butanoic acid, pentanoic acid, glycolic acid, etc.	
	Esters	methyl acetate, methyl formate, methyl propionate, butyrolactone, dimethyl pentanedioate; etc.	
	Alcohols	butane-2,3-diol, methanol, ethanol, 2-propene-1-ol, isobutanol; ethylene glycol etc.	
	Ketones	acetone, 2-butanone, 2,3-butanedione; cyclopentanone; Methylcyclopentanone, 2-pentanone, cyclohexanone; 3-hexanone, methylcyclohexaneone, cyclopentanedione; etc.	
	Aldehydes	2-butenal, glyoxal, formaldehyde; acetaldehyde; 2-propenal; 2-methyl-2-butenal, pentanal, benzaldehyde; etc.	
Miscellaneous oxygenates		glycolaldehyde	
		1-hydroxyl-2-propanone	
		methyl 2-oxopropanoate; 1-hydroxy-2-butanone; etc.	
Sugars		Levoglucosan	
		glucose, fructose, D-xylose, D-arobinose, etc.	
Furans		furan; methylfuran; dimethyl-furan;	
		furfural	
		2-furanone, 4-methyl-2-furanone, furfuryl alcohol, 2-acetyl furan; 5-methyl furfural; etc.	
Hydro-carbons	Alkene	hexane, methyl-propene, alkylated cyclohexene, etc.	
	Aromatics	toluene; xylene, ethylbenzene, naphthalene, etc.	
Phenolics	Phenols	phenol, methylphenol, dimethylphenol	
	Anisoles	anisole	
		methyl-anisole, ethyl-anisole, etc.	
	Catechols	Catechol	
		methyl-catechol, ethyl-catechol, methoxycatechol, etc.	
	Guaiacols	guaiacol	
		methyl-guaiacol, ethyl-guaiacol, eugenol, etc.	
	Syringols	syringol, methyl-syringol, ethyl-syringol, propyl-syringol, 4-propenylsyringol, etc.	
	Others	vanillin, vanillic acid, sinapaldehyde, syringaldehyde, acetosyringone, etc.	
High-molecular-weight species		Dimmer, trimmer, and oligomer of cellulose, hemicellulose, and lignin pyrolysis products.	

I.1.3.3 Analytical characterization techniques

In a wider perspective, Oasmaa and Meier (Oasmaa and Meier 2002) illustrated the difficulty to analyze bio-oils by explaining that a bio-oil is globally composed by around about 35 % of volatiles, analyzable by gas chromatography-mass spectrometry (GC-MS), 15 % of non-volatiles, often analyzed by high-performance liquid chromatography (HPLC), 25 % of pyrolytic lignin (a water insoluble fraction) and

25 % water, as exemplified in Figure I-2. Due to its high complexity, bio-oils require powerful analytical techniques and even combinations of several techniques to access to their molecular information (see Table I-8). Moreover, to achieve more detail information, preparation methods must be also applied (see Table I-9) as it was described by Staš et al. (Staš et al. 2014).

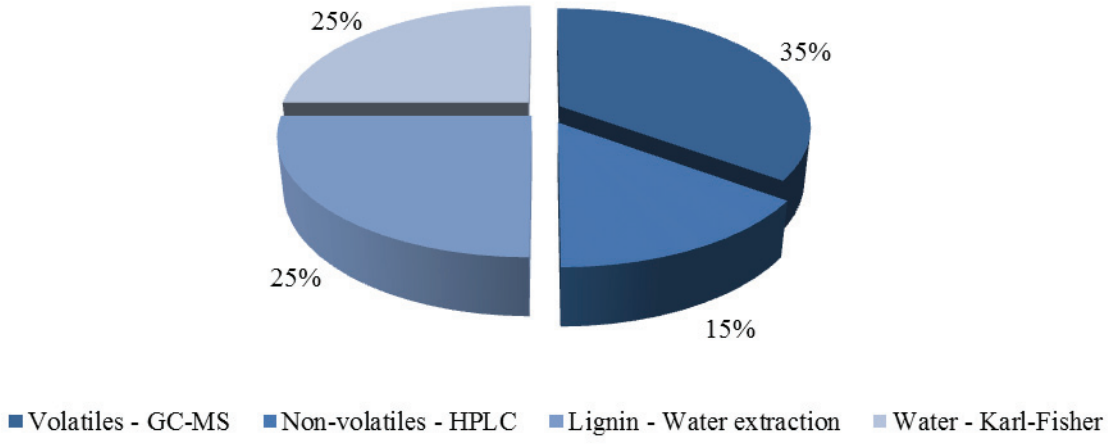


Figure I-2 - Methodology of the characterization of bio-oils (Oasmaa and Meier 2002)

From the many analytical techniques applied to bio-oil, the following section will only focus on the most used ones, as Gas Chromatography (GC), High Resolution Mass Spectrometry (HRMS), Nuclear Magnetic Resonance (NMR), and Size Exclusion Chromatography (SEC) (designated Gel Permeation Chromatograph (GPC) in Table I-9).

Regarding GC, the apparatus may be coupled with an instrument for identification, as Mass Spectroscopy (MS) or quantification, as flame Ionization Detector (FID). Thanks to both, it is possible to identify and quantified several compounds in a mixture. According to the review of Staš et al. (Staš et al. 2014), when grouping all information retrieved from the open literature, both GC and (GC × GC) should allow the identification and quantification of more than three hundred molecules.

The application of HRMS to bio-oils has been increasing all over the years. In the open literature, three main mass spectrometry techniques can be found: FT-ICR (Fourier-transform ion cyclotron resonance), OrbiTrap and ToF. (Time of Flight). The first method is the most powerful mass spectrometers currently available (Staš et al. 2014). Besides different MS, HRMS is associated with several ionization techniques, as shown in Table I-8. The negative electrospray ionization (ESI) is the most applied to bio-oils, as the majority of the compounds are polar (Staš et al. 2017).

Concerning the NMR, the most applied to bio-oil is ^{13}C NMR. This analysis has the advantage of covering almost a complete bio-oil sample. Through the identification of the carbon atom type, as

illustrated in Table I-7, this analysis can quantify the different chemical functional groups present in bio-oils.

Table I-7 - 13C NMR chemical shifts and results for different types of bio-oils (Staš et al. 2014)

Carbon Atom Type	Aliphatic carbons	Carbons connected to hydroxyls and ethers	Aromatic carbons	Carboxyl	Carbonyl
Chemical shift (ppm)	0 – 55	55 – 95	95 – 165	165 – 180	180 – 215

The SEC analysis, in particular GPC, liquid chromatographic technique used to separate molecules according to their molecular size, more specifically, according to their hydrodynamic volume (Bayerbach et al. 2006). Usually SEC chromatography is coupled with a refraction index (RI) detector and/or a UV detector. Commonly, linear polystyrene standards are used to correlate retention volumes with known molecular weights. Resulting chromatograms consist in a plot of normalized detector signal intensity versus polystyrene equivalent molecular weight (see Figure I-3). In the end, this technique results in a molar mass distribution of the sample, but in polystyrene equivalent molecular weight ($MW_{eq}PS$), which is different from the true molecular weight of the compounds in the sample.

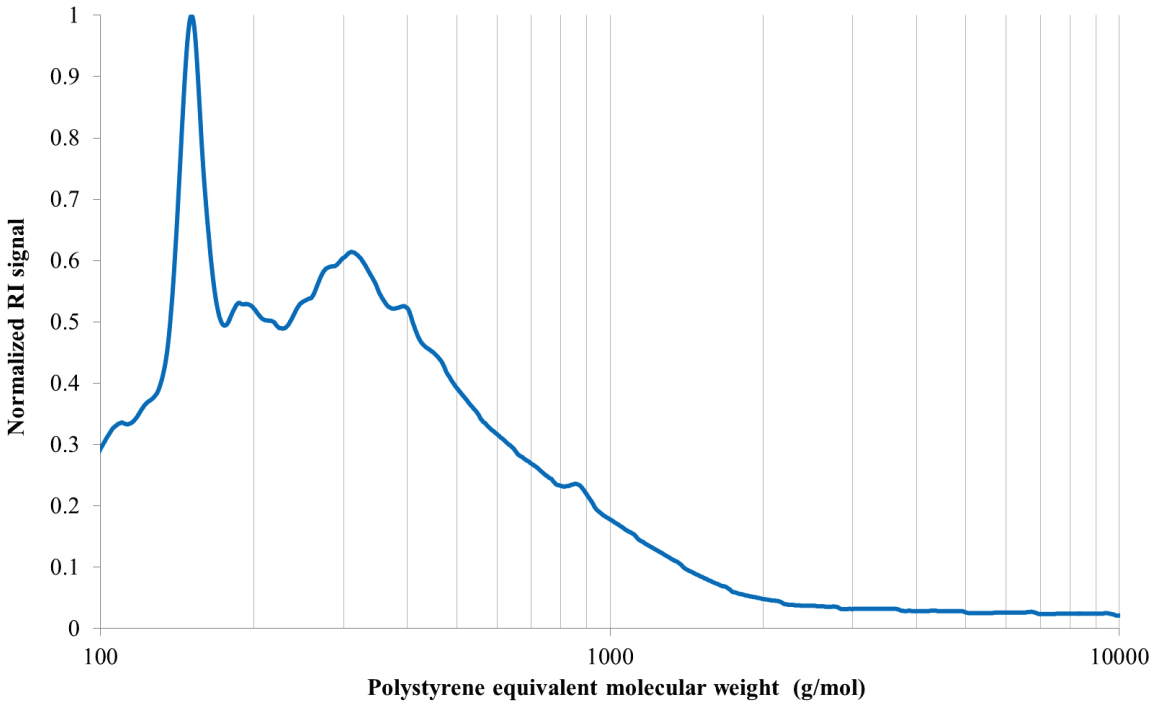


Figure I-3 - Normalized SEC analysis for bio-oil (Ozagac 2016).

Table I-8 - Characterization techniques of bio-oils (Staš et al. 2014).

Characterization technique	Method	Function	Disadvantages	Advantages	References
Gas Chromatography (GC)	Conventional GC	Extraction of information about the chemical structure	Unable to characterize nonvolatile compounds; Insufficient chromatography resolution, peak co-elution; Lack of analytical standards	Acceptable price and wide availability	(Oasmaa and Meier 2002); (Bertero et al. 2012); (Kantarelis et al. 2013); (Branca et al. 2003); (Torri and Fabbri 2009); (Evans and Milne 1987); (Liu et al. 2012); (Lu et al. 2008); (Qiang et al. 2009); (Faix et al. 1990); (Achladas 1991); (Murwanashyaka et al. 2001); (Sipilä et al. 1998); (Faix et al. 1991)
	Comprehensive two-dimensional GC	Characterization by boiling point and after, polarity	High purchase price; Only until boiling point up to 400 °C	Detection of more compounds than GC; Great theoretical peak capacity; Enhance the chromatography resolution	(Fullana et al. 2005); (Marsman et al. 2008); (Marsman et al. 2007); (Sfetsas et al. 2011); (Djokic et al. 2012); (Tessarolo et al. 2013); (Migliorini et al. 2013); (Omais et al. 2013)
Liquid Chromatography (LC)	Adsorption Chromatography	Differentiation by polarity	Unable to detect heavy fractions of bio-oil	(-)	(Pütün et al. 1999); (Özçimen and Karaosmanoğlu 2004);
	Gel Permeation Chromatography	Differentiation by molecular size			(Harvey 2000)
	HPLC	Identification of acids, aldehydes, sugars and volatile compounds			(Christensen et al. 2011)

Table I-8 - Characterization techniques of bio-oils (Staš et al. 2014). [Continuation]

Characterization technique	Method	Function	Disadvantages	Advantages	References
High Resolution Mass Spec* (HRMS)	Negative Ion Laser-Desorption Ionization	Identification of elemental compositions by measure of its mass	High purchase price; Large space requirements; Discrimination of some compounds in the ionization process	The addition of an ionization technique, detailed in the method column, improves the performance of the technique	(Liu et al. 2012); (Smith and Lee 2010); (Smith et al. 2012); (Jarvis et al. 2012); (Olcese et al. 2013); (Gross 2010); (Hoffmann and Stroobant 2007)
	Electrospray Ionization				
	Atmospheric Pressure Chemical Ionization				
	Atmospheric Pressure Photoionization				
Nuclear Magnetic Resonance (NMR)	^{1}H NMR	Determination of the amount of hydrogen atoms	(-)	Characterization of almost all the bio-oils molecules	(Mullen et al. 2009); (Strahan et al. 2011); (Joseph et al. 2010)
	^{13}C NMR	Determination of the amount of carbon atoms			
	^{31}P NMR	Quantification of hydroxyl and carboxyl functional groups			
Fourier Transform Infrared Spectroscopy (FTIR)	(-)	Characterization of molecules by infrared spectrum of absorption or emission	(-)	Simple; Low purchase; Fast	(Lu et al. 2008); (Pütün et al. 1999); (Lievens et al. 2011); (Scholze and Meier 2001); (Lu et al. 2011); (Özbay et al. 2008); (Xu et al. 2010); (Hilten and Das 2010); (Tripathi et al. 2009)

* Associated with ionization methods referred in the "Method" column.

Table I-9 - Pretreatment of bio-oil samples (Staš et al. 2014).

Method	Experimental technique		Function	Characterization technique	References
Solvent	Adsorption Chromatography		Differentiation of molecules by polarity	Gas Chromatography (GC)	(Pütün et al. 1999); (Özçimen and Karaosmanoğlu 2004)
	Solvent Extraction	Liquid-Liquid Extraction	Differentiation of molecules by water solubility	Gas Chromatography – Mass Spectrometry (GC-MS)	(Oasmaa et al. 2003); (Sipilä et al. 1998)
		Supercritical Fluid Extraction	Extraction of low-polar compounds		(Rout et al. 2009); (Wang et al. 2010); (Wang 2013)
	Gel Permeation Chromatography		Differentiation of molecules by size	Multidimensional Liquid Chromatography (MLC)	(Andersson et al. 2000)
	Sample Derivatization		Adapts the analyte to the characterization technique by increasing/decreasing the volatility, reducing the adsorption of the analyte and improving the detector response, peak separation and peak symmetry.	GC High Performance Liquid Chromatography (HPLC)	(Tessini et al. 2012); (Christensen et al. 2011); (Orata 2012)
	Solid Phase Extraction		Purification of the sample and selection of the desired molecules	(Only cleanup proposes)	(Mahinpey et al. 2009); (Poole 2003)
Solvent--free	Solid Phase Microextraction		Analysis of volatile and semi-volatile compounds	GC	(Oasmaa and Meier 2002); (Torri and Fabbri 2009); (Busetto et al. 2011); (Prosen and Zupančič-Kralj 1999)
	Molecular Distillation		Separation and purification of thermal unstable chemicals.	GC-MS	(Wang 2013); (Guo et al. 2010a); (Guo et al. 2010b); (Wang et al. 2009)

I.1.4 Conclusion of biomass to bio-oil

Biomass, in particular lignocellulosic biomass, is a useful source for the production of energy. However, to arrive at the desired product, biomass needs to be pretreated and subsequently converted. Although there are several pretreatment and transformation processes, the present work focuses on the conversion through fast pyrolysis, which in general generates a solid, a liquid and a gas in the absence of oxygen at high temperatures (450-550°C) and short vapor residence times (1-2 s).

The liquid product, normally designated as bio-oil, is a unique mixture with a vast number of oxygenated molecules, and consequently, a great diversity of characteristics. These molecules are generally classified into families as alcohols, aldehydes, carbohydrates, carboxylic acids, furans, ketones, phenols and methoxy-phenols. Due to the undesirably high amount of oxygen contained in these molecules, bio-oils are incompatible with conventional fuels composed of hydrocarbons, are chemically instable, and are extremely difficult to characterize. Nonetheless, several analytical techniques or combinations of such, as FT-ICR/MS and GC × GC, are being used with some success in the partial characterization of these mixtures.

I.2 Bio-oil to hydrocarbons

Due to its incompatibility with conventional fuels, its solids content, its high density, its high acid content, its high oxygen content, its low NHV (Net Heating Value) and its chemical and thermal instability (Bridgwater 2003), bio-oils require a dedicated and complex upgrading process. Between conversion processes as hydrotreating, zeolite upgrading, aqueous-phase processing, condensation and fractionation, hydrotreating is the most common route to upgrade bio-oils (Wang et al. 2013). Studied for several years, hydrotreating started being applied to bio-oils using the same conditions as the hydrodesulfurization processes used in petroleum refineries (Melero et al. 2011). Nowadays, these hydrotreating processes and conditions have been evolving in order to match the bio-oil requirements. However, and as mentioned before, the bio-oil reactivity is extremely complex. Overall, the reaction network of a bio-oil depends greatly on the catalyst, solvent and operating conditions. Such a dependency leads to a vast range of reactivities for the molecules. For this reason, the scientific community started studying single model compounds and mixtures of these model compounds. The following section will present a brief description of the hydrotreating of bio-oils and some studies on bio-oil model compounds and their mixtures. For more information about the topic, it is recommended the reading of the following articles (Elliott 2007; Furimsky 2000; Wang et al. 2013).

I.2.1 Hydrotreating process

Hydrotreating (HDT) is an exothermic process where the feedstock is treated with hydrogen generating a new mixture which does not contain the initial impurities. In the case of bio-oils, the aim consists in

the removal of the undesirable oxygen and also in the fragmentation of the macromolecules. Typically, the hydrotreating process occurs at a high pressure, typically in the range of 35 to 170 bar, and at temperatures around 300 to 450°C in catalyst beds with sulfided CoMo or NiMo catalysts.

I.2.1.1 Catalyst

For bio-oil hydrotreating, the catalyst and the hydrotreating conditions may change in order to improve the efficiency of the process and the quality of the products. Regarding the catalyst, transition metal sulfides, as presented before, have a poor performance due to the absence of sulfur in the feed, leading to the reduction of the catalyst activity and then deactivation. Although sulfur may be added in parallel with bio-oil, this approach would contaminate the products with sulfur (Wang et al. 2013). Other catalysts may be proposed as noble metals, transition metal phosphides, transition metal nitrides and carbides, or even other non-precious metals. However, in comparison with CoMo or NiMo, these tend to be more costly, as noble metals, more sensitivity to water, as transition metal phosphides, to have less activity, as transition metal nitrides and carbides (He and Wang 2012).

Concerning the support of the catalyst, several materials as γ-Al_2O_3, silica (SiO_2), activated carbon or even TiO_2 or ZrO_2 can be used as support. The first material has a low tolerance to water, the main product of HDO. When both react, alumina is converted into boehmite, reducing the catalyst activity and later on deactivation (He and Wang 2012). For silica supports, although these can reduce the formation of macromolecules, Centeno et al. (Centeno et al. 1995) found that CoMo/SiO_2 is less active than CoMo/γ-Al_2O_3. Concerning activated carbon, although is more active than alumina and silica, the material is fragile and has micro pores, which are easily blocked by macromolecule or coke. For the zirconia and titania supports, Yakovlev et al. (Yakovlev et al. 2009) detected a weak acidity, which benefits the activity of the oxy-compounds on the support surface.

I.2.1.2 Process development

Regarding the conditions, and according to Venderbosch et al. (Venderbosch et al. 2010), HDO of bio-oil can be done through three distinct steps. The initial step reduces, at a temperature below 300°C, all the aldehydes and ketones into alcohols, which are less reactive than the former. By increasing the mixture stability, it is possible to avoid the formation of undesired macromolecules. The two following steps consist solely in the conversion of the oxygenated compounds, at a temperature around $350°C$ (Furimsky 2000), into alkanes and water by hydrogenation, hydrogenolysis, hydrodeoxygenation, cracking/hydrocracking, decarboxylation, decarbonylation and polymerization reactions (Wildschut 2009). To finish, the water generated through the process may impact the mixture state, creating a phase separation between the molecules with organic and aqueous character, as previously discussed. More detailed information about the possible hydrotreating process of bio-oil can be found in the following review: (Elliott 2007).

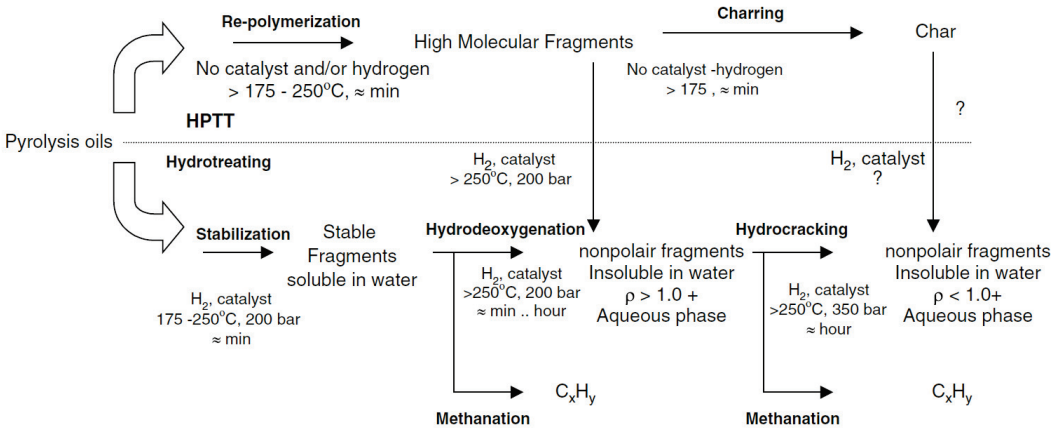


Figure I-4- Pathways of bio-oils hydrotreating (Venderbosch et al. 2010).

I.2.2 Bio-oil reactivity

It is known that oxygen containing groups follow a certain reactivity order. Therefore, some molecules will react more easily than others as a function of temperature. In 2007, Elliott (Elliott 2007) proposed a scheme, Figure I-5, based on data of the open literature.

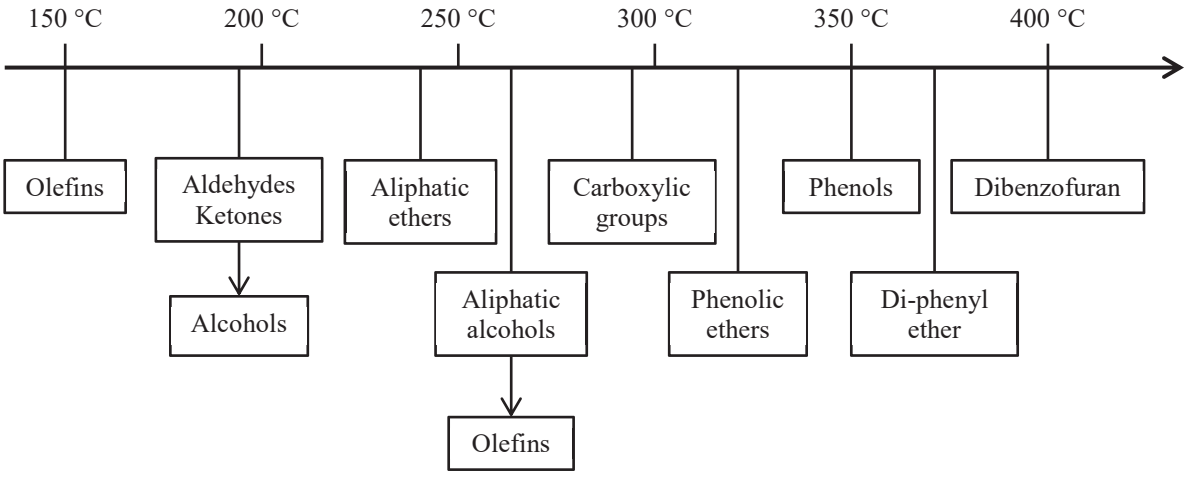


Figure I-5- Reactivity scale of oxygenated groups under HDO (adapted from (Elliott 2007)).

At lower temperatures, olefins, aldehydes and ketones are the first to react with hydrogen. This enables the reduction of the most reactive groups in bio-oils allowing the stabilization of the mixture. After, aliphatic ethers and alcohols are treated at a temperature around $250 - 300°C$, while around $300°C$ carboxylic groups are reduced, followed by aromatic compounds (Ar-OH), such as phenolic ethers, phenols and methoxy phenols, until the great majority of the oxygen is removed (Elliott 2007).

The order seen by Elliott (Elliott 2007), and presented in Figure I-5, is based on the bond dissociation energies. As seen in the review of Furimsky (Furimsky 2000) (exemplified by Table I-10), the reactivity of aliphatic ethers is lower than aliphatic alcohols, as the latter require more energy to break the atomic

bond. For the phenolic compounds, the exact same conclusion can be then assumed. Furthermore, it is possible to verify that the presence of an aromatic group increases the dissociation energy. This effect may be caused by a higher stability of the molecule, which implies that the higher the number of aromatic rings, the more difficult is their reactivity with other molecules, and, as consequence, the harder is their treatment (see Figure I-5).

Table I-10 – Examples of bond dissociation energies (kJ/mol) (Benson 1968).

Type of bond	RO-R	R-OH	RO-Ar	Ar-OH
Energy (kJ/mol)	339	385	422	468

I.2.2.1 Kinetic models

Despite the complexity associated to bio-oil, the scientific community still adventured into the understanding of the hydrotreating of bio-oil, through experimental studies. The following section will briefly present some of these studies, in particular those associated to a kinetic model.

For sulfided CoMo catalyst supported on alumina, Zhang et al. (Zhang et al. 2003) studied the conversion of bio-oil from fast pyrolysis of biomass. At 360 °C, 375 °C and 390 °C, the authors detected the reduction of hydroxyl groups thanks to the comparison between the Fourier Transform Infrared spectrum of raw and upgraded bio-oil. For the deoxygenation rate, the authors considered an n^{th} order kinetics with respect to the total of oxygen in the medium. In the end, a reaction order of 2.3 was obtained together with an apparent overall activation energy of 91.4 kJ/mol.

Later on, the same authors (Zhang et al. 2009) re-studied the process kinetics by creating a lumped model. The lumps were based on a chromatographic simulated distillation. As shown in Figure I-6, the model was divided in four lumps: feedstock, light fraction oil (below 250 °C), heavy fraction oil (above 250 °C) and gas, water and char (same lump).

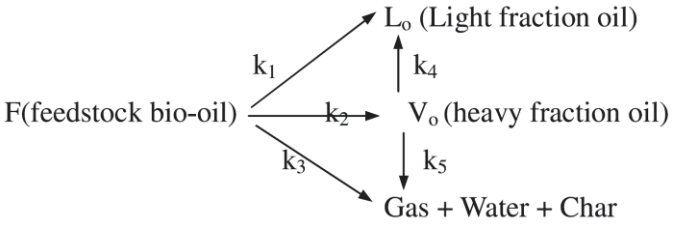


Figure I-6 - Reaction network proposed by Zhang et al. (Zhang et al. 2009) for the HDT of bio-oil over a sulfided $CoMo/Al_2O_3$ catalyst at 360 °C, 375 °C and 390 °C.

Alike the first model, Zhang et al. (Zhang et al. 2009) assumed a first order kinetic expression for each lump. A good agreement between the experimental and predicted data is given by the results in Table I-11.

Table I-11 - Kinetic parameters of upgrading of bio-oil (Zhang et al. 2009).

	1	2	3	4	5
k_{i0} (min^{-1})	2.15×10^3	2.20×10^4	9.57×10^2	2.51×10^5	4.28
Ea (kJ/mol)	64.81	75.44	66.281	97.12	27.88

For a NiMo catalyst supported on alumina, but in its oxide forms, Grilc et al. (Grilc et al. 2014) studied the hydrotreating of liquefied wood biomass. The hydrotreating tests were performed in a batch reactor at a temperature range of 200 °C – 300 °C. Like Zhang et al. (Zhang et al. 2003), the authors characterized the liquid product through Fourier Transform Infrared spectrum and consequently divided the effluent into lumps (see Figure I-7). By assuming a first order kinetics to all lumps, Grilc et al (Grilc et al. 2014) proposed a reaction network and a kinetic model. A suitable fit for the hydrotreating of liquefied biomass was achieved thanks to the parameters in Table I-12.

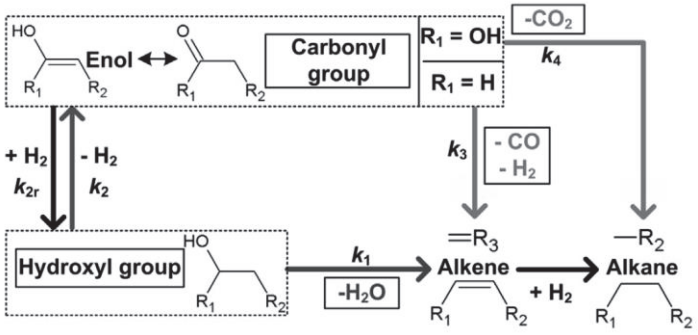


Figure I-7 – Reaction network proposed by Grilc et al. (Grilc et al. 2014) for the HDT of liquefied biomass over a NiMo/Al_2O_3 catalyst at a temperature range of 200 °C – 300 °C.

Reaction	Ea (kJ/mol)	A (/min)
1	73.8	9.78×10^4
2	11.0	2.38×10^7
2r	55.9	1.87×10^2
3	80.6	8.20×10^4
4	96.7	3.69×10^4

Table I-12 – Kinetic parameters for liquefied biomass Grilc et al. (Grilc et al. 2014).

I.2.3 Model compounds

As mentioned before, to overcome the vast range of reactivities of bio-oil, the scientific community started by studying single model compounds representative of bio-oil. Further on, mixtures of these model compounds began to be studied as well. The aim is to understand, firstly the main reaction of each bio-family (see Table I-5) and secondly the interconnected pathways between these families through mixtures. The following section will present a brief summary of the work done by the scientific

community, trying to focus on the exemplification of all possible reactions for each family. Moreover, as the objective of the present work is a kinetic study, the following section will also focus on reaction networks associated to kinetic studies, if available.

In the present section, the families of the aldehydes and ketones were merged, as their reactivity is similar. Also, in this section all phenolic compounds are grouped in the same family together with all the methoxy compounds. For more information the following references are recommended: (Furimsky 2000), (Elliott 2007), (He and Wang 2012) and (Wang et al. 2013).

I.2.3.1 Alcohols

Alcohols are compounds characterized by a hydroxyl group (OH). These molecules are mainly originated by the degradation of cellulose and hemicellulose. However, during HDT, compounds from other families, as aldehydes and ketones, may produce alcohols. The conversion of alcohols into hydrocarbons passes by the removal of the hydroxyl group.

For bifunctional catalysts, as Pt/SiO_2-Al_2O_3, sorbitol is converted by dehydration or retro-aldol condensation, as shown in Figure I-8. Furthermore, the derivatives of sorbitol may be transformed by the same reactions or by hydrodeoxygenation or decarbonylation as seen by Li and Huber (Li and Huber 2010) at 245 °C. For smaller alcohols, as butanol, the authors proposed only dehydration and hydrogenation of the hydroxyl group as potential reactions. A similar reaction network was seen by Moreno et al. (Moreno et al. 2013) also at 245 °C. Furthermore, the authors added ring closure by dehydration and hydrogenolysis of the ring as possible reactions for sorbitol and its derivatives, respectively. In its work, Moreno et al. (Moreno et al. 2013) created a kinetic model to described the conversion of sorbitol and also its products. Although the kinetic model was in agreement with the experimental data, a low product quantification for several temperatures enabled the inclusion of the temperature dependence.

For noble catalysts, as Pt/Al_2O_3, sorbitol follows mainly dehydration or dehydrogenation of the hydroxyl group, as shown in Figure I-9. According to Kirilin at al. (Kirilin et al. 2014) at 225 °C, the generated products may then follow hydrogenation, in the case of dehydration, or decarbonylation, in the case of dehydrogenation. For these reactions, the authors (Kirilin et al. 2014) created a kinetic model, which, although it predicted well the formation of the compounds in the liquid phase, was not able to describe minor components in the gas phases, as alkanes. For the same catalyst and in a temperature range of 120 °C to 250 °C, smaller alcohols, as propanol, propanediol and glycerol are also converted though dehydration and dehydrogenation (Peng et al. 2012).

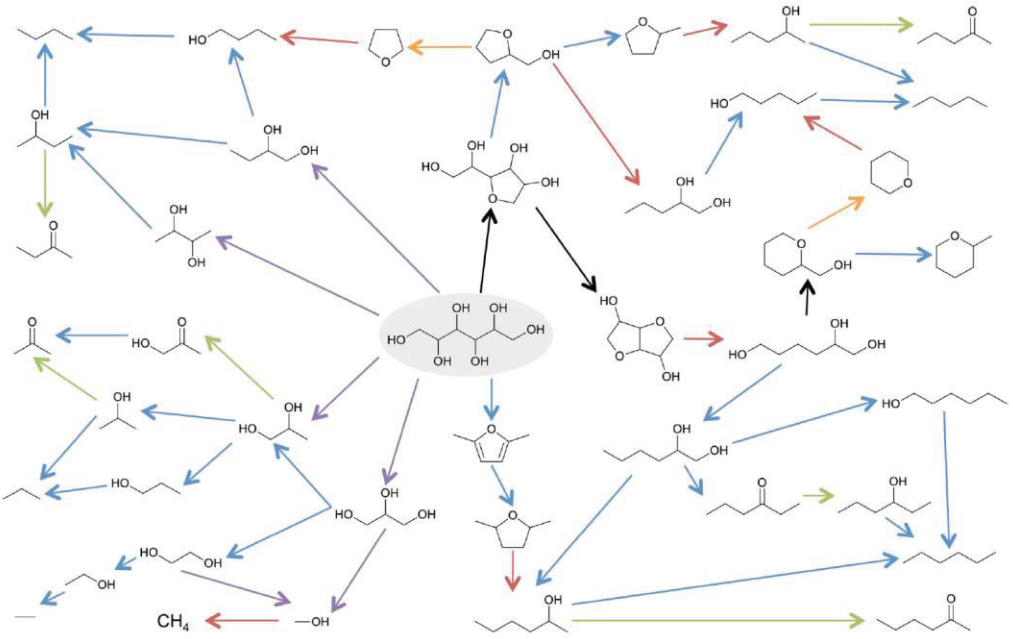


Figure I-8 – Reaction network proposed by Moreno et al (Moreno et al. 2013) for the HDT of sorbitol over a Pt/SiO2-Al2O3 catalyst at 245 °C.

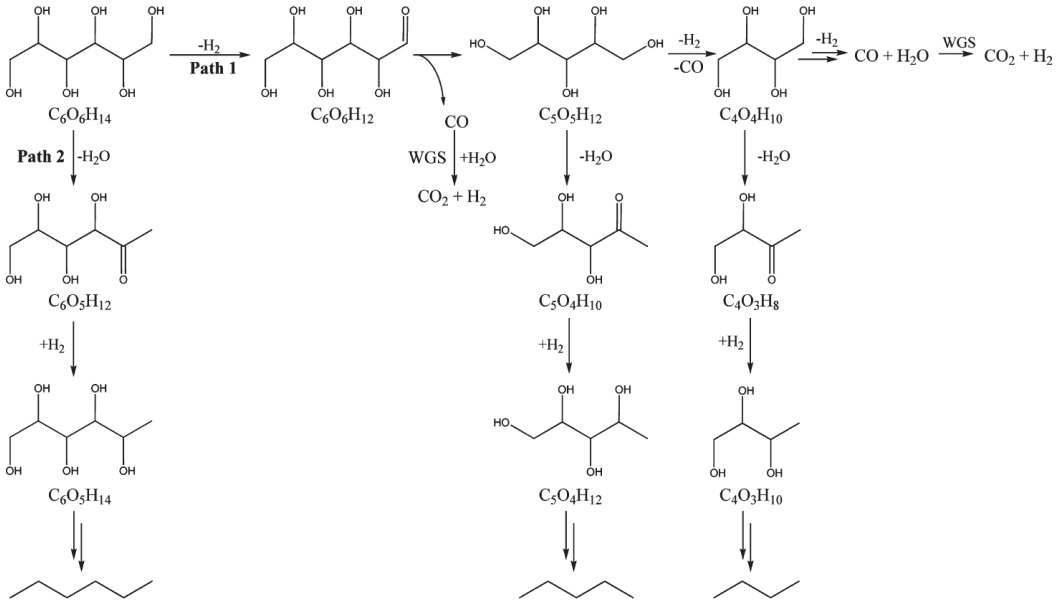


Figure I-9 – Reaction network proposed by Kirilin et al. (Kirilin et al. 2014) for the HDT of sorbitol over a Pt/Al2O3 catalyst at 225 °C.

In summary, alcohols can be converted by dehydration, dehydrogenation, hydrogenation, and retro-aldol reactions. Furthermore, for molecules with several hydroxyl groups, as sorbitol, dehydration and retro-aldol are the main routes for bifunctional catalysts, while dehydration and dehydrogenation are more prosperous to happen in noble catalysts. For smaller compounds with few hydroxyl groups, molecules tend to convert by hydrogenation and dehydration for bifunctional catalyst and dehydration and

dehydrogenation for noble catalysts. Independently of the catalyst, all the proposed steps may eventually lead to the formation of hydrocarbons through decarbonylation or hydrogenation.

I.2.3.2 Aldehydes and ketones

Aldehydes and ketones are compounds characterized by a carbonyl group (CO). In the case of aldehydes, this group is connected to a carbon and a hydrogen, while in the case of ketones the carbonyl is connected to two carbons. Both types of molecules are mainly originated by the degradation of cellulose and hemicellulose. However, during HDT, compounds from other families, as furans or carbohydrates, may produce aldehydes and ketones. The conversion of aldehydes and ketones into hydrocarbons passes by the removal of the carbonyl group.

In the case of aldehydes, Procházková et al. (Procházková et al. 2007) tested benzaldehyde, dissolved in hexane, on supported palladium catalysts, namely zeolites and active carbon at a temperature range of 30 °C – 130 °C. When Pd/C was used, benzaldehyde converts into benzylalcohol and later toluene by hydrogenation, as shown in Figure I-10. Benzaldehyde may undergo as well direct hydrogenolysis of the carbonyl group, forming toluene. In the case of the zeolites, when palladium was supported on zeolite beta, hydrogenation was the main transformation pathway for benzaldehyde. However, when palladium was supported on ZSM-5 zeolite, the main reaction would be direct hydrogenolysis. Through the adaptation of a kinetic model, Procházková et al. (Procházková et al. 2007) denoted that hydrogenation of the carbonyl is the fastest reaction. In addition, and as mentioned in the previous chapter, aldehydes may undergo decarbonylation. Both for bifunctional and noble catalyst supported on alumina, products of dehydrogenation of sorbitol (formation of a carbonyl) could undergo decarbonylation. The authors also studied the impact of the solvent. When dissolved into methanol, for the same catalysts and temperatures, benzaldehyde generates all products shown in Figure I-10. The formation of ethers, as molecule D, is undesirable. However, the acid medium promotes its formation. With this brief example, it is clear that besides the temperature, the catalyst and the structure of the molecule, the solvent also affects the molecules reactivity.

In the case of ketones, González et al. (González et al. 2015) tested acetophenone in gas phase over noble metals supported on alumina at a temperature range of 275 °C – 375 °C. Briefly, acetophenone is converted by hydrogenation either of the carbonyl or of the ring, as shown in Figure I-11. Its products may then undergo hydrogenation of the carbonyl/hydroxyl groups or dehydration of the hydroxyl group. Independently of the precious metal used, ethylbenzene was the main product, indicating that acetophenone follows preferentially hydrogenation of the carbonyl, then dehydration of the hydroxyl and finally hydrogenation of the double bond into ethylbenzene. In this study, the author included a kinetic model to acetophenone over Pt/Al_2O_3. The experimental data fitted well the pseudo-first-order kinetic model proposed by the author.

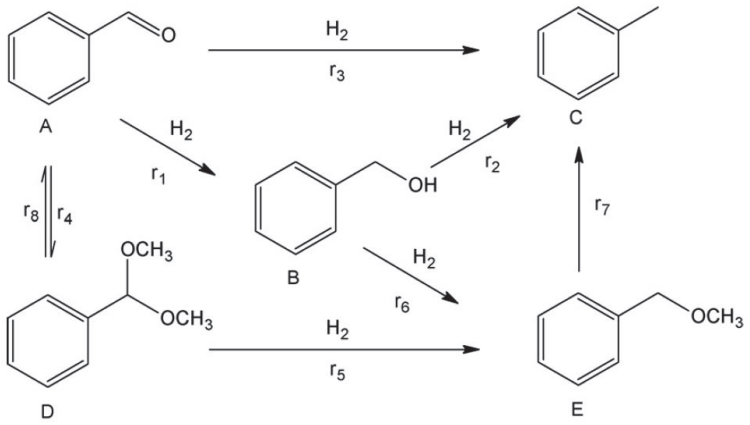


Figure I-10 – Reaction network proposed by Procházková et al. (Procházková et al. 2007) for the HDT of benzaldehyde over a Pd/C catalyst at 30 °C – 130 °C.

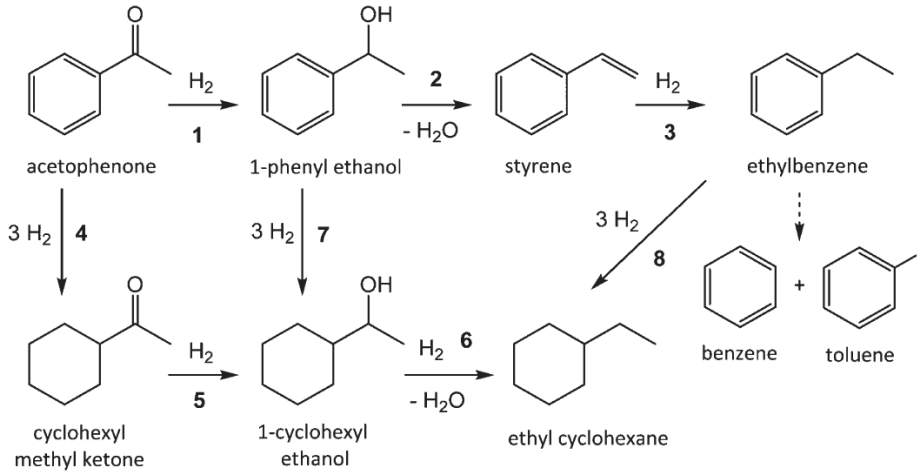


Figure I-11 – Reaction network proposed by González et al. (González et al. 2015) for the HDT of acetophenone over a Pt/Al_2O_3 catalyst at 275 °C – 375 °C.

For conventional HDT catalyst, Laurent et Delmon (Laurent and Delmon 1994) analyzed the conversion of methyl-acetophenone at 260 °C. The main product was ethyl-methyl-benzene, and no methyl-benzylalcohol or methyl-styrene was detected at lower temperatures. For this reason, either direct hydrogenolysis may be the only reaction undergone by the methyl-acetophenone, or hydrogenation of the carbonyl, followed by hydrogenation or dehydration of the hydroxyl and then hydrogenation of the double bond are extremely fast reactions. From these tests, Laurent and Delmon (Laurent and Delmon 1994) concluded that CoMo and NiMo catalyst present as activation energies for this reaction 12 kcal/mol and 17.5 kcal/mol respectively.

In summary, aldehydes and ketones, over noble metals, may be converted by hydrogenation of the carbonyl group or direct hydrogenolysis of the carbon-oxygen double bond. In the first case, the alcohol is normally transformed by dehydration which is later on hydrogenated. In the case of conventional HDT catalyst, ketones tend to convert through direct hydrogenolysis.

I.2.3.3 Carbohydrates

Carbohydrates (or sugars) are compounds characterized by a saturated ring with one oxygen atom and four or five carbons. These compounds also present hydroxyl group adjusted to the ring. Carbohydrates are originated by the dehydration of cellulose. Unlike any other family so far, carbohydrates are not intermediate products of other families. These molecules have a high reactivity and are easily transformed into other compounds by the opening of the ring.

For Ru/C catalysts, Vispute and Huber (Vispute and Huber 2009) released that levoglucosan was transformed into glucose by hydration. The same pathway was seen by Bindwal and Vaidya (Bindwal and Vaidya 2013) at a temperature range of 125 °C – 160 °C. Moreover, the authors proposed for the same catalyst the conversion of levoglucosan into glucose, which through hydrogenation forms sorbitol. A similar reaction network was also seen by Wildschut et al. (Wildschut et al. 2009) at a temperature range of 250 °C – 400 °C. However, the author denoted as well the formation of carboxylic acids and furanic species, which are usually related to the thermal reactions. The sorbitol generated by the hydrotreating of D-glucose undergoes retro-aldol and dehydration reactions. The furanic and acid species continue their conversion through the thermal route, as shown in Figure I-12.

Besides the reactional pathways, summarized in Figure I-13, Bindwal and Vaidya (Bindwal and Vaidya 2013) generated a kinetic model for the transformation of levoglucosan. As the authors did not see glucose and sorbitol in the product mass, the model included a direct pathway from levoglucosan to small alcohols as ethylene glycol, propylene glycol and 1,4-butanediol. For this transformation, the authors found an activation energy of 94.5 kJ/mol.

The transformation of glucose into sorbitol had been already studied by several authors, including Crezee (Crezee 2003). The author analyzed the kinetics of glucose conversion into sorbitol over a Ru/C catalyst at a temperature range of 100 °C – 130 °C. At lower temperatures than Bindwal and Vaidya (Bindwal and Vaidya 2013), glucose converts totally into sorbitol. According to the author, this transformation is associated to an observed activation energy in the order of 55 kJ/mol.

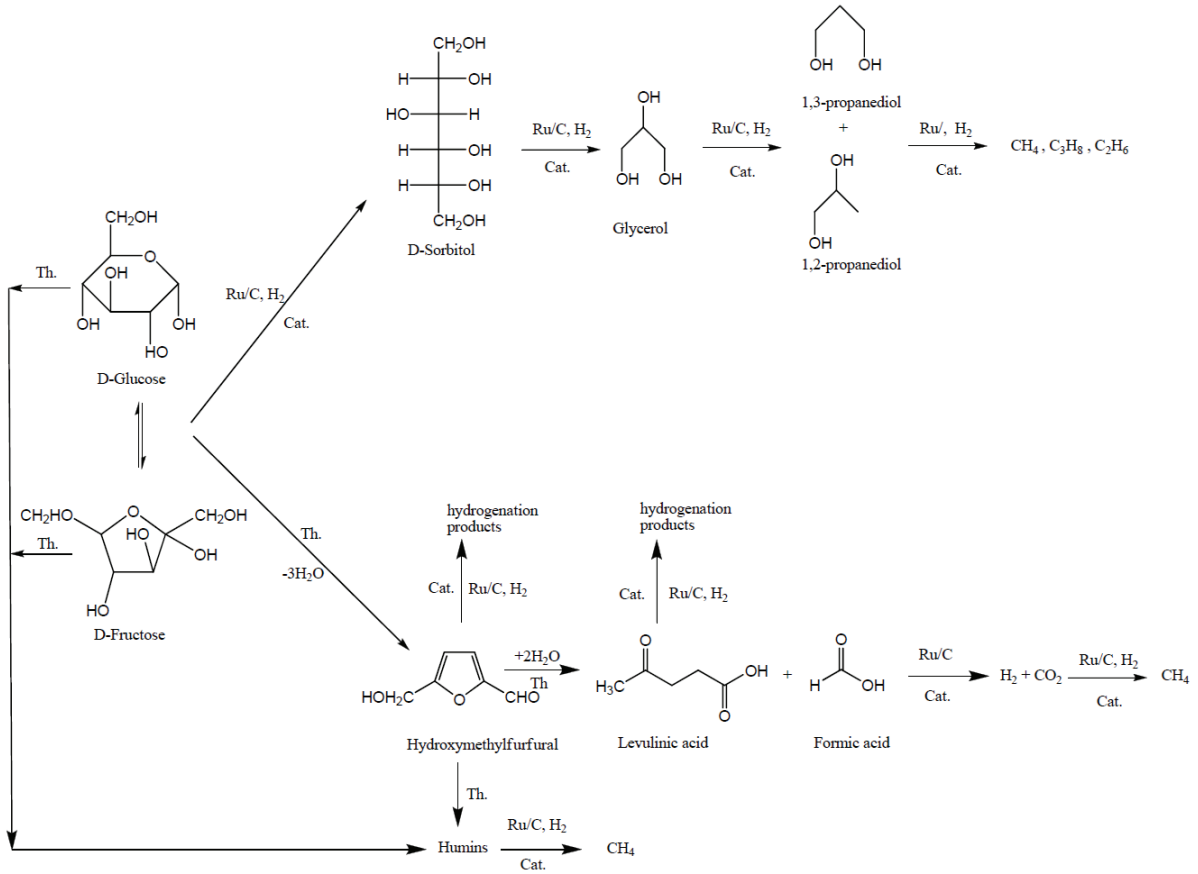


Figure I-12 – Reaction network proposed by Wildschut et al. (Wildschut et al. 2009) for the HDT of D-glucose over a Ru/C catalyst at250 °C – 400 °C.

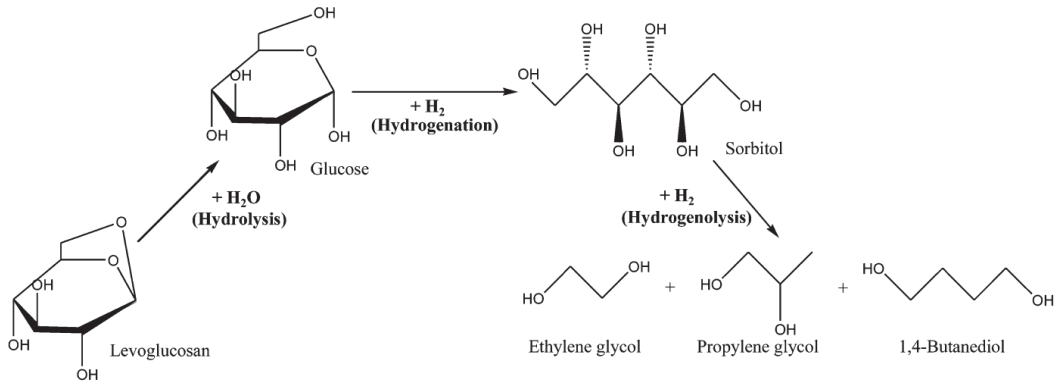


Figure I-13 – Reaction network proposed by Bindwal and Vaidya (Bindwal and Vaidya 2013) for the HDT of levoglucosan over a Ru/C catalyst at 125 °C – 160 °C.

For reduced NiMo catalyst supported on alumina and temperatures of 200 °C, 250 °C and 300 °C, Ozagac et al. (Ozagac et al. 2016b) saw a similar behavior as Wildschut et al. (Wildschut et al. 2009). D-glucose, in isomerization with D-fructose, either undergoes retro-aldol into sorbitol or is dehydrated into cyclic species, as shown in Figure I-14. Later on, sorbitol undergoes once again retro-aldol or dehydration reactions. The conversion of furanic species will be discussed in the following section.

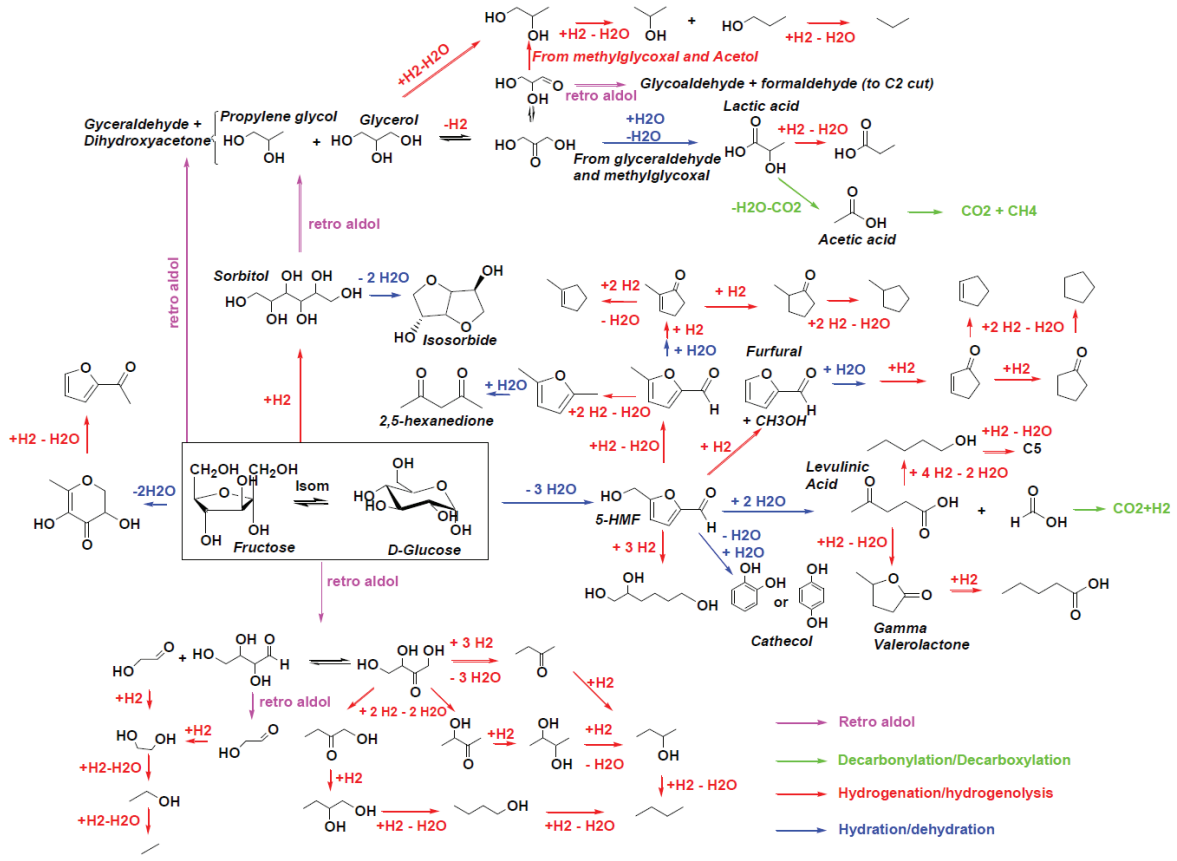


Figure I-14 - Reaction network proposed by Ozagac et al. (Ozagac et al. 2016b) for the HDT of D-glucose over a reduced NiMo/Al_2O_3 catalyst at 200 °C, 250 °C and 300 °C.

In summary, carbohydrates, over noble metals, may be converted by hydrogenation or dehydration (thermal degradation). In the first case, carbohydrates react through opening of the tetrahydropyran ring with the help of hydrogen, forming polyalcohols. In the second case, carbohydrates form furanic species through removal of water from the molecule. For reduced NiMo catalysts, carbohydrates may undergo the same reactions and also retro-aldol reaction, which open the tetrahydropyran ring.

I.2.3.4 Carboxylic Acids

Carboxylic acids are compounds characterized by a carboxyl group (COOH). This type of molecules mainly originates by the degradation of cellulose and hemicellulose. However, during HDT, compounds from other families, as carbohydrates or even poly alcohols, may produce carboxylic acids. The conversion of these into hydrocarbons passes by the removal of the carboxyl.

For a InNi/Al_2O_3 and a commercial reduced NiMo/ Al_2O_3 catalysts, Onyestyák et al. (Onyestyák et al. 2012) studied the reactivity of the acetic acid at a temperature range of 220 °C – 380 °C. During hydrotreating, the acetic acid hydrogenates into an acetaldehyde, which will then be hydrogenated into ethanol. This alcohol can proceed with the HDO process and produce alkanes or it may react with the acetic acids in the medium through esterification generating esters. The acetic acid may also undergo

decarboxylation producing methane and carbon dioxide or react with another acetic acid to form acetone and carbon dioxide. For the commercial catalyst decarboxylation in the main pathway, while for $InNi/Al_2O_3$, acetic acid prefers the hydrogenation and esterification steps.

Also for a reduced $NiMo/Al_2O_3$, Joshi and Lawal (Joshi and Lawal 2012) analyzed the reactivity of acetic acid at 300 °C, as shown in Figure I-15. At low temperatures, acetaldehyde was the majority product. However, hydrogenation stated to loss field when at higher temperatures, ethyl acetate was detected in addition to ethanol. With the increase of the temperature, the quantity of methane, carbon dioxide and acetone also increased. This indicates that decarboxylation and other side reactions are preferable at higher temperature.

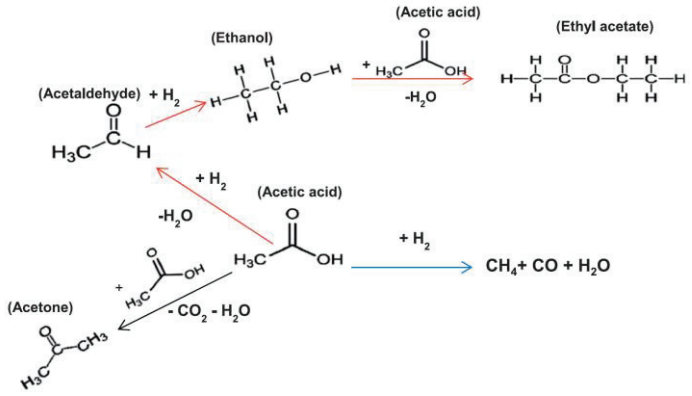


Figure I-15 - Reaction network proposed by Joshi and Lawal (Joshi and Lawal 2012) for the HDT of acetic acid over a reduced $NiMo/Al_2O_3$ catalyst at 300 °C.

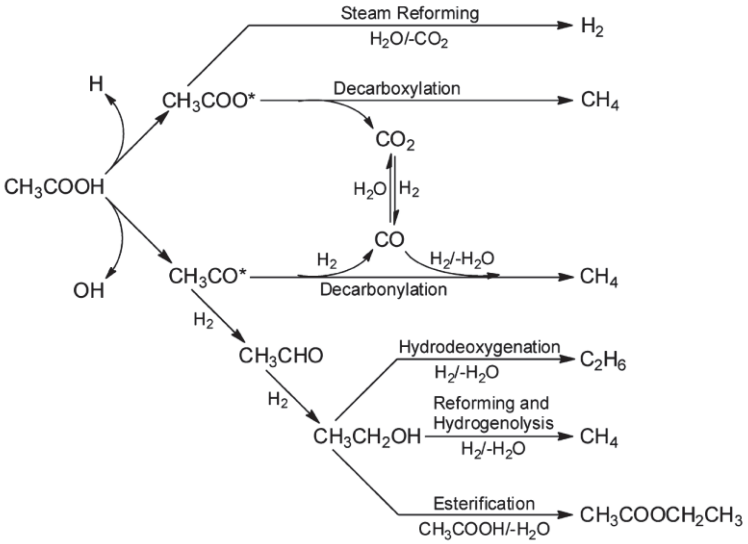


Figure I-16 - Reaction network proposed by Wan et al. (Wan et al. 2013) for the HDT of acetic acid over a Ru/C catalyst at 150 °C – 300 °C.

As well for acetic acid, Wan et al. (Wan et al. 2013) detected a similar reaction network over a Ru/C catalyst at a temperature range of 150 °C – 300 °C, as shown in Figure I-16. As already discussed, acetic acid generates acetaldehyde which will produce ethanol. This alcohol can proceed with the HDO process and produce alkanes or it can react with the carboxylic acids in the medium through esterification. Moreover, acetic acid may undergo decarboxylation. The formation of acetone was not detected by Wan et al. (Wan et al. 2013).

Independently of the catalyst, carboxylic acids, may be converted by hydrogenation of the hydroxyl in the carboxyl group or may undergo decarboxylation. The latter reaction forms products as methane and carbon dioxide. However, the former reaction generates an aldehyde which may continue in the hydrogenation path until alcohols and then alkanes. For any carboxylic acid was found a kinetic study.

I.2.3.5 Furans

Furans are compounds characterized by an aromatic ring with one oxygen atom and four carbons. These compounds, like carbohydrates, may present hydroxyl or carbonyl group attached to the ring. Furans are originated by the degradation of cellulose and hemicellulose. However, during HDT, compounds from other families, as carbohydrates, may produce furanic compounds. The conversion of these compounds into hydrocarbons passes by the saturation and/or opening of the ring.

For Pt/C, Vaidya and Mahajani (Vaidya and Mahajani 2003) studied the conversion of furfural into furfuryl alcohol at a temperature range of 130 °C – 175 °C. For the reaction, the authors created a kinetic model, where a zeroth-order with respect to furfural was considered. For this Vaidya and Mahajani (Vaidya and Mahajani 2003) concluded that an activation energy of 28 kJ/mol is required to convert furfural to furfuryl alcohol.

As well for furfural, Sitthisa et al. (Sitthisa et al. 2011) studied the hydrogenation studies over a Cu/SiO_2 catalyst at 230 °C – 290 °C. As the previous authors, they identified the conversion of furfural into furfuryl alcohol as the main path. Methyl furan was also detected, but as a minor product of furfural and the main product of furfural alcohol. For the observed results, Sitthisa et al. (Sitthisa et al. 2011) created a simple model, where direct hydrogenolysis of furfural was neglected, as shown in Figure I-17. As results, the authors found an activation energy of 11.8 kcal/mol and 12.4 kcal/mol for the hydrogenation of furfural and furfuryl alcohol respectively.

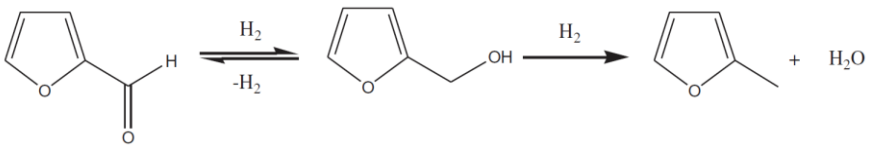


Figure I-17 - Reaction network proposed by Sitthisa et al. (Sitthisa et al. 2011) for the HDT of furfural over a Cu/SiO_2 catalyst at 230 °C – 290 °C

For traditional sulfided and reduced NiMo/γ-Al_2O_3, Bunch et al. (Bunch et al. 2007) studied the hydrodeoxygenation of benzofuran at 200 °C and 320 °C, as shown in Figure I-18. Both catalysts promote the saturation of the furanic ring. However, the sulfide catalyst favors the opening of the tetrahydrofuran ring, while the reduced catalyst favors the hydrogenation of the benzene ring. For these reasons, the authors concluded that the major reaction pathway for the hydrotreating of benzofuran over sulfide catalysts was the hydrogenolysis, while for reduced catalysts, it was the hydrogenation. From these results, Bunch et al. (Bunch et al. 2007) concluded as well that the apparent activation energy of the hydrotreating of benzofuran over sulfide catalysts (192.4 kJ/mol) was higher than for the reduced catalysts (65.5 kJ/mol).

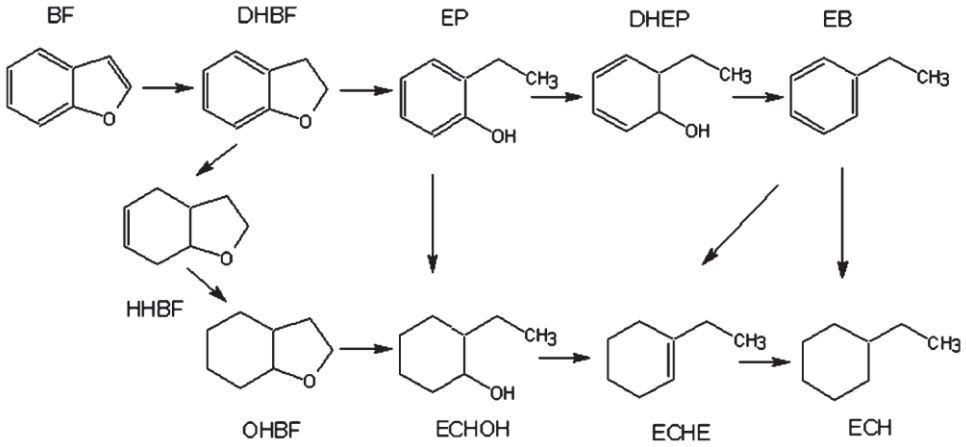


Figure I-18 - Reaction network proposed by Bunch et al. (Bunch et al. 2007) for the HDT of benzofuran over a sulfided and reduced NiMo/γ-Al_2O_3 catalysts at 200 °C and 320 °C.

As well for reduced NiMo catalyst supported on alumina, Ozagac et al. (Ozagac et al. 2016b) studied the conversion of furfural at 200 °C, 250 °C and 300 °C. Contrary to the previous studies, the authors identified the hydrogenation and decarbonylation of the carbonyl in the furfural, as the two possible pathways. Similar to glucose, the author tried as well to comprehend the full reaction network of the hydrotreating of furfural and its derivatives, as shown in Figure I-19.

In summary, furanic species may react by hydrogenation, saturation and also ring opening. Although the first and last reaction may be similar in the presence of hydrogen, the opening of the ring may also happen by water addition, as seen by Ozagac et al. (Ozagac et al. 2016b). The saturation reaction is a typical pathway for the conversion of aromatic rings, as furans. To finish, and as also seen by Ozagac et al. (Ozagac et al. 2016b), aldehydes may undergo decarbonylation. Although this pathway was not seen in section I.2.3.2, it is a possibility at a given set of conditions.

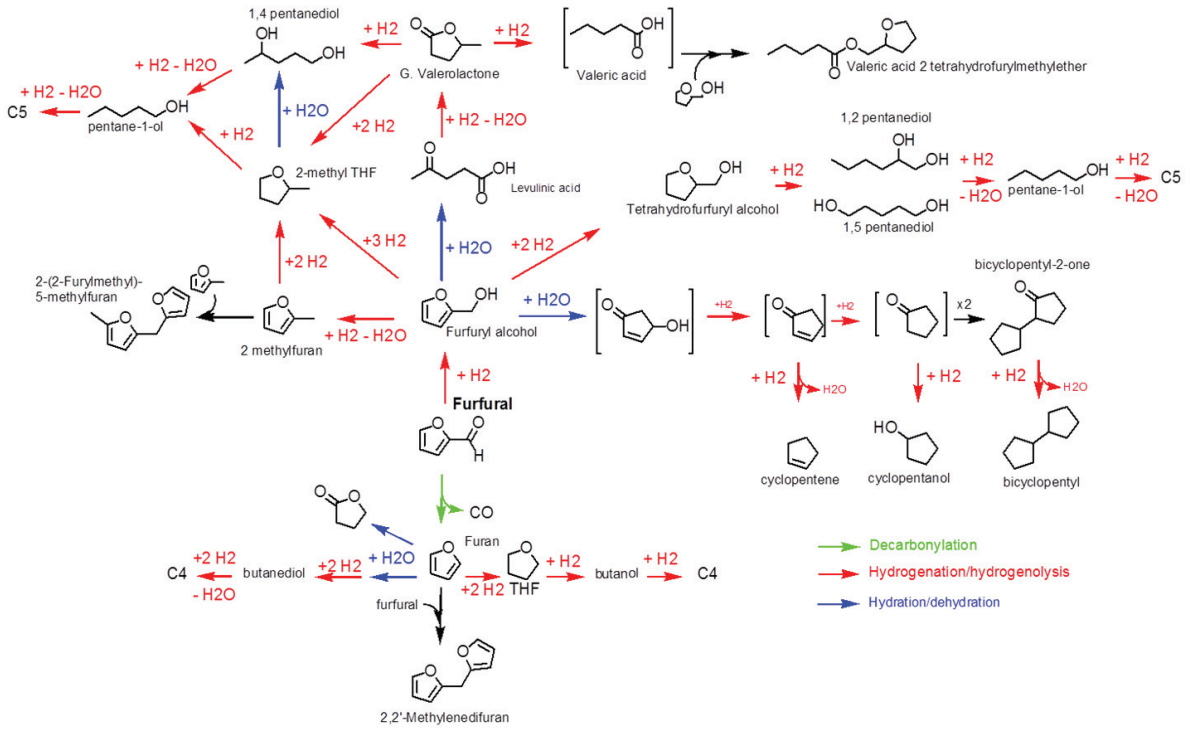


Figure I-19 - Reaction network proposed by Ozagac et al. (Ozagac et al. 2016b) for the HDT of furfural over a reduced NiMo/γ-Al_2O_3 catalyst at 200 °C, 250 °C and 300 °C.

I.2.3.6 Phenols and methoxy-phenols

Phenols and methoxy-phenols are compounds characterized by an aromatic ring with six carbons and hydroxyl or methoxy groups attached to the ring. Both types of molecules are originated by the degradation of lignin. Similar to carbohydrates, phenols and methoxy-phenols are not intermediated products. Although quite stable, when converted these molecules tend to form aromatic or saturated six carbons rings.

For a commercial sulfided $CoMo/Al_2O_3$ catalyst, Gevert et al. (Gevert et al. 1987) studied the HDO of mono and dimethyl substituted phenols at 300 °C. In the article, an example is given for 4-methylphenol, which, when under HDT conditions, either undergoes the hydrogenation of the hydroxyl, or the incomplete/complete hydrogenation of the ring, as shown in Figure I-20.

For the kinetic study, which accounted with the heat-up of the batch reactor, the authors considered pseudo first-order reactions. In the end, and for several molecules, Gevert et al. (Gevert et al. 1987) obtained the pseudo first order rate constants, as seen in Table I-13.

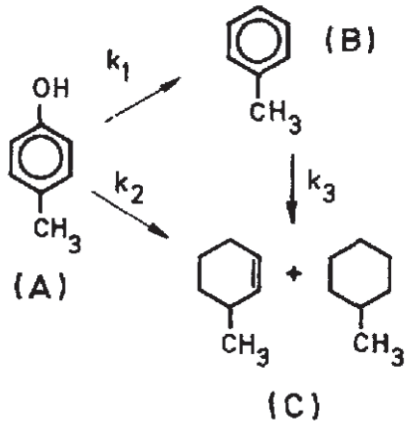


Figure I-20 - Reaction network proposed by Gevert et al. (Gevert et al. 1987) for the HDT of 4-methylphenol over a sulfide CoMo/Al_2O_3 catalyst at 300 °C.

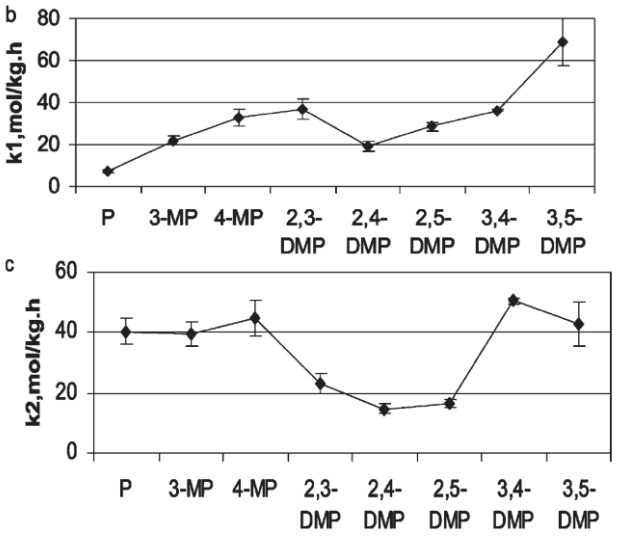


Figure I-21 – Effect of the methyl position on (b) hydroxyl HDO rate constant, k_1, and (c)ring HDT rate constant, k_2,over a sulfide CoMo/Al_2O_3 catalyst at 300 °C.

Table I-13 – Pseudo first-order rate constants for methyl-substituted phenols (Gevert et al. 1987).

Molecules	Hydrodeoxygenation k_1 (cm^3/min/g)	Hydrogenation k_2 (cm^3/min/g)
4-Methylphenol	2.40 ± 0.10	0.52 ± 0.15
2-Methylphenol	1.26 ± 0.08	0.17 ± 0.01
2,4-Dimethylphenol	1.50	0.18
2,6-Dimethylphenol	0.27 ± 0.01	0.11 ± 0.01
2,4,6-Trimethylphenol	0.45 ± 0.15	0.05 ± 0.02

Gevert et al. (Gevert et al. 1994), continued their studies for 2,6-dimethylphenol (2,6-DMP) and 3,5-dimethylphenol (3,5-DMP) over the same catalyst and at the same temperature. Some years later, Massoth et al. (Massoth et al. 2006) continued the studies by analyzing the effects of the methyl position in the rate constants of methyl-substituted phenols. Also at 300 °C and over a sulfided CoMo/Al_2O_3 catalyst, the authors tested the conversion of phenol (P), 2-methylphenol (2-MP), 3-methylphenol (3-MP), 4-methylphenol (4-MP), 2,3-dimethylphenol (2,3-DMP), 2,4-dimethylphenol (2,4-DMP), 2,5-dimethylphenol (2,5-DMP), 3,4-dimethylphenol (3,4-DMP), 2,3,5-Trimethylphenol (2,3,5-TMP), 2,3,6-Trimethylphenol (2,3,6-TMP), 2,4,6-Trimethylphenol (2,4,6-TMP). Thanks to a Langmuir-Hinshelwood kinetic model based on Figure I-20, the authors were able to follow the dependence of the rate constants and the methyl position (see Figure I-21).

For a Pt/γ-Al_2O_3 catalyst, Nimmanwudipong et al. (Nimmanwudipong et al. 2011) studied the conversion of guaiacol at 300 °C. Several products were identified, but the majority consisted in phenol, catechol and 3-methylcatechol, in addition to water and methanol. For the following results, the author considered a first-order kinetics for guaiacol and its primary products. The results of the first-order rate constants if given by Figure I-22. As reported by the author, it is clear that "(…) deoxygenation by removal of the methyl group is more favorable than deoxygenation by removal of the hydroxyl group.".

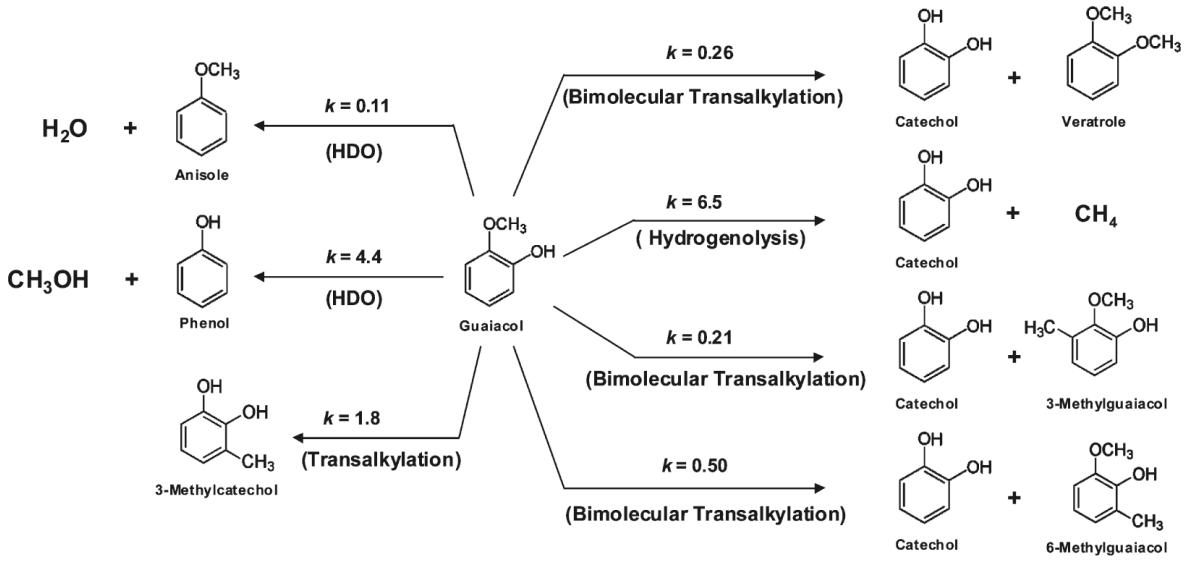


Figure I-22 - Reaction network proposed by Nuimmanwudipong et al. (Nimmanwudipong et al. 2011) for the HDT of guaiacol over a sulfide Pt/γ-Al_2O_3 catalyst at 300 °C.

For the same catalyst and temperature, Runnebaum et al. (Runnebaum et al. 2011), from the same research group, analyzed the conversion of anisole. Similar to guaiacol, several products were identified, but the most abundant molecules were phenol, 2-methylphenol, benzene and 2,6-dimethylphenol. By considering a first-order kinetics for the reactant and its derivatives, the authors realized that demethylation assumes the lead in the conversion of anisole, followed by transalkylation to the ortho position (see Figure I-23).

For a reduced CoMo/Al_2O_3 catalyst, Otyuskaya et al. (Otyuskaya et al. 2017) studied also the conversion of anisole. In a gas phase plug-flow reactor between 275 °C – 350 °C, the authors saw as main products phenol, cresol, benzene, toluene and some traces of methyl-substituted phenols, benzenes and anisoles. For these results, the authors considered as the main reaction pathways demethylation, transalkylation, hydrodeoxygenation and methylation. Since no traces of methanol were detected, demethoxylation was not considered, as shown in Figure I-24.

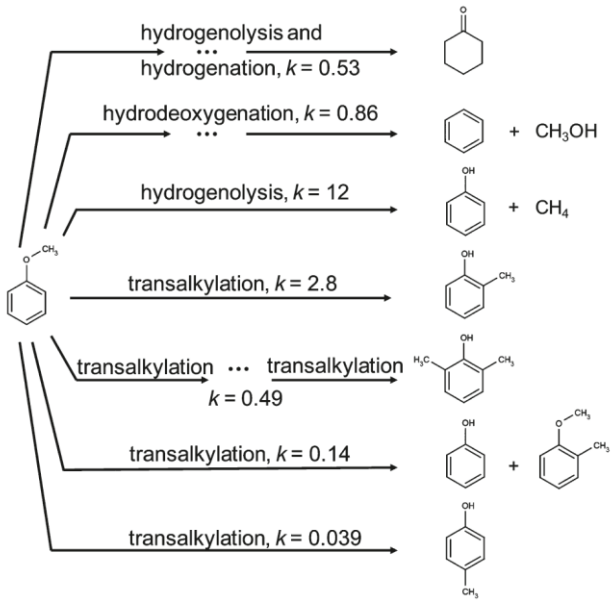


Figure I-23 - Reaction network proposed by Runnebaum et al. (Runnebaum et al. 2011) for the HDT of anisole over a Pt/γ-Al_2O_3 catalyst at 300 °C.

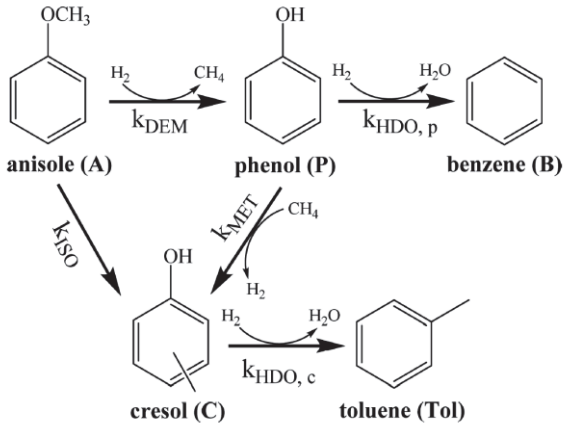


Figure I-24 - Reaction network proposed by Otyuskaya et al. (Otyuskaya et al. 2017) for the HDT of anisole over a CoMo/Al_2O_3 catalyst at 275 °C - 350 °C.

Through the application of a Langmuir-Hinshelwood-Houston-Watson type kinetic, the authors developed a model with 27 adjustable parameters. In order to reduce this number, the pre-exponential factors were calculated by transition state theory, while the remaining parameters were fitted to the experimental data. Table I-14 summarizes the pre-exponential and activation energies values.

For sulfided CoMo and NiMo/γ-Al_2O_3 catalyst, Laurent and Delmon (Laurent and Delmon 1994) studied the hydrotreating of guaiacol at 260 °C, 280 °C and 300 °C. According to the authors, guaiacol has two regions of conversion: the initial and middle, where the rate conversion of the initial data is superior to the middle data. The authors also determinate a third reaction rate constant, which was

defined for the production of catechol and phenol, indicating that demethoxylation and demethylation were the main reactions. In the end, Laurent and Delmon (Laurent and Delmon 1994) obtained for the CoMo/γ-Al_2O_3 catalyst, 17 kcal/mol, 26.9 kcal/mol and 25.1 kcal/mol as apparent activation energies for the initial, middle and products conversion respectively. For a NiMo/γ-Al_2O_3 catalyst, values of 14 kcal/mol, 26.5 kcal/mol and 25.0 kcal/mol were obtained as apparent activation energies for the initial, middle and products conversion respectively.

Table I-14 – Pre-exponential and activation energies values for anisole HDO (Otyuskaya et al. 2017)

Molecules	Pre-exponential (kg/s/mol or /s)	Activation energy (kJ/mol)
Anisole demethylation	10^{14}	82.6
Anisole transalkylation	10^{12}	108.0
Phenol methylation	10^{11}	43.8
Phenol HDO	10^{13}	81.6
Cresol HDO	10^{13}	74.0

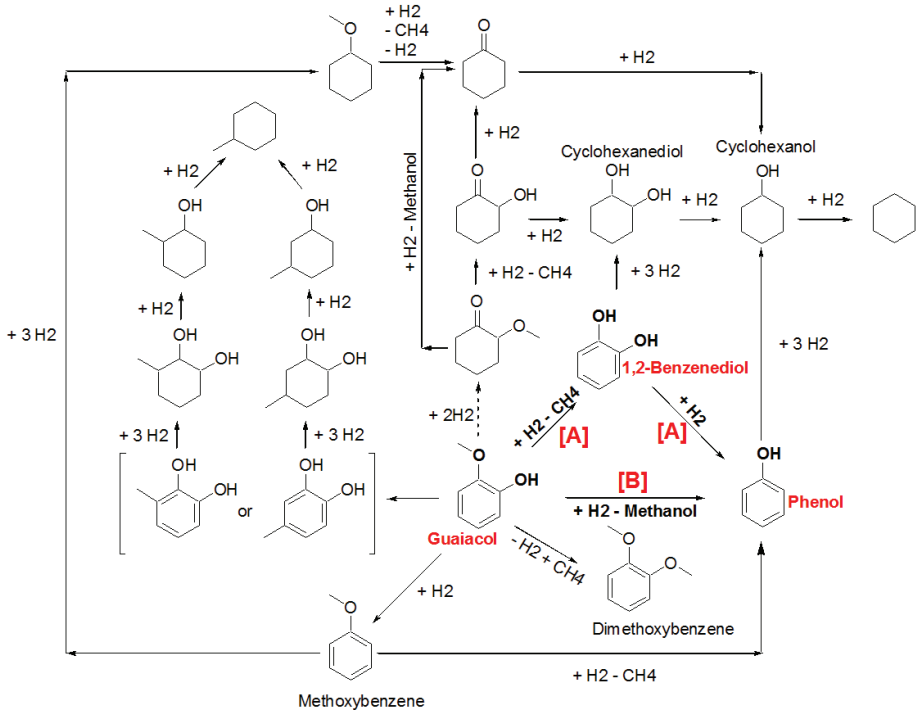


Figure I-25 - Reaction network proposed by Ozagac (Ozagac 2016) for the HDT of guaiacol over a reduced NiMo/γ-Al_2O_3 catalyst at 200 °C, 250 °C and 300 °C.

For a reduced NiMo/γ-Al_2O_3 catalyst, Ozagac (Ozagac 2016) studied the conversion of guaiacol at 200 °C, 250 °C and 300 °C. Similar to the previous studies, the author identified several products, which were organized in a single reaction network. According to Ozagac (Ozagac 2016), demethoxylation and

demethylation were the main pathways for the conversion of guaiacol. Similar to glucose and furfural, the author tried as well to comprehend the full reaction network of the hydrotreating of guaiacol and its derivatives, as shown in Figure I-25.

In summary, phenols and methoxy-phenols may undergo hydrogenation, saturation, demethylation, demethoxylation and transalkylation. As expected, these three last reactions depend on the presence of a methoxy group. The saturation is a typical pathway for the conversion of aromatic rings, as benzene, whereas hydrogenation consists mainly in the removal of the oxygen element. To finish, according to several authors (Otyuskaya et al. 2017; Ozagac 2016), phenols and methoxy-phenols may still be methylated. Although this reaction may be a desirable one, as it adds a carbon atom to the molecules, it will also increase the complexity of the feed (formation of dimethoxybenzene).

I.2.4 Mixtures of model compounds

Although single model compounds are useful to gain a perspective of the reactivity of a particular molecule, they lack the general view of a bio-oil feedstock. According to Wang et al. (Wang et al. 2013), "(…) the competitive adsorption of molecules in (…) mixtures and subsequent inhibiting and poisoning effects, [causes] significant differences (…) in reaction rates and selectivities measured from single model compounds and from the same model compound in various mixtures". For these reasons, the scientific community expanded their investigation from single model compounds to mixtures of these same compounds.

As for the single model compounds, the following section will present a brief summary of the work done by the scientific community regarding model compound mixtures. Once again, the aim is to illustrate the main interactions between the molecular species and always considering the kinetic point of view, if this is available. Moreover, the following section will present a detailed study of the reaction behind the formation macromolecules and coke.

I.2.4.1 Open literature studies

Elliott et al. from the Pacific Northwest National Laboratory (Elliott et al. 2006) studied the conversion of three synthetic bio-oil mixtures at low temperature and over a Ru/C catalyst at a temperature range of 150 °C – 200 °C. From the three mixtures, represented in Table I-17, the authors concluded that under hydrogenation conditions, the phenolic compounds would undergo dealkylation of any side chain, saturation of the ring and demethylation followed by the removal of the hydroxyl or demethoxylation. Furthermore, the molecules with carbonyl groups would undergo hydrogenation into alcohols. In a wider perspective, the conversion process applied by Elliott et al. (Elliott et al. 2006) at low temperature allowed the transformation of unstable oxygenated compounds, as aldehydes and ketones, into stable molecules, as aliphatic and phenolic alcohols.

Elliot and Hart (Elliott and Hart 2009) studied another mixture of model compounds, as shown in Table I-17, but this time over Pd/C catalyst and at a temperature range of 150 °C – 300 °C. From this work, the authors concluded that while Ru catalysts are suitable for low temperatures, Pd catalysts are more appropriate for high temperatures. At high temperatures, the authors also denoted the conversion of stable oxygenated molecules into smaller hydrocarbons through hydrogenation. Finally, based on the single thermal test (no catalyst), Elliot and Hart (Elliott and Hart 2009) enlightened the production of polymers which, according to the authors, it is originated from the conversion of furfural.

In the same year as the latter work, Fish et al. (Fisk et al. 2009) proposed the upgrading of a synthetic bio-oil mixture (see Table I-17) over a Pt/Al_2O_3 at 300 °C. The authors concluded that light oxygenates can be converted via reforming or via hydrogenation, as shown in Figure I-26. The former generates carbon dioxide (and hydrogen) with the help of water, while the latter produces alkanes (and water) thanks to the hydrogen in the medium. Since Pt catalysts are more selective towards C-C bond cleavage, these experiments showed a small number of alkanes as output. Regarding the aromatic molecules, since their reforming is thermodynamically disfavored, small hydrocarbons are obtained under HDO conditions. Furthermore Fish et al. (Fisk et al. 2009) denoted the formation of polymers through thermal reactions, mainly during the heating process, thermal reactions, and the final reaction temperature, catalyzed reactions.

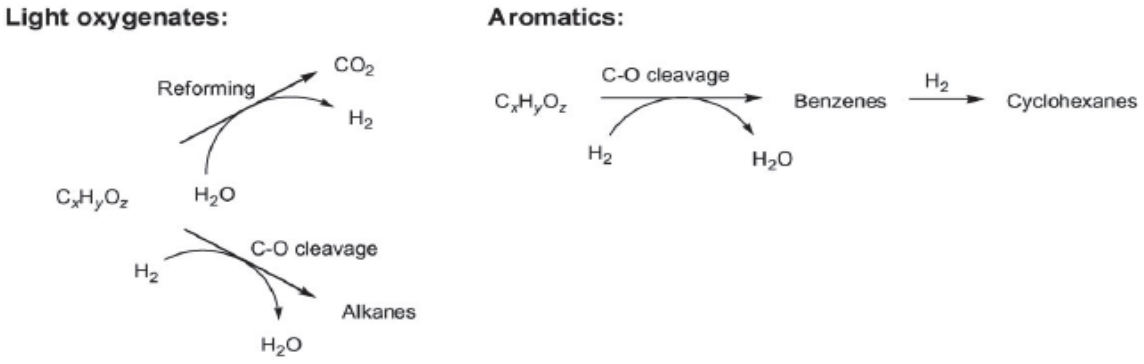


Figure I-26 - Simplified reaction scheme for model bio-oil upgrading over Pt catalysts (Fisk et al. 2009).

Wildschut (Wildschut 2009) studied the conversion of phenol and carbohydrates over a Ru/C catalyst at 250 °C. For these, the author found that phenol leads to the formation of alkanes, while carbohydrates provide insoluble macromolecules. Deeper into the complexity of bio-oils, Wildschut (Wildschut 2009) analyzed three mixtures presented in Table I-17. In the presence of an acidic medium, the effluents would present a high fraction of solids, as shown in Table I-15. As a conclusion, the authors stated that the possible formation of insoluble macromolecules, may generate unstable solids, if their acidic character is not removed. For this reason, the authors finished by recommending the removal of the carboxylic acids and other unstable species by reducing them into alcohols through a stabilization step. As a final remark, the nitrogen test presented in Table I-15 represents the thermal conversion of D-

glucose over a Ru/C catalyst at 250 °C. According to the author, the value of solids yield "(…) is the upper limit of solids formation by the thermal pathway to humins under these conditions.".

Table I-15 - Overview of solids formation at 250 °C (Wildschut 2009).

Reagent	Co-reagent	Gas phase	Catalyst	Solids yield (%-wt.)
D-cellobiose	-	Hydrogen	Ru/C	1.1
D-glucose	-	Hydrogen	Ru/C	1.5
D-glucose	acetic acid	Hydrogen	Ru/C	6.8
D-glucose	acetaldehyde	Hydrogen	Ru/C	0.3
D-glucose	acetone	Hydrogen	Ru/C	1.3
D-glucose	-	nitrogen	Ru/C	43.2
D-glucose	-	Hydrogen	Pd/C	2.6

With a larger focus on the polymerization reactions, Hu et al. (Hu et al. 2013) studied the role of typical bio-oil compounds in the formation of macromolecules over a Amberlyst 70 catalyst and between 90 ºC and 190 ºC. The authors found that carbohydrates have a great impact on the production of polymers, due to their products. According to the authors, when submitted to hydrolysis, levoglucosan is transformed into glucose which, by successive dehydration steps, generates small oxygenated compounds. These are mainly composed by furans, aldehydes and ketones, which are extremely prone to polymerization reactions. Although the three former components are favorable for the production of macromolecules by their own polymerization or by reaction with other compounds, the carboxylic acids take no part in the formation of the polymeric products. In fact, these acids are just the triggers responsible for the maintenance of the medium acidity, as clearly stated by Table I-16. The author discovered that other acidic compounds such as phenols can enhance polymerization reactions. However, since phenols are weaker acids, this phenomenon is less important than for carboxylic acids. Hu et al. (Hu et al. 2013) concluded additionally that, unlike the carboxylic acids, phenols interfere with the polymerization reactions. In fact, the presence carbonyl molecules, increases the phenols reactivity towards polymerization. Furthermore, in the absence of both acids and phenols, the polymerization may yet occur, as the dehydration of glucose generates carboxylic acids. Finally, Hu et al. (Hu et al. 2013) studied the impact of the solvent. As mentioned in section I.2.3.2, the solvent can redirect the reaction network of a compound. Therefore, and as seen by the authors, in a methanol medium the polymerization reactions are reduced, since this alcohol suppresses the dehydration of glucose and, as a result, the formation of the reactive products.

Table I-16 - Polymer formation at different runs (Hu et al. 2013).

	Special conditions of the runs	Polymer to the total reactants (wt.%)
Run 1[a]	With carboxylic acids in water medium	14.7
Run 2[b]	Without carboxylic acids in water medium	3.4
Run 3[c]	Without carboxylic acids and the phenolics	2.3
Run 4[d]	In the methanol medium	4.6

[a]Experimental conditions: Reactants (All the compounds listed in Table 1 plus water): formic acid, acetic acid, levoglucosan, hydroxyl aldehyde, hydroxyl acetone, cyclopentanone, furan, furfural, phenol, guaiacol, and vanillin. Reaction medium: water. Temperature: from 90 to 190°C with 30 min of residence time at each reaction temperature.
[b]Experimental conditions: Reactants (Without the acids): levoglucosan, hydroxyl aldehyde, hydroxyl acetone, cyclopentanone, furan, furfural, phenol, guaiacol, vanillin, and water. Others were same to that in Run 1.
[c]Experimental conditions: Reactants (Without the acids and the phenolics): levoglucosan, hydroxyl aldehyde, hydroxyl acetone, cyclopentanone, furan, furfural, and water. Others were same to that in Run 1.
[d]Experimental conditions: Reaction medium: methanol; Catalyst: Amberlyst 70 (3 wt %); the reactants were all the compounds listed in Table 1 plus methanol. Other reaction conditions were same as that in Run 1.

As the previous authors, Ozagac (Ozagac 2016) (Ozagac et al. 2016a) (Ozagac et al. 2016b) dedicated his thesis to the understanding of the reactions of bio-oil compounds and their interactions over a NiMo/γ-Al_2O_3 catalyst and for 200 ºC, 250 ºC and 300 ºC. According to both authors, when the right set of molecules is hydrotreated together as a mixture, they can represent a similar behavior as a real bio-oil fraction. In terms of mixtures, the author edified a diverse plan of experiments constituted by binary, ternary and as well quaternary mixtures of guaiacol, furfural, glucose, and acetic acid (see Table I-17). From these, Ozagac (Ozagac 2016) (Ozagac et al. 2016a) conclude that mixtures with acid generate more solids, while mixtures with guaiacol produce less solids. The impact of the acids in the proliferation of solids has been already discuss, however the effect of guaiacol never was exposed before. As presented in Figure I-27, in the beginning of the HDO, this compound assumes his role in the conversion process, however, a short while after, guaiacol stops been converted, but is still consumed. According to the author, guaiacol can also react with the macromolecules, thereby justifying its consumption and non-conversion. However, also according to the author, guaiacol can act as an inert molecule.

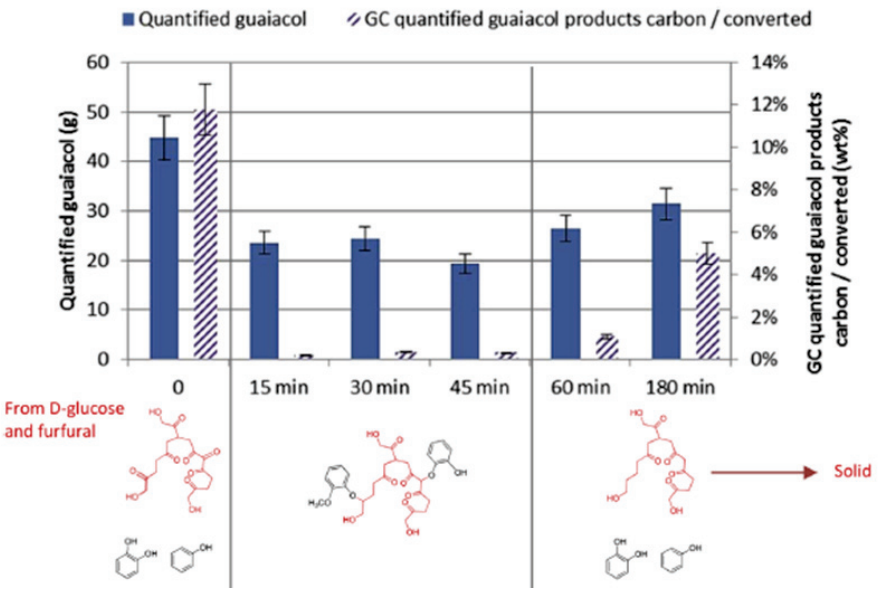


Figure I-27 – Impact of guaiacol on the solubilization of macromolecules (Ozagac et al. 2016a).

Ozagac (Ozagac 2016) (Ozagac et al. 2016a) (Ozagac et al. 2016b), the reaction between guaiacol and the macromolecules inhibits the precipitation of these, but does not eliminate them. In fact, through the formation of a hemiacetal and dehydration reactions, guaiacol joints with the macromolecules turning them soluble. However, this inhibition does not last, as the continuous addition of hydrogen promotes the hydrogenolysis of such components, leading to the release of guaiacol and once again to the precipitation of the macromolecules. At this point, a competition for this molecule is triggered, since it is consumed in the solubility of the macromolecules and as well in the production of the guaiacol derivatives. Figure I-28 summarizes the behavior of guaiacol in the studied quaternary mixture.

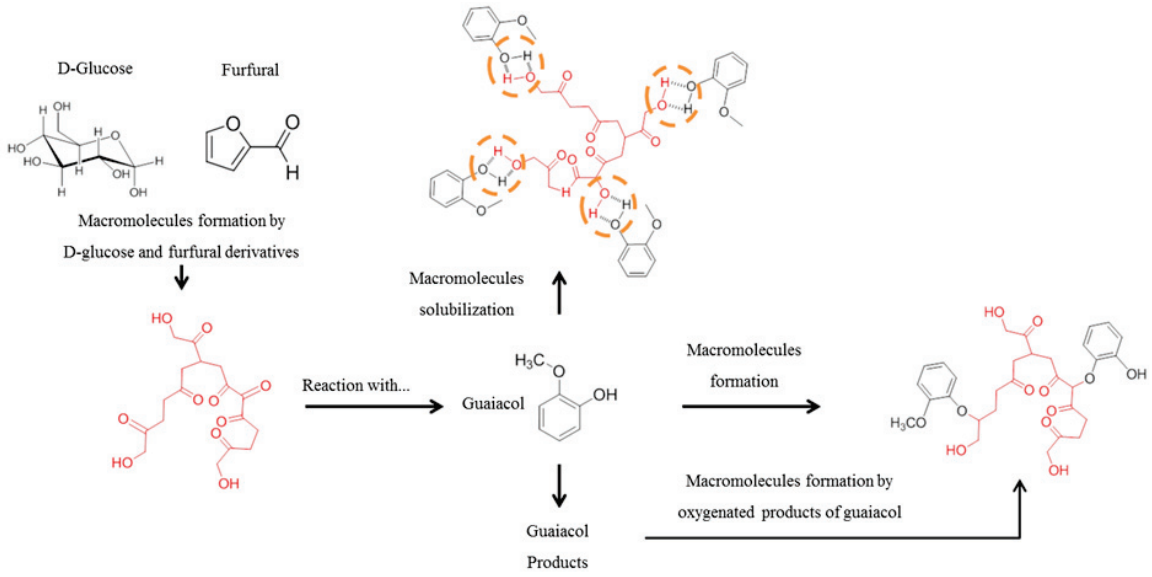


Figure I-28 - Impact of guaiacol in the hydrotreating of the quaternary mixture (Ozagac 2016).

Table I-17 – Studies of mixtures of model compounds.

Authors	**Catalyst**	**Reactants**			
Elliot et al. (Elliott et al. 2006)	Ru/C	1st Mixture: hydroxyacetaldehyde, acetol, eugenol and ethyl guaiacol	2nd Mixture: furfural, guaiacol, homovanillic acid and isoeugenol	3rd Mixture: methyl-cyclopentenone, methyl guaiacol, acetovanillone and oleic acid	
Elliot and Hart (Elliott and Hart 2009)	Pd/C	furfural, guaiacol, acetic acid and water			
Fish et al. (Fisk et al. 2009)	Pt/ Al_2O_3	methanol, acetaldehyde, acetic acid, glyoxal, acetol, glucose, guaiacol, furfural, vanillin and water			
Wildschut (Wildschut 2009)	Ru/C	1st Mixture: D-glucose and acetic acid	2nd Mixture: D-glucose and acetaldehyde	3rd Mixture: D-glucose and acetone	
Hu et al. (Hu et al. 2013)	Amberlyst 70	1st Mixture: levoglucosan, formic acid, acetic acid, furan, furfural, glycolaldehyde, phenol, guaiacol, vanillin, acetol, cyclopentanone and water	2nd Mixture: levoglucosan, furan, furfural, glycolaldehyde, phenol, guaiacol, vanillin, acetol, cyclopentanone, and water	3rd Mixture: levoglucosan, furan, furfural glycolaldehyde, acetol, cyclopentanone, and water	4th Mixture: levoglucosan, formic acid, acetic acid, furan, furfural, glycolaldehyde, acetol, cyclopentanone phenol, guaiacol, vanillin, and water
Ozagac (Ozagac 2016)	NiMo/ γ-Al_2O_3	Binary mixtures: D-glucose, furfural, acid acetic, guaiacol, water and n-hexadecane (Combinations between BO compounds – total of ten mixtures – water variation)	Ternary mixtures: D-glucose, furfural, acid acetic, guaiacol, water and n-hexadecane (Combinations between BO compounds – total of eight mixtures – water variation)	Quaternary mixtures: D-glucose, furfural, acid acetic, guaiacol, water and n-hexadecane (Total of two mixtures – guaiacol variation)	

I.2.4.2 Formation of undesirable molecules

As seen, mixtures of model compounds end mostly with the generation of undesirable compounds via polymerization or other bimolecular reactions. These, when already in the stage of insoluble macromolecules, tend to precipitate either on the catalyst or in the reaction medium. These solid compounds are called coke. The origin of undesirable molecules not always depends on mixtures and their interactions. Simple hydrotreating tests to single molecules may also originate undesirable molecules, since the products molecules may react among themselves, as seen for 5-HMF in (Ozagac et al. 2016b).

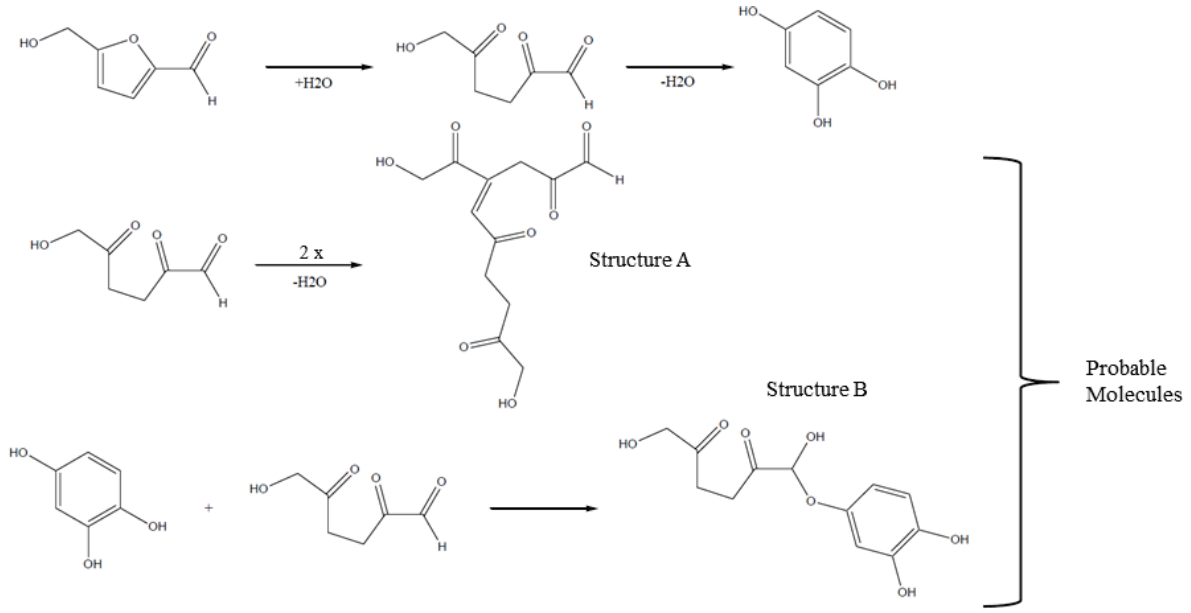


Figure I-29 - Formation of macromolecules from the hydrotreating of 5-HMF over a reduced NiMo/γ-Al_2O_3 catalyst at 200 °C, 250 °C and 300 °C (Ozagac et al. 2016b).

The formation of undesirable molecules depends on several parameters, as the temperature, the pH of the medium and the presence of water (Hu et al. 2013; J. P. Diebold 2000). When exposed to long periods of time under undesirable conditions of any of these characteristics, unwanted reactions tend to propagate. For this reason, storage of bio-oil is extremely complicated and is one of the biggest disadvantages regarding crude oil (J. P. Diebold 2000).

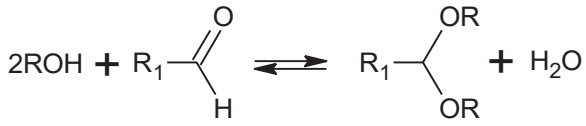
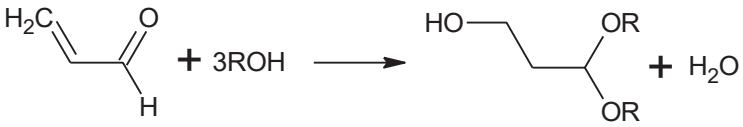
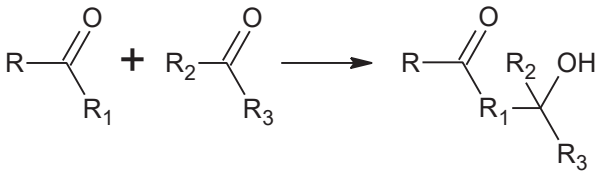
In reactional terms, the increase of temperature promotes degradation reactions as thermal cracking or dehydration (see Table I-18). The generation of small unsaturated molecules (mostly from aromatic rings typical from lignin derivatives) are easily expose to other unwanted reactions as for example Diels-Alder reaction (see Table I-18), suspected of begin the main reaction for the formation of polyaromatics. Gas reactions, as decarbonylation or decarboxylation, are also trigged by heat. Besides forming toxic gases, these reactions lead to carbon loss, reducing the bio-oil energetic performance.

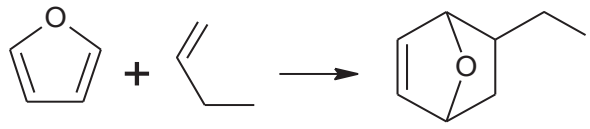
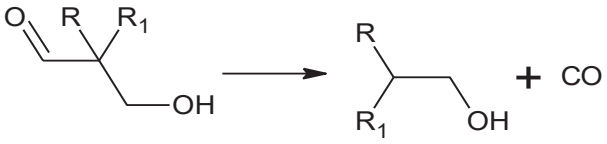
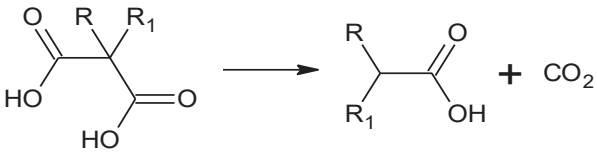
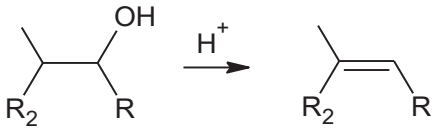
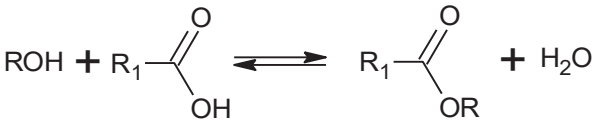
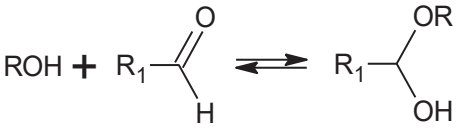
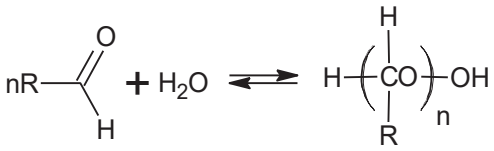
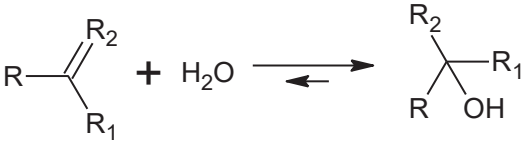
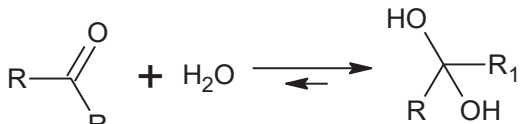
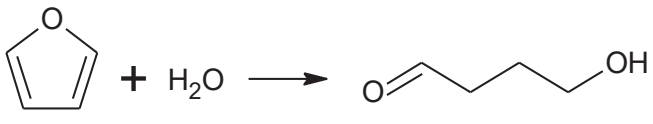
Regarding the pH, under an acid medium, molecules with carbonyls groups can undergo dehydration, esterification, transesterification, aldol or hemiacetal and then acetal reactions (see Table I-18), forming unsaturated molecules as olefins or others with ester and ether groups. Furthermore, reactions between phenols and aldehydes, addition of alcohols and even polymerization of aromatics and olefins (see Table I-18) may generate macromolecules. These, as seen before, may be converted into coke, which not only reduces the carbon content in the liquid, but also reduces the process performance through catalyst deactivation.

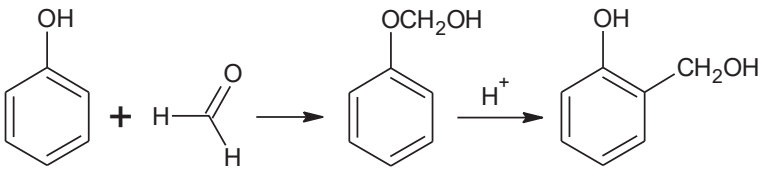
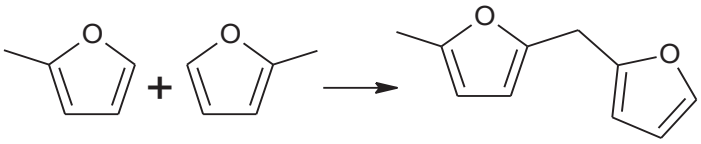
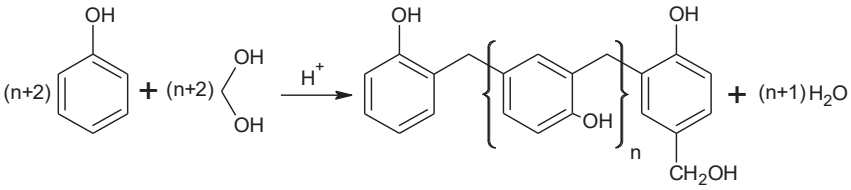
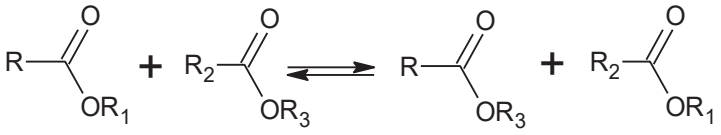
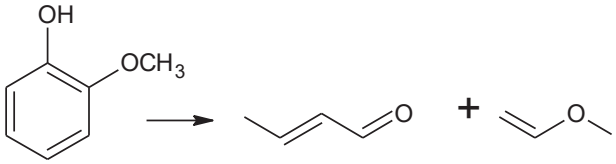
Concerning the water, besides increasing the probability of phases separation, water promotes homopolymerization (see Table I-18) and hydration of unsaturated molecules (see Table I-18) and also carbon-oxygen liaison in rings (as tetrahydrofurans) (Ozagac 2016). These reactions, together with the previous ones, help the formation of macromolecules and undesirably increase the amount of oxygen in the mixture. This last consequence may also originate from the oxidation of bio-oil. When in contact with air, bio-oil ignites the amount of carboxylic acids from aldehydes and alcohols. Furthermore, through autoxidation, certain molecules of bio-oil may be transformed into peroxides. If these are derived of ethers, then besides unstable, bio-oil became explosive.

Despite all, the unwanted molecules, which are still soluble, may be converted into lighter molecules. According to the functional groups present in the molecules, these may be transformed by the reactions given in section I.2.3. Regarding the newly formed functional groups as ethers and esters, the following section will briefly present their pathways under hydrotreating. For peroxides groups, any information has found in the open literature.

Table I-18 – Reaction rules for the formation of undesirable molecules (J. P. Diebold 2000; Solomons and Fryhle 2004).

Name / Type	Reaction
Acetal reaction	
Alcohol addition	
Aldol reaction	

Diels-Alder	
Decarbonylation	
Decarboxylation	
Dehydration	
Esterification	
Hemiacetal reaction	
Homopolymerization	
Hydration	
	
	

Phenol/aldehydes reaction	
Furans polymerization	
Resins formation	
Transesterification	
Thermal cracking	

Ester and Ethers

Concerning ethers, Kirby et al. (Kirby et al. 1996) studied the conversion of dinaphthyl ether over a CoMo catalyst at 300 °C, 350 °C and 400 °C. The authors saw that hydrogenation of the benzene rings and the C-O scission were the two main steps. The breaking between the oxygen and carbon bond generates an alcohol and an aromatic hydrocarbon which will continue the hydrogenation process through the alcohols and aromatics networks until they reach light hydrocarbons.

The scheme proposed by Kirby et al. (Kirby et al. 1996) was later on validated by Artok et al. (Artok et al. 1996), as shown in Figure I-30. The authors also studied the conversion of dinaphthyl dissolved in dihydrophenanthrene over a MoS_3 catalyst at a temperature range of 375 °C – 425 °C. Through the implementation of a first-order kinetics, the authors estimated an activation energy of 75 kJ/mol and 90 kJ/mol for the hydrogenation of the ring and the C-O scission, respectively. Moreover, Artok et al. (Artok et al. 1996) saw as pre-exponential factors 16390 /min and 517622 /min for the same reactions.

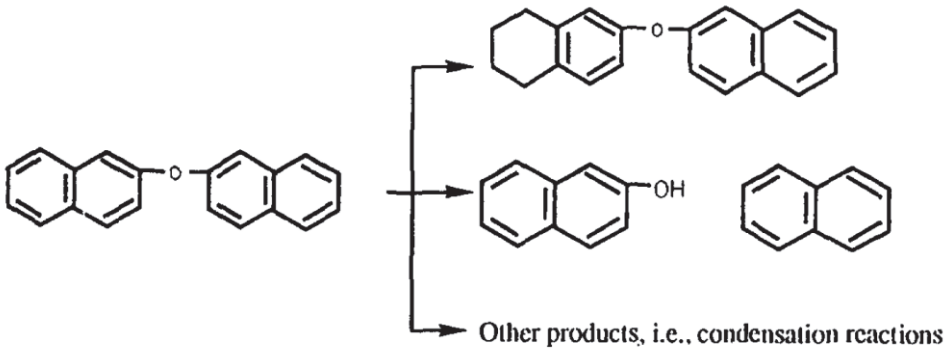


Figure I-30 - Reaction network proposed by Artok et al. (Artok et al. 1996) for the HDT of dinaphthyl over a MoS3 catalyst at a temperature range of 375 °C – 425 °C.

Regarding the esters, Şenol et al. (Şenol et al. 2005) (Şenol et al. 2007) proposed a reaction pathway for methyl heptanoate and methyl hexanoate over NiMo/γ-Al_2O_3 and CoMo/γ-Al_2O_3 catalysts at a 250 °C, 275 °C and 300 °C. The authors proposed two pathways for the conversion of esters. The first is the hydrogenolysis of the carboxyl group generating an alcohol, while the second is the hydrogenolysis of the molecule into a carboxylic acid. The generated molecules then continue to react according to the reaction networks shown in section I.2.3.

Besides the former authors, Bie et al. (Bie et al. 2013) (Bie et al. 2015) also defined a reaction scheme, and even a kinetic model, for the conversion of methyl heptanoate over a Rh/ZrO_2 catalyst at a temperature range of 250 °C – 330 °C. As presented in Figure I-31, heptanoic acid is formed via hydrogenolysis, which has a 10 times higher reaction rate than the production of heptanol. Afterwards, the acid compound is converted into an alcohol through hydrogenation of the carbonyl group into an aldehyde and later on of the hydroxyl group. Furthermore, heptanoic acid can be transformed through decarboxylation into an alkane or once again into an ester by addition of water. Regarding the kinetics, the authors considered the scheme summarized in Figure I-32. For the activation energy, the authors obtained as results a value of 45 kJ/mol for the ester hydrogenolysis, 140 – 180 kJ/mol for the acid transformation into an alcohol or/and alkane and a value of 110 kJ/mol for the hydrogenation of the alcohol into an alkane.

In summary, the presence of an ether group in the hydrotreating process ends in the transformation of its oxygenated bridge into an alcohol which will soon after be converted into an alkane. In the case of esters, a similar procedure is expected. However, due to the presence of the carbonyl group, the oxygenated bridge will produce either carboxylic acids or alcohols. This process may be reinverted leading once again to the formation of an undesirable ester.

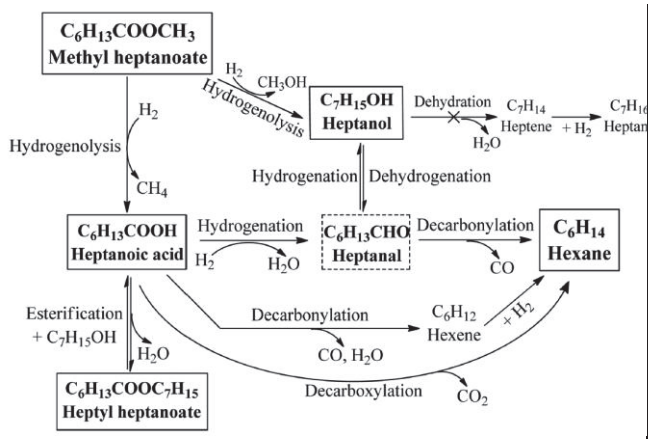


Figure I-31 - Reaction network proposed by Bie et al. (Bie et al. 2015)for the HDT of methyl heptanoate over a Rh/ZrO2 catalyst at a temperature range of 250 °C - 330 °C.

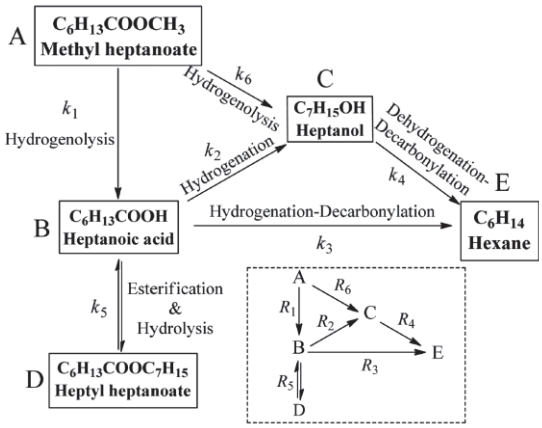


Figure I-32 – Simplified scheme of the kinetic model proposed by Bie et al. (Bie et al. 2015)for the HDT of methyl heptanoate over a Rh/ZrO2 catalyst at a temperature range of 250 °C - 330 °C.

I.2.5 Conclusions for bio-oil to hydrocarbons

Hydrotreating is the most applied process for the removal of oxygen from bio-oils. However, this process is not simple, in particular due to the complex reactivity of the bio-oil mixture. Therefore, in order to avoid polymerization reactions, the bio-oil must first be stabilized at low temperature, following a conversion process at a higher temperature.

Concerning the understanding of the reaction network, several results of hydrotreating experiments are reported, but few kinetic models exist in the open literature. Even so, they are often quite simple and cannot unveil the reactions behind the hydrotreating of bio-oil. For this reason, model molecules were tested as single molecules or mixtures of several model compounds. The first type of studies gives a detailed insight into the mechanisms of specific molecules (and functional groups), whereas studies on mixtures lead to a global characterization of the interactions between model molecules, such as the formation of macromolecules and coke. From all this, it is possible to conclude that the reactivity of bio-oils strongly depends on the solvent, on the catalyst, on the temperature and on the structure of the molecule (amount and positions of functional groups).

The pathways followed by the bio-oil molecules can be represented as a network of reactions. The impact of the previously mentioned characteristics (solvent, catalyst, temperature and structure) then shapes this theoretical exhaustive reaction network into the reality, i.e. going from the reaction pathways to those actual reactions that have a higher activity than others.

Table I-19 presents a brief summary of the possible reactions of bio-oil and its derivatives under hydrotreating conditions. As seen before, hydrogenation, hydrodeoxygenation, saturation, ring opening, direct hydrogenolysis and typical methoxyls reactions, as demethylation, demethoxylation and transalkylation, are the typical (and desirable) reactions for hydrotreating of bio-oil molecules. The remaining reactions of Table I-19 are atypical and undesirable, however usual. These reactions promote the addition of oxygen, the removal of carbon, the production of complex and/or unsaturated molecules, such as olefins and carbonylic compounds, and also the degradation of potential high-value molecules, such as aromatic compounds. These reactions are enhanced by the presence of water, an acid medium and the increase of temperature. As also seen in Table I-19, carbonyl compounds are the most prominent species in the formation of undesirable molecules, thereby justifying the use of a stabilization step in the hydrotreating process of bio-oils (see section I.2.1.2).

Table I-19 – Summary of possible reaction of bio-oil and its model compounds.

Bio-oils Family	Typical hydrotreating reactions	Undesirable reactions
Alcohols	Hydrodeoxygenation **(1)**	Dehydration, Dehydrodeoxygenation, Retro-aldol, Alcohol addition, Resins formation **(5)**
Aldehydes & Ketones	Hydrogenation, Hydrogenolysis **(2)**	Decarbonylation, Hydration, Alcohol addition, Acetal, Aldol and Hemiacteal reactions, Homopolymerization, Phenol/Aldehydes reaction **(8)**
Carbohydrates	Hydrogenation, Ring opening **(2)**	Dehydration, Retro-aldol **(2)**
Carboxylic acids	Hydrodeoxygenation **(1)**	Decarboxylation, Esterification **(2)**
Ethers & Esters	Hydrogenation, Hydrogenolysis **(2)**	Hydration, Transesterification **(2)**
Furans	Saturation, Ring opening, Hydrogenation **(3)**	Hydration, Diels-Alder, Furans polymerization **(3)**
Phenols	Saturation, Demethylation, Demethoxylation, Transalkylation, Hydrodeoxygenation **(5)**	Methylation, Diels-Alder, Phenol/Aldehydes reaction, Resins formation, Thermal cracking **(5)**

I.3 Modeling of Complex Mixtures

As seen before, bio-oil is a complex oxygenated mixture. The vast number of distinct molecules, along with their numerous reactions, complicate the understanding of bio-oil and their upgrading processes. When analyzing bio-oil as a whole, any type of experimental test or even modelling technique is reduced to imprecise information and lumped models, respectively. Even through dedicated studies of typical

bio-oil molecules and mixtures of these, the development of a hydrotreating process to the overall feed is complicated.

Be that as it may, modeling complex mixtures, such as bio-oil, is not an impossible task. As already applied to heavy oil fractions, there are advanced modeling techniques which may be implemented. These tools not only help in the development of the process, but also in the understanding of the molecular composition and reactional behavior of the complex mixtures. The following section will present a summary of several methods that were successfully applied to complex mixtures. The aim of the section is to understand how to molecularly describe a feedstock, how to create a kinetic model for an overall reaction network and finally, how to simulate this network with less computer effort. For more information regarding the modeling of complex mixtures, the following review is recommended (Oliveira et al. 2016).

I.3.1 Molecular reconstruction techniques

A crucial step for the creation of kinetic models is the knowledge of its reactants. Usually, analytical data is able to provide molecular information about the feed. However, when trying to analyze complex mixtures, such task becomes extremely demanding, especially due to the lack of analytical information. As a consequence, the scientific community proposed several molecular reconstruction techniques, which can overcome the lack of experimental data. These compositional models can generate the composition of a given mixture, together with their molecular detail.

In the open literature, there is no information available concerning the bio-oil mixtures. Therefore, the following section will present and discuss the most significant reconstruction methods so far applied to complex mixtures. These techniques were divided in three groups according to the approach used to represent the molecular composition: (i) approach by model molecule, (ii) deterministic molecular reconstruction approach, and (iii) stochastic molecular reconstruction approach.

I.3.1.1 Approach by model molecule

The approach by model molecule was the first method to compositional models for the heavy cuts of petroleum. The main idea is to generate a single model molecule which, through its average properties, is able to represent the entire mixture.

Initially created for the study of coal, these methods began to be applied for the oil fractions, as seen in the work of Speight (Speight 1970) for the asphaltenes. Afterwards, these methods undergo several developments, through the time, with the addition of new analytical techniques and further improvement (Hirsch and Altgelt 1970) (Takegami et al. 1980) (Suzuki et al. 1982) (Sato 1997) (Artok et al. 1999) (Faulon et al. 1990) (Faulon 1991) (Faulon 1994).

Regarding the bio-based models, a similar concept was created by Freudenberg and Neish (Freudenberg and Neish 1968) for the reconstruction of the lignin. In fact, the scheme edified by the author does not consist in a chemical structure, but in a schematic assemble of the types of bonds present in the lignin. This work was further continued by Sarkanen and Ludwig (Sarkanen and Ludwig 1971), Glasser et al. (Glasser et al. 1981), Medeiros (Medeiros, D. C. M. 2013) and Rita (Rita, A. I. B. 2014). Figure I-33 is an example of the reconstruction of lignin done by Rita (Rita, A. I. B. 2014). The presented molecule aims to represent all the structure of lignin.

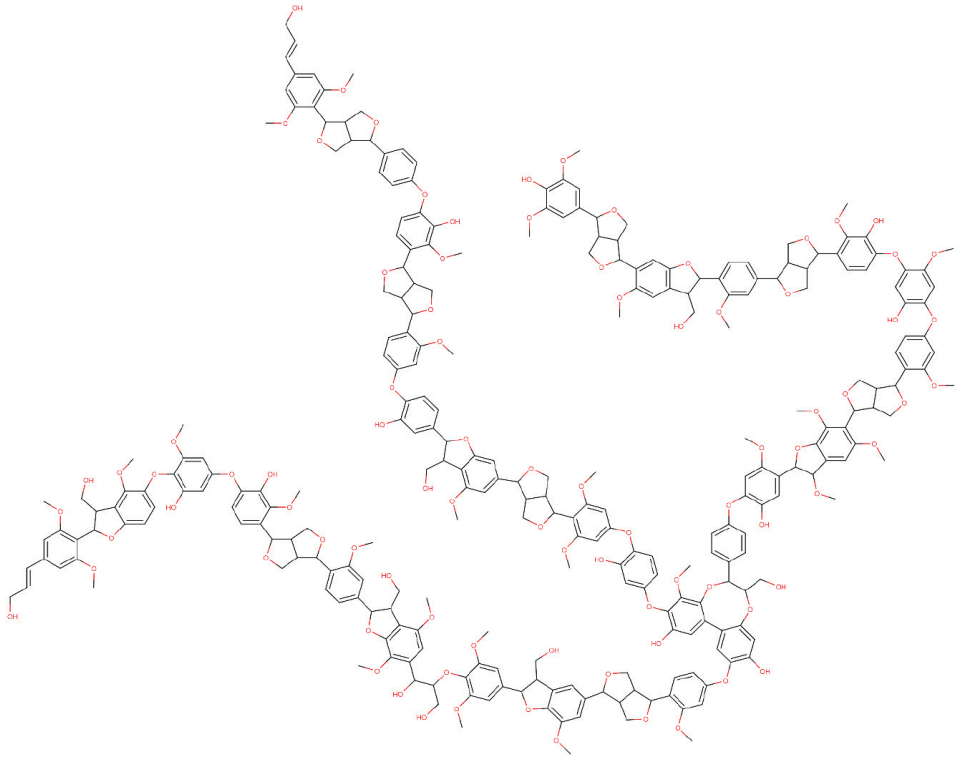
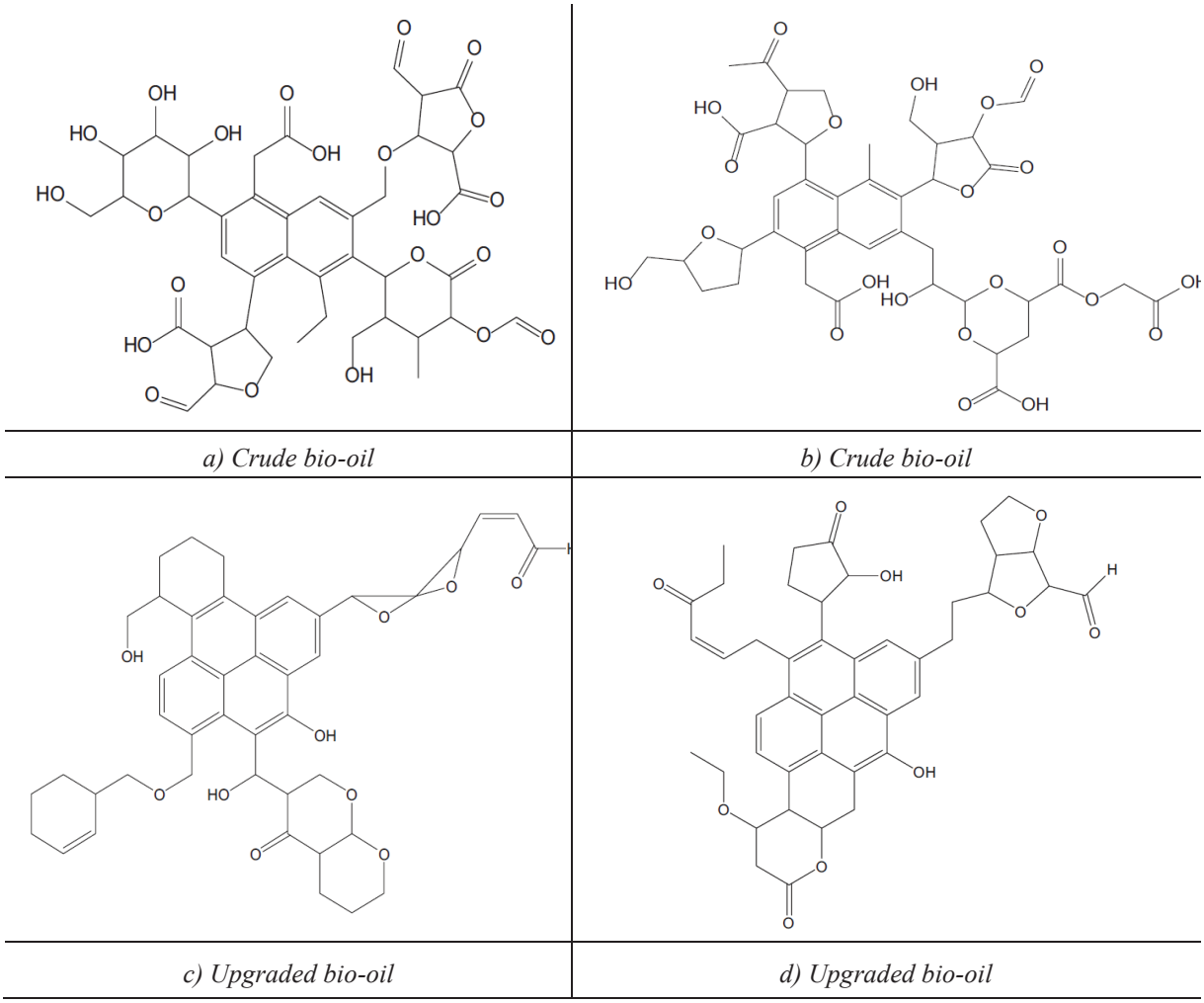


Figure I-33 – Proposed structure for lignin (Rita, A. I. B. 2014).

More recently, at IFPEN, López-Abelairas (paper to be submitted) model the structure of lignin, through the methodology of Klein (Klein et al. 2006) (Hou et al. 2010). In this technique, lignin is reconstructed thanks to pre-defined molecule blocks, which are associated to probability distribution functions obtained by the experimental data. Through the sampling of this pseudo-compounds and then assembling, it is possible to achieve a single molecule representative of lignin.

As mentioned before, there is no information available concerning the bio-oil mixtures. However, Xu et al. (Xu et al. 2014) created a molecular model for the phenolic oligomers in crude and upgraded bio-oil. In the paper, the authors generated structure models of asphaltenes thanks to the Brown-Ladner method. This approach estimates several average molecular structure parameters, as aromaticity, mass and atomic ration of C/H and H/C, respectively, in the alkanes, total number of rings and number of aromatic and saturated rings. By comparing with the average molecular weight, the FT-IR and ^{1}H and ^{13}C NMR

for both crude and upgraded bio-oil, Xu et al. (Xu et al. 2014) were able to predict were able to predict two structures for both samples. The results are present in Figure I-34



a) Crude bio-oil | *b) Crude bio-oil*

c) Upgraded bio-oil | *d) Upgraded bio-oil*

Figure I-34 – Structure models of asphaltenes: a) and b) crude bio-oil, c) and d) upgraded bio-oil (Xu et al. 2014).

Nevertheless, this simple approach is associated to several flaws, as the poor representation of the polydispersity of complex mixtures. Therefore, this method is not an adequate tool for the reconstruction of the pyrolytic oil or any others as well complex mixtures.

I.3.1.2 Deterministic molecular reconstruction approach

The deterministic molecular reconstruction strategy consists in adjusting the molar fractions of a pre-defined synthetic mixture until the convergence of the properties of both real and synthetic systems. In summary, these types of methods follow a creation line where first a pre-defined database must be constructed, an objective function based mostly on analytical constraints must be defined; and finally, the molar fractions are calculated.

From the several existing methods, it is important to highlight the approach proposed by Allen and Liguras (Allen 1991) (Allen and Liguras 1991), on the catalytic cracking process. This methodology

selects a set of predefined molecules and modifies their respective molar fractions in order to obtain a mixture whose properties are closer to the desired analytical data.

Based on a different perspective, Quann and Jaffe, (Quann and Jaffe 1996), created a method, designated of Structure Oriented Lumping (SOL), which consists in the representation of the oil fractions through vectors of structural blocks. This method consists in the idea where each molecule in the mixture is described by a vector of twenty-two elements each one representing a structural block of a molecule. Later, Jaffe et al. (Jaffe et al. 2005) proposed to extend the vectors size from twenty-two elements to twenty-four, allowing the representation of multicore species.

Based on the previous method, Peng (Peng 1999) and Zhang (Zhang 1999) created a procedure, named Method Type and Homologous Series (MTHS). This method characterizes a petroleum cut through a matrix of pseudo-compounds, presented in Figure I-35, classified by chemical families, as columns, and carbon number, as lines. Each matrix cell represents then the fraction of a given family specific that number of carbons. This method was as well extended (Gomez-Prado et al. 2008) by modifying the number of carbons to normal boiling points and also by detailing the families.

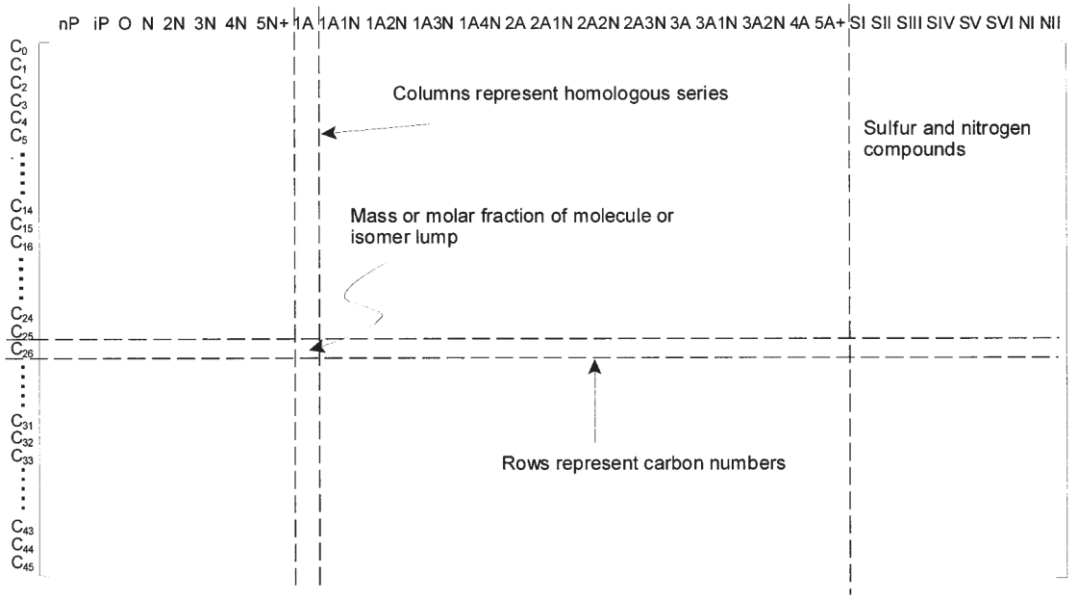


Figure I-35 - Matrix of pseudo-components (Hu et al. 2002).

In the same perspectives as the MTHS, two new methodologies were created. Both the method of Pyl et al. (Pyl et al. 2011) and the statistical reconstruction method edified by (Hudebine 2003) (Hudebine et al. 2011; López García et al. 2010) intend to represent the petroleum cuts via a matrix of chemical families versus molecules size. Furthermore, and thanks to probability distribution functions (PDF), the authors were able to retrieve the molecular fractions. To increase the fitting performance, both authors added an optimization loop. Figure I-36 briefly describes the statistical method.

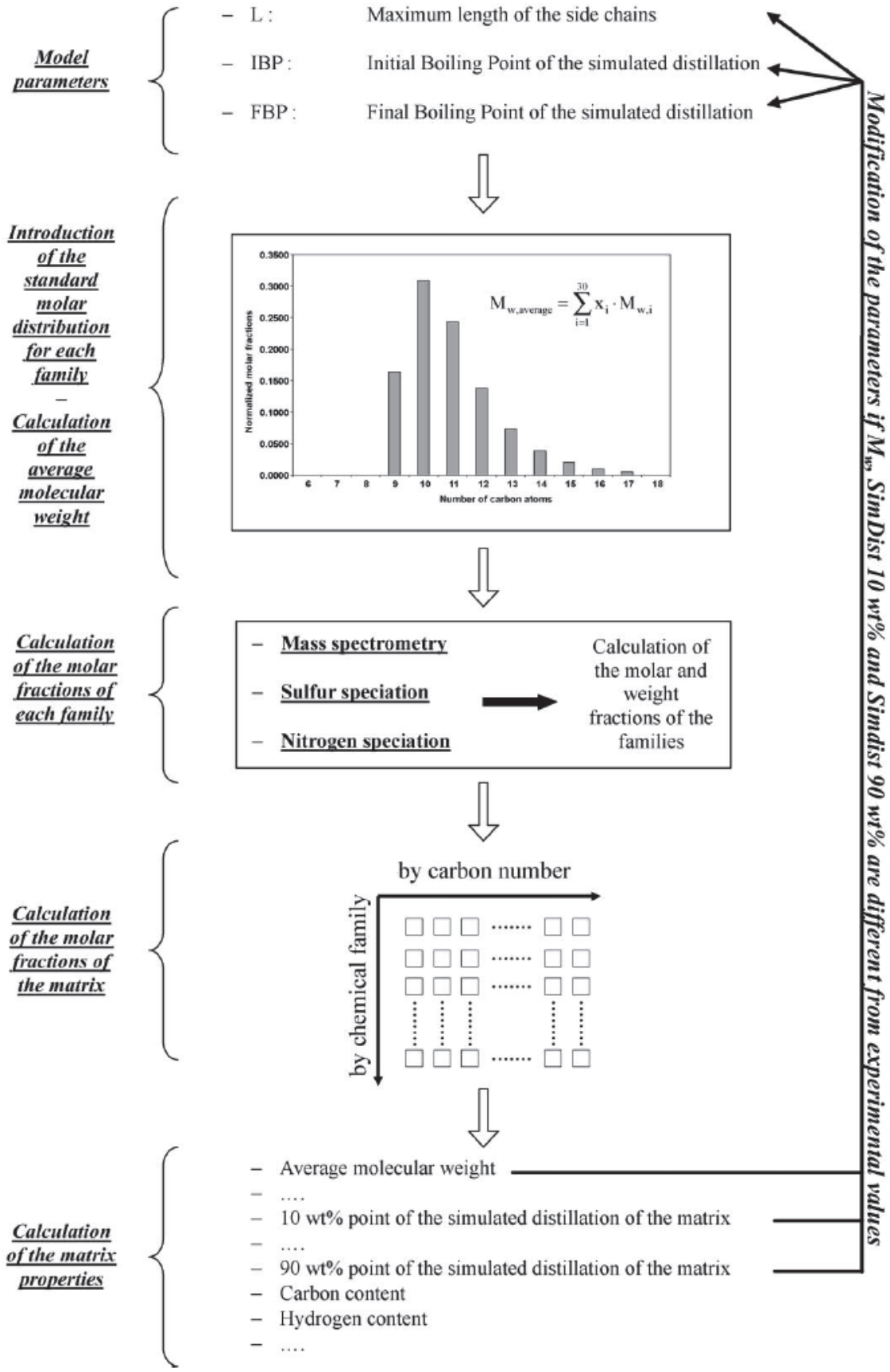


Figure I-36 - Diagram of the statistical reconstruction of gas oils (López García et al. 2010).

Hudebine et al. (Hudebine 2003) (Hudebine and Verstraete 2004) (Hudebine and Verstraete 2011) (Van Geem, K. M. et al. 2007) proposed another method, termed Entropy Maximization (EM). Although both share the same aim, the new approach uses a criterion to retrieve the molecular fractions, as in the method by Allen and Liguras (Allen 1991) (Allen and Liguras 1991). In the case of the EM, the fractions are obtained through the maximization of an information entropy criterion, Equation I-1, also known as the Shannon Criterion (Shannon 1948). In order to improve the fractions adjustment, the author added to this criterion, restrictions which represent not only the analytical constraints, but also the concept of normalized fractions. Furthermore, the author applied as well to each constraint, a Lagrange multiplier reducing the complexity of the problem by changing the resolution for just J Lagrange parameters (J constraints), instead of the N molecular fractions (N molecules). Through the maximization of a conjugate gradient method, these parameters can be obtained, as well as the desired mole fractions.

$$H(x_i) = -\sum_{i=1}^{N} (x_i \cdot \ln(x_i)) \qquad \textit{Equation I-1}$$

In conclusion, all the deterministic approaches mentioned are able to construct a synthetic mixture whose properties match those of the feedstock. Since these techniques are based on a vast library of pre-defined molecules and on constraints supported by analytical data, they are considered the most efficient. However, when faced with complex mixtures, these methods fail since not only the molecular library becomes extremely time consuming and difficult to build, but also because the analytical information is insufficient or even inexistent.

I.3.1.3 Stochastic molecular reconstruction approach

The stochastic molecular reconstruction (SMR) approach focuses on the modeling of a mixture composition by assuming that any molecule can be represented by an ensemble of molecular attributes (e.g. the number of rings, chains length, etc.). For one of these attributes, a probability distribution function is adjusted in order to represent the probability of finding the given attribute in the given molecule. In the end, the probability of a specific molecule is the product of the probabilities of the attributes which compose the given molecule.

The present approach started with the work of Boduszynski, (Boduszynski 1987) (Boduszynski 1988). The structure of a petroleum cut can be characterized by statistic distributions which vary according to their boiling point, molecular weight and number of carbons. Taking such in consideration, Neurock and Trauth, led by Klein (Neurock 1992) (Trauth 1993) (Neurock et al. 1994) (Trauth et al. 1994), proposed a method, designated of stochastic reconstruction, for the characterization of the oil fractions. As a result, the team was able to generate a synthetic mixture based on a Monte Carlo procedure and on the referred probability distributions functions.

To describe asphaltenes, Neurock et al. (Neurock 1992) (Neurock et al. 1994) declared seven attributes along with their PDFs, based on the molecule model works (Hirsch and Altgelt 1970) (Speight 1970) (Boduszynski 1988): polycyclic cores; naphthenic cycles; aromatic cycles; substitutions in the naphthenic cycles; substitutions in the aromatic cycles; length of the aliphatic chains in the naphthenic cycles and finally, length of the aliphatic chains in the aromatic cycles. Besides, and in order to generated a coherent molecule and consequently mixtures, a decision tree, named building diagram, was edified. The construction of a molecule is done according to this diagram and the random sampling of the probability distributions functions associated to each branch of the diagram, as shown in Figure I-37.

Nevertheless, even if the molecule models are constantly applied to support the attributes definition, synthetic mixture will never be a true reflection of the real one. To overcome this shortcoming, Trauth et al., (Trauth 1993) (Trauth et al. 1994), proposed the addition of a optimization loop that modifies the PDF parameters until the properties of the synthetic mixture match those of the feedstock. To apply this new methodology, the PDFs stop being statistical distributions and were converted into mathematic laws, as Gaussian, gamma, exponential and etc. distributions. However, due to the rigorous building algorithm, the team lead by Klein still continues the improvement of this method through the application to several feeds as oil fractions (Horton et al. 2013) (Campbell et al. 2009) (Zhang et al. 2014), biomass (Horton et al. 2015a; Horton et al. 2015b; Moreno 2014) and also to plastics (Horton et al. 2015c). Nowadays, their stochastic reconstruction technique, designated by Composition Model Editor (CME), can predict the compositions in addition to the molecules.

In addition to the previous developments, Hudebine et al. (Hudebine 2003) (Hudebine and Verstraete 2004), proposed to join the stochastic reconstruction method and the reconstruction by entropy maximization procedure (Hudebine 2003) (Hudebine and Verstraete 2011), in order to suppress the drawbacks of both techniques, as illustrated in Table I-20.

Table I-20 - Comparison between the SR and REM method (Hudebine and Verstraete 2004).

SR		**REM**	
Advantages	**Disadvantages**	**Advantages**	**Disadvantages**
• Good match between the results and the analytical data; • The reconstructed molecules are always typical of studied fractions, since this process uses expert knowledge and specific chemical rules.	• Enables to fit into all analysis, due to intrinsic constraints of the method; • Monte Carlo technique is computationally intensive.	• Small computational effort; • Obtains a mixture whose properties are very close to the analytical data.	• Strongly depends on the initial set of molecules; • Difficult to find a solution when the properties of the initial set of molecules are too far from the analytical data; • Very sensitive to inconsistencies in the analytical data.

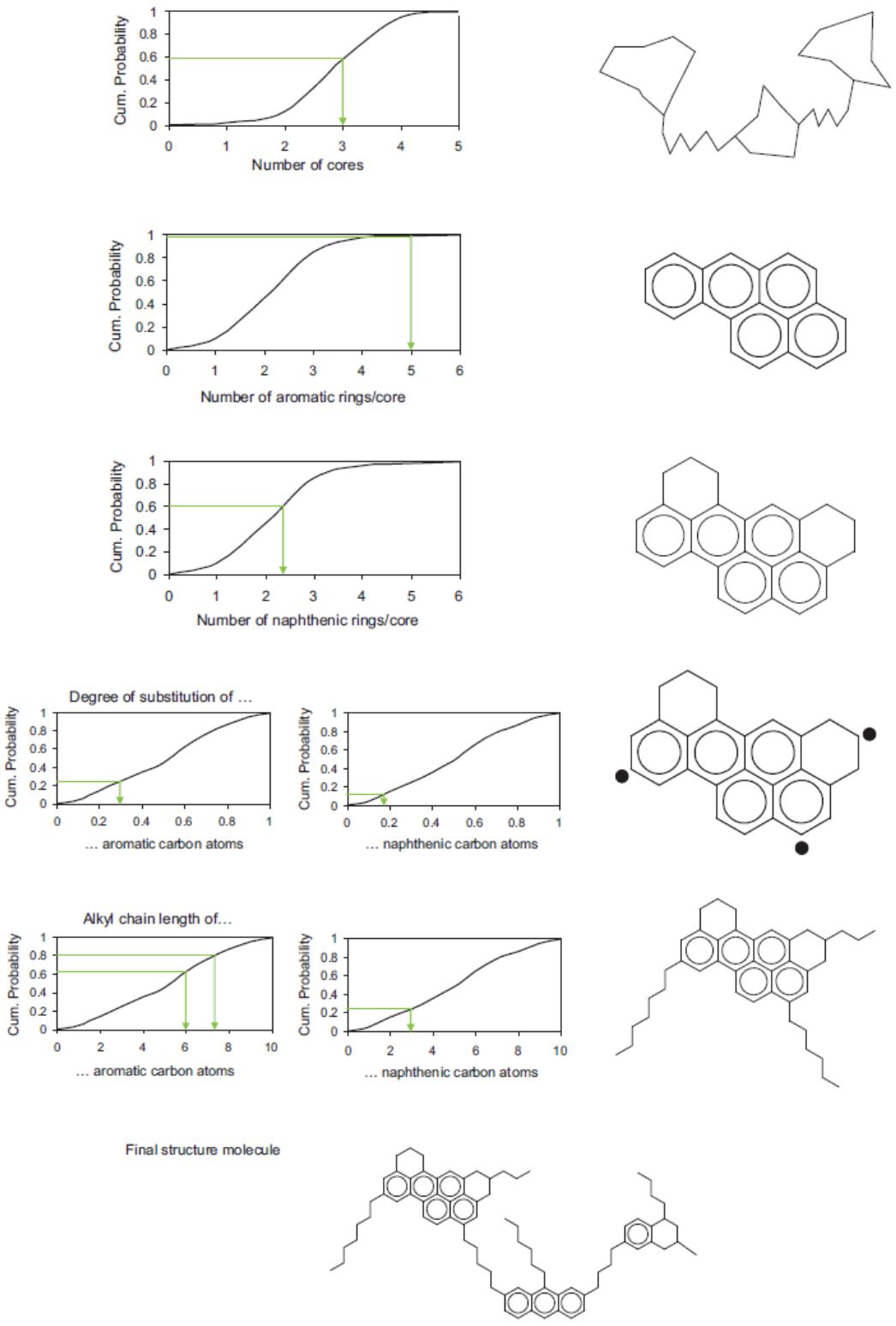


Figure I-37 - The logic of the stochastic algorithm for assembling molecular asphaltenes structures. The hierarchy of the construction process involves determining and assembling (A) the unit sheets which make up each molecule, (B) the aromatic core of each unit sheet, (C) the naphthenic rings of the polycyclic core, (D) the number of aliphatic substituents, (E) and the length of each substituent. The final structure of the assembled molecule is illustrated in part F (Neurock et al. 1994).

As shown in Figure I-38, the two-step methodology starts by the SR step, where a set of equimolar molecules are generated. Afterwards, by the REM, the molar fractions of the synthetic mixture are adjusted until these converge into those of the feedstock. This technique was later validated by IFPEN for the light cycle oils (Hudebine and Verstraete 2004) (Oliveira et al. 2012a) (Lopez-Abelairas M. et al. 2016), vacuum gas oils (Verstraete et al. 2004) (Charon-Revellin et al. 2011) and also vacuum residues (Verstraete et al. 2010) (Oliveira et al. 2013a) (Oliveira et al. 2013b) (Oliveira et al. 2014) (Oliveira et al. 2013c). Besides these authors, also Alvarez-Majmutov et al. (Alvarez-Majmutov et al. 2014; Alvarez-Majmutov et al. 2015) applied the two-step SR-REM technique to middle and vacuum distillates. Thanks to all these studies, the authors proved the method flexibility, expanding the reconstruction of further complex mixtures.

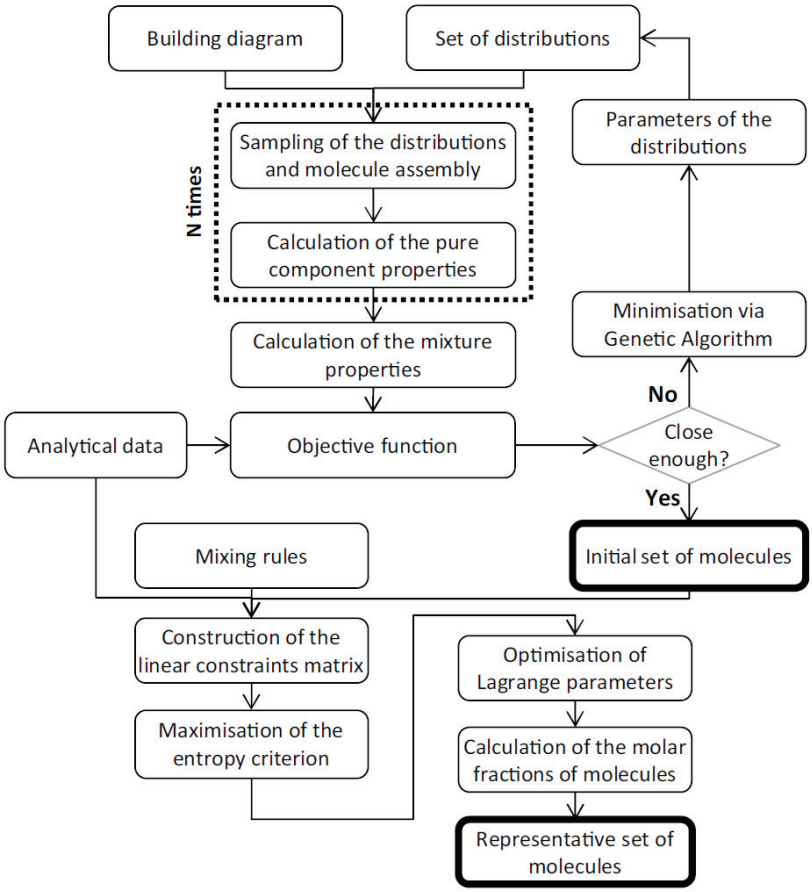


Figure I-38 - Flowchart of the SR-REM algorithm (Lopez-Abelairas M. et al. 2016).

Based on the stochastic reconstruction method, Sheremata et al. (Sheremata et al. 2004), created a method designated Quantitative Molecular Representation (QMR). Originally developed for heavy oil fractions, such as asphaltenes, this method also assembles attributes through a Monte Carlo procedure. However, unlike SMR, the molecular attributes are more detailed, assuming a configuration of structures (see Figure I-39). For example, in the case of asphaltenes (Sheremata et al. 2004), instead for sampling

the number or cores, number of rings, etc., the QMR samples structural groups as naphthalene, dibenzothiophene, etc. and the possible connections between these. After generating a library of molecules, the algorithm activates an optimization loop which will select the best molecules according to the experimental data. The Quantitative Molecular Representation methodology have been since then improved and developed by other authors as Boek et al. (Boek et al. 2009) and Iwase et al. (Iwase et al. 2018).

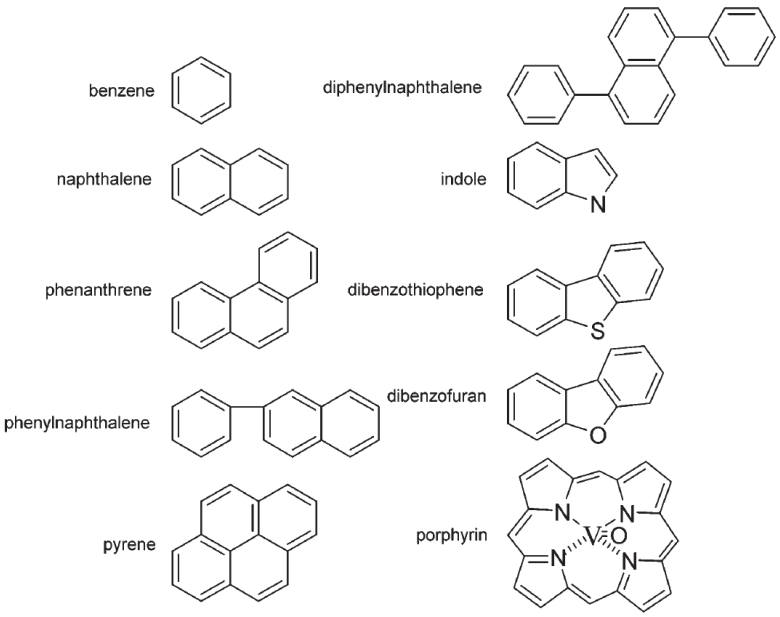


Figure I-39 - Aromatic groups present in the asphaltenes representation (Sheremata et al. 2004).

I.3.2 Reaction network generation techniques

The difficulty in generation of a reaction network strongly depends on the chemical system under question. If this is simple, its network is relatively straight forward. However, confronted with a complex mixture, the generation of the network is no longer an easy task. Complex mixtures are usually composed by a vast number of distinct molecules, which have different reactivities and consequently innumerous reactions.

According to (Ugi et al. 1993), network generation algorithms can be classified as empirical, semi-formal and formal methods. Although this classification is not really applied nowadays, it helps organizing the state of the art related to kinetic modeling, and therefore the understanding of the topic.

The empirical methods are based on pre-established reaction libraries and expertise. This technique is mostly used for relatively simple chemical systems largely studied by the scientific community. Although quite simple to apply, these methods not always describe correctly the behavior of the reaction in the medium (unsuitable reaction network). The semi-formal method is based on heuristic algorithms that generate reactions just by knowing basis reaction rules. This technique can be view as an algorithm,

which knowing certain reaction rules, extrapolate them for the entire chemical system without any limit. For this reason, this approach will consider a large number of reactions even if some are improbable. The formal method generates a reaction network at the mechanistic level (elemental steps) by account only with physical and chemical description of the molecules. Thanks to such molecular precision, this technique can create a detailed reaction network without the improbable reactions.

As it will be further on presented, the generation of a reaction network is related to the type of desired kinetic model. Higher precision from the model will equally require a more precise network. The inverse is also valid. However, a precise reaction network can always be reduced to a less detail model. For the reduction of reaction network, especially those innumerous reactions, there are several techniques that can be used as presented by Oliveira et al. (Oliveira et al. 2016). Detailed reduction is the most applied technique. By identifying and removing non-contributing reactions, this method can decrease the level of complexity of any given network.

I.3.2.1 Automated network generation

As mentioned before, there are algorithms that can create reaction networks and the respective kinetic model automatically and *in silico*. Designated as automated network generators, these can recreate the behavior of the reactions with more or less detail according to their application. For example, InGen (Bennett 2009; Wei et al. 2008) is a tool which presents a vast library of reactions. Therefore, this algorithm can generate reaction networks with the detailed desired by the user. Furthermore, an automated network generator is extremely useful for chemical systems with a vast reaction network. Once *in silico*, these tools can generate massive networks in a small period of time without taking much from the user.

Besides InGen (Bennett 2009; Wei et al. 2008), there are other algorithm which can generate networks, as Genesys (Vandewiele et al. 2012), RING (Rule Input Network Generator) (Rangarajan et al. 2012), etc. All the algorithms here presented were already applied and validated for several chemical systems. Although none of these applications was bio-oil, InGen (Horton et al. 2015a; Moreno 2014) and RING (Rangarajan et al. 2010, 2014) were already used for the generation of a reaction network for the conversion of biomass.

I.3.3 Kinetic modeling techniques

There are three main types of kinetic modeling strategies. Each type is more suitable for a specific chemical system than others. As it will be presented in the following section, kinetic models may be divided into lumped, molecular or mechanistic models. The latter type is the most detailed at the reaction level, whereas the first one is the least detailed. As consequence, lumping is very often applied to more

complex feedstocks, such as bio-oil (see section I.2.2.1), while mechanistic models are more easily implemented for smaller chemical systems.

As mentioned before, and as will be seen in the following section, there is a connection between the reaction network and kinetic models. Both are intricately related, as a high precision from a model requires a precise reaction network, whereas the opposite is not needed.

I.3.3.1 Lumped models

Lumped models, as the name indicates, consist in grouping (lumping) molecules with the same properties into families, or lumps. Therefore, in this type of models, the reactivity of each molecule is somehow neglected, meaning that these models are insensitive to the chemical impact. In the end, the reaction network associated to lumped models is represented by families of reactants and families of products. The reactions consist in global transformations between lumps. For each of these transformations, a first-order rate law is often used, although lumped models can also include thermodynamic and physical phenomena, such as external mass transfer, adsorption and/or equilibrium constraints, and include Langmuir-Hinshelwood and other rate equations. The estimation of the pre-exponential factor and the activation energy is performed based on the experimental data available.

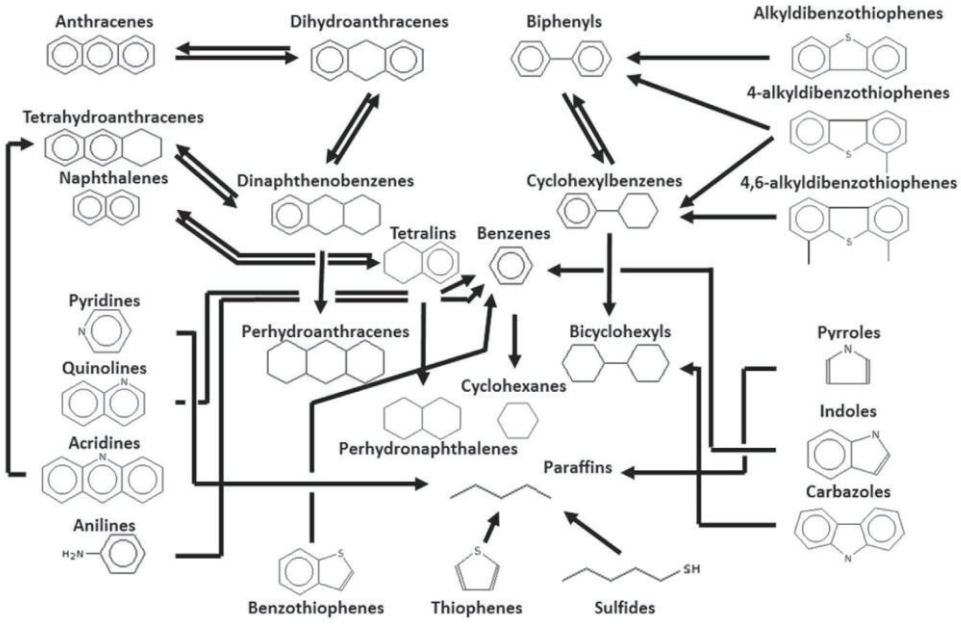


Figure I-40 - Reaction scheme of gas oils hydrotreating for each lump (28 chemical families by n carbon number up to C_{30}) (Oliveira et al. 2016).

Figure I-40 shows a the reaction network behind a lumped model for the hydrotreating of gas oils (López García et al. 2010). In total, the model has 28 chemical families that are distributed over 30 carbon numbers, resulting in 597 lumps. For the model, López García et al. (López García et al. 2010) considered 14 reaction types related to hydrogenation/dehydrogenation, hydrodesulfurization and

hydrodenitrogenation pathways. Thanks to the inclusion of 24 different types of industrial gas oil,s the authors were able to create a robust and performant model.

To conclude, although lumped models neglect the specific reactivity of individual molecules, giving them fewer chemical insights, when well created, as the example below, these types of models may be extremely useful.

I.3.3.2 Mechanistic models

Mechanistic models, unlike lumped models, are created thanks to the molecules and their reactivity. These take into account in their creation the chemical mechanism, alongside the kinetic theories. As a consequence, these types of models are usually described at the radicals and intermediates level by elemental reactions, as radical recombination, protonation/deprotonation, etc. Each reaction is then associated to a set of kinetic parameters. The estimation of these can be done not only through available experimental data, but also thanks to theoretical approaches, as ab initio calculations, single events or Quantitative Structure Reactivity Correlations (QSRC). These last approaches are extremely interesting, as they allow the reduction of the parameters to be estimated by experimental data. However, these are not always ease to apply to a chemical system.

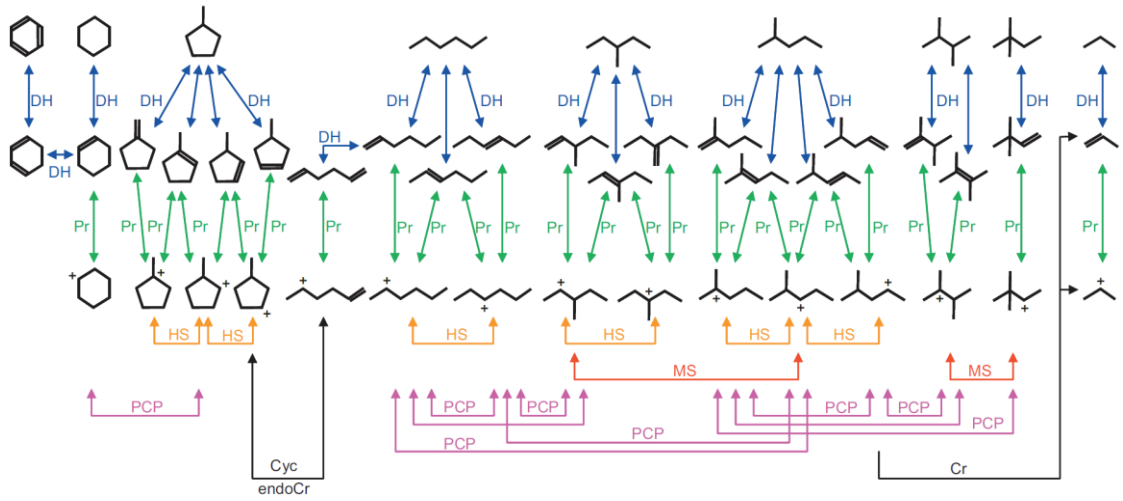


Figure I-41 - Reaction network for the catalytic reforming of n-hexane (without accounting for hydrogenolysis reactions) (Oliveira et al. 2016).

Figure I-41 shows a the reaction network behind a mechanistic model (single events type) for the catalytic reforming of n-hexane (Surla et al. 2011). For the model, 5 different transformations were considered: Protonated Cyclo Propane (PCP), protonation (Pr), hydride shift /methyl shift (HS/MS), dehydrogenation (DH) and cracking (Cr). For the first three isomerization reactions, 5 rate coefficients were considered, whereas for the two last cracking reactions, one rate coefficient was used for each reaction. In the end, the authors (Surla et al. 2011) obtained suitable model for the catalytic reforming of n-hexane.

To conclude, although mechanistic models allow the development of a detailed kinetic model, these are not always suitable as modeling approaches, in particular for complex mixtures with several distinct molecules and reactions.

I.3.3.3 Molecular models

Molecular models are an intermediate technique between lumped and mechanistic models. Similar to the last type models, the molecular approach describes the reactants and the products as molecules and not families. However, these molecules are later on grouped into families, which unlike the lumped models, are related by the molecules structures and reactivity instead their properties. In the end, the reaction network associated to molecular models is represented by transformations between reactants and products, which may be grouped in molecular reactivity families. Each of these reactions is then associated to a reaction rate and usually a global kinetic constant. The kinetic parameters of the model can be estimated by experimental data or by Quantitative Structure Reactivity Correlations (QSRC), as the mechanistic models.

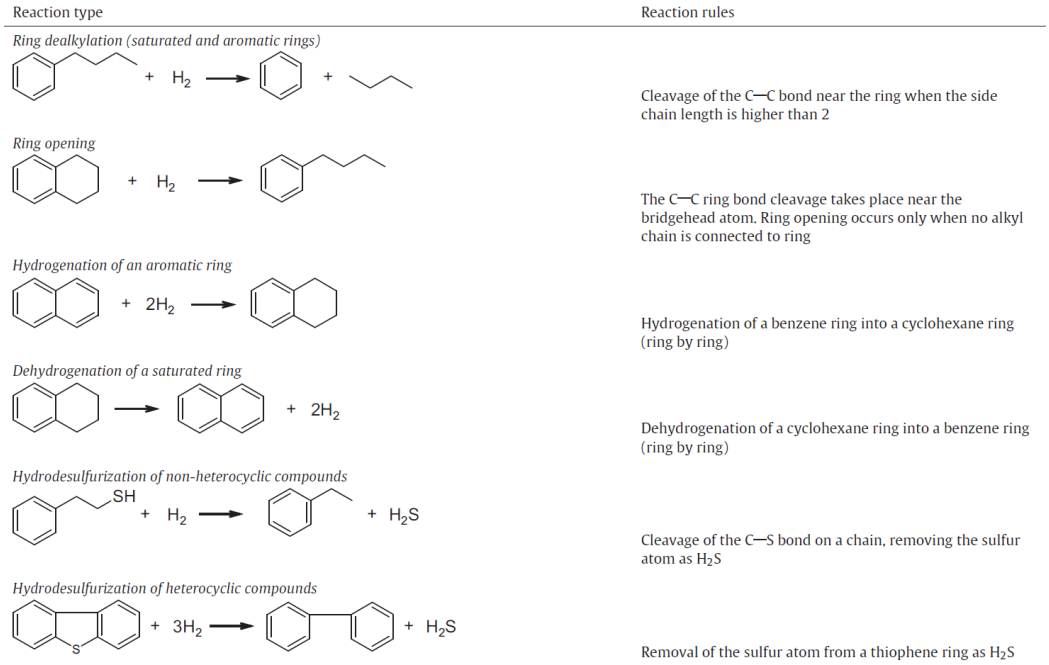


Figure I-42 - Reaction type and rules for the hydroconversion of vacuum residues (Oliveira et al. 2014).

Figure I-42 shows a the reaction considered for a molecular model for the hydroconversion of vacuum residues (Oliveira et al. 2014). The model considered six reaction families. Although Figure I-42 specifies several molecules, the reactions are applied to all molecules which can convert via these. In total, the authors (Oliveira et al. 2014) estimated four kinetic parameters, which was enough to retrieve a suitable model to the hydroconversion of vacuum residues.

To conclude, molecular models can be seen as an intermedium type between the lumps and the mechanistic models. Besides account with the chemical detail, as the mechanist models, the molecular

models are easily managed, as the lumped models, since only a reduced number of reaction types is required to describe a chemical system.

I.3.4 Reaction simulation techniques

Even if a suitable kinetic model for a complex mixture is achieved, the simulation of the associated reaction network is still difficult. In the open literature, there are two ways to simulate a reaction network of a given process: deterministic approach and stochastic approach. As will be presented in the following section, the first technique implies the description of the entire reaction network, which, for complex mixtures, it is not always possible. The second technique is based on random events, which, although requires more simulation time to achieved a suitable precision, can describe the reaction network of a complex network. For bio-oil mixtures, little information was found in the open literature. Apart for lumping models, which deterministically are ease to simulate, any other type of more complex models was not found.

I.3.4.1 Deterministic simulation

From a general perspective, deterministic simulation is classic method without any random event, where the inputs are previously known and a single 'mean field' output is generated, making it unique and uniform. For a chemical system, this technique can be described by ordinary differential equations (ODEs), as for ideal plug flow or batch reactors, or by algebraic equations, as for ideal continuous stirred tank reactors (Oliveira et al. 2016). In more detail, this method follows the continuous evolution of a chemical system through time/space, mostly by ODEs. Thus, for each chemical species present in the mixture, an ODE is created and afterwards, all the equations are integrated by analytical or numerical methods. Although the integration procedure has a rapid performance, it depends on the integration step and the equations under question, which, if complex, may increase the difficulty of resolution and damage the precision of the method.

In the end, a macroscopic scale description of the reaction behavior of each molecule through time is obtained as a function of their extensive properties, such as concentration, partial pressure, etc. However, not always a deterministic simulation suits the chemical system, leading to flaws in the approximation. As revealed above, when the mixture under study is associated to a certain degree of complexity, as molecular fluctuations (biological systems) or non-linear equations, this method fails, since it is not able to take into account such type of evolutions (Gillespie 1976). Additionally, in the case of complex mixtures, not only the inputs required, as the pre-established reaction network, are incredible difficult to obtain, but also the number of ODEs would be as large as the number of active species, turning the mathematical resolution of the system unreasonable (Oppenheim 1969) (Kurtz 1972) (Libaniti 1992). For these reasons, when faced with a complex mixture, the description of the system must be modified from a macroscopic to microscopic scale. Consequently, instead of a chemical evolution based on

extensive properties, a model supported on the evolution of a population and their reactions should be created.

I.3.4.2 Stochastic simulation

Contrary of the previous method, the stochastic simulation is related to the randomness of events, which may be represented by their probability of occurrence. The different nature of this method implies a description of the chemical system at the microscopic scale, which is based on the discrete evolution of a population and their reactions. In summary, the outputs of this technique depend strongly on probabilities of the events, making them suitable for large systems.

In this methodology, the time evolution is given by a single differential-difference equation for a grand probability functions named Chemical Master Equation (CME). This equation, represented in the form of Equation I-2, possesses a discrete nature and, as a result, calculates the different states of the reaction system in the time. However, this expression cannot be easily solved due to its high complexity. Therefore, two Monte Carlo methods were proposed.

$$\frac{\partial P(\overrightarrow{X_\iota}, t)}{\partial t} = \sum_{\overrightarrow{X_J}} \left[P(\overrightarrow{X_J}, t) \cdot T(\overrightarrow{X_J} \rightarrow \overrightarrow{X_\iota})\right] - P(\overrightarrow{X_\iota}, t) \cdot \sum_{\overrightarrow{X_J}} T(\overrightarrow{X_J} \rightarrow \overrightarrow{X_\iota}) \qquad \textit{Equation I-2}$$

On the one hand, Mc Dermott et al. (McDermott et al. 1990) built a procedure that follows the states of a system at fixed time integral and calculates which reactions happen in each interval. On the other hand, Gillespie (Gillespie 1976) (Gillespie 1992) created a method were the evaluation of the system is followed event-by-event (reaction by reaction) leading to disparate time intervals between two consecutive reactions. Although both are easily applicable, they also present their own drawbacks. The former requires the definition of the molecular pathways as an input data, while the latter, because of its stochastic nature, requires several runs that must be afterwards averaged in order to obtain an accurate representation of the reactions. Either way, and since the information required by McDermott (McDermott et al. 1990) method is extremely difficult to obtain, the implementation of this technique is futile in the case of complex mixtures. For this reason, the proposed methodology will be based on the Gillespie procedure (Gillespie 1976) (Gillespie 1992).

As referred before, the algorithm developed by Gillespie (Gillespie 1976) (Gillespie 1992) predicts results in exact accordance with the CME. This procedure is designated as the Stochastic Simulation Algorithm (SSA) and, due to its Monte Carlo character, is able to generate discrete time trajectories of molecular populations and handle chemical systems with many chemical species and complex reactions networks. As the CME, this procedure can describe the time evolution of a system step by step, ending in several events, each one corresponding to a reaction. To each of these events, a probability of

occurrence is associated. Since the SSA has a microscopic support, these probabilities are function of the molecules involved in the process and consequently function of their kinetic parameters.

In addition to the differences mentioned, the SSA requires less memory than the deterministic methods when processing complex mixtures. However, and due to its stochastic character, the former has a higher simulation time than the latter. As it happens, the SSA produces a different outcome for each simulation. Hence, a single simulation is not enough to construct the evolution of the reaction system over time. Therefore, this procedure requires a large number of simulations in order to obtain an appropriate result. After averaging, these allow to construct the time evolution of the system, as for the deterministic methods. Furthermore, the averaged results of all simulations of the SSA converge to the results of the classic procedure, as shown in Figure I-43. The SSA is therefore simply a different way of integrating the same differential model equations.

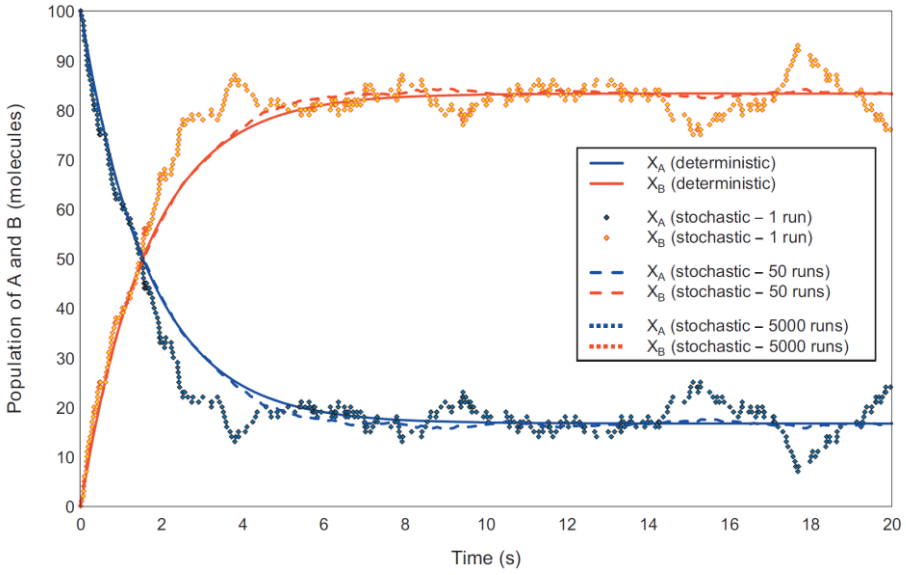


Figure I-43 - Example of results of stochastic and deterministic simulation (adapted from (Oliveira et al. 2016)).

As mentioned before, the stochastic simulation is quite useful for the simulation of processes involving complex mixtures. For oil fractions, this approach has already been tested and validated for Light Cycle Oils (Lopez Abelairas et al. 2016; Oliveira et al. 2012a), Vacuum Gas Oils (Alvarez-Majmutov et al. 2016; Alvarez-Majmutov and Chen 2017) and Vacuum Residues (Oliveira et al. 2012b; Oliveira et al. 2014). Furthermore, this approach was also used and validated for the liquefaction of lignin (Lopez-Abelairas M. et al. 2016). Similar to the two-step molecular reconstruction technique, this method is easily adaptable and very flexible, covering the simulation of several complex mixtures.

I.3.5 Conclusion of modeling complex mixtures

To model complex mixtures, there are four main concerns: the description of feedstock, the description of its reaction network, the suitable model and the suitable simulation strategy.

For the description of a complex feedstock, there is the molecular, deterministic and stochastic approach. The first one generates a single molecule to match the experimental mixture properties, the second method calculates, by optimization and with experimental data, the corresponding composition for a given mixture, and the last technique reproduces a library of molecules (w/ or w/out composition), whose mixture properties match the experimental properties. The first approach lacks molecular insight, while the second approach requires the prior knowledge of the molecules in the feed. For both reasons, none of these two approaches is suitable for mixtures with many unknowns, leaving only the last approach (stochastic molecular reconstruction) to be applicable for bio-oils.

For the description of the reaction network, there are several procedures. Nowadays, for complex mixtures, automated reaction network generators are the most frequently used tools. These do not only create the entire reaction network, but they also provide a kinetic model for the generated network. The type of model generated may vary between lumped, molecular or deterministic models depending on the mechanism insight desired by the user. For complex mixtures, lumped models (grouping molecules with similar properties) or molecular models (grouping molecules by their reactivity) are the ones usually applied.

For the simulation of a complex reaction network, it was seen that the deterministic approaches becomes difficult, not only because of the difficulty to generate the vast network exhaustively and to solve the huge set of differential equations, but also because the inputs contain many approximations, due to the difficulty in obtaining them. Alternatively, the stochastic simulation can directly follow the actual reaction pathways, step by step, based on random events, thereby generating the "useful" part of the exhaustive reaction network "on-the-fly". Since this method is able to overcome the lack of information regarding the reactional network, it is suitable to be applied to complex mixtures, such as heavy oil fractions and bio-oils.

I.4 Conclusions of Chapter I

As exposed in this chapter, biomass is an organic feedstock with strong potential to produce fuels and valuable chemical building blocks. In order to achieve these compounds, biomass must first be pre-treated and converted into a matter easily manageable, preferably liquids. One of these conversion processes is fast pyrolysis, which generates a solid, a gas and a liquid, called bio-oil. This mixture is complex oxygenated matrix, with a vast number of molecules, and consequently, a great diversity of characteristics. In bio-oil, it is possible to find compounds as alcohols, aldehydes, carbohydrates,

carboxylic acids, furans, ketones, and phenols and methoxy-phenols. More details of the molecular characteristics of the bio-oil is still quite complicated to acquire analytically. The undesired amount of oxygen in bio-oil enables its mixing with conventional fuels composed of hydrocarbons and a suitable characterization. Even so, several analytical techniques or combinations of techniques have been applied.

Through the conversion of bio-oil, it is possible to finally obtain fuels and other chemical building blocks. These mixtures can be transformed by several processes, but hydrotreating is the most applied. This process is wide-spread in conventional oil refineries for the removal of undesirable heteroatoms, which, in the case of bio-oils, would be the removal of oxygen. Be that as it may, this process is not simple, in particular due to the complex reactivity pattern and the unstability of the mixture. Therefore, in order to avoid polymerization reactions, the mixture must first be stabilized at low temperature, and followed by a conversion process at a higher temperature. The behavior of bio-oil under hydrotreating conditions is quite difficult to predict. The scientific community has been analyzing this complexity of processing of real bio-oil fractions for several years, mainly through studying model molecules representative of bio-oils. Although working with actual bio-oil fractions is rather difficult, the model molecule approach allows generating a better understanding of the bio-oil reaction network. Concisely, bio-oil and its model molecules, when under hydrotreating conditions, can be converted into hydrocarbons and water by hydrogenation, hydrogenolysis, hydrodeoxygenation, decarboxylation, decarbonylation, cracking/hydrocracking and polymerization reactions (Wildschut 2009). It is important to note that these reactions are strongly impacted by the solvent, catalyst, temperature and structure of the molecule.

Even if these model molecules are extremely useful in order to study, identify and retrieve intrinsic information about the bio-oils kinetics, the overall reaction network of bio-oil cannot be totally covered by such an approach. Therefore, modeling tools, tailored for complex mixtures, may be considered. For bio-oils, it is required:

- 1st – A molecular description technique, in order to reconstruct the feed,
- 2nd – A tool for the generation of the reaction network, as it is too large to be done by the user,
- 3rd – A suitable kinetic model approach, which is able to deal with molecular details, but that present a low complexity,
- 4th – A suitable simulation strategy that can cover the vast reaction network of bio-oils without a great complexity.

Among all techniques previously explain, it was considered that coupling the stochastic reconstruction and the reconstruction by maximization of entropy would be a good approach for the molecular description of bio-oils. To cover the remaining aspects, it was considered that the stochastic simulation algorithm would be an appropriate method. Besides being able to generate the entire reaction network,

this technique can easily simulate the hydrotreating of a complex mixture according to the desirable level of kinetics. The kinetics should at least follow a molecular approach, in order to capture the molecular details of bio-oil, but in a simple way.

To conclude, the chapter of the State of the Art had as objective to provide some answers to the three questions asked in the introduction of the present thesis. The answers to the main problematics of the present subject are given below:

1st – How to describe a bio-oil mixture in molecular terms?

Bio-oil is an complex oxygenated mixture. According to the analytical characterization available, bio-oil can be divided in alcohols, aldehydes, carbohydrates, carboxylic acids, furans, ketones, phenols, methoxys-phenols and other unknown molecules.

2nd – How does a bio-oil mixture react under hydrotreating conditions?

Under hydrotreating conditions, bio-oil can react through hydrogenation, hydrogenolysis, hydrodeoxygenation, decarboxylation, decarbonylation, cracking/hydrocracking and polymerization reactions. The impact of each reaction in the network depends on the solvent, the catalyst, the temperature and also the molecular structure.

3th – How to create an *in silico* tool for hydrotreating of bio-oil?

A compositional model can be made by coupling the stochastic reconstruction and the reconstruction by entropy maximization, while the kinetic model can be based on a stochastic simulation algorithm, which can simulate the hydrotreating of bio-oil.

Chapter II. Objectives and Strategy

As seen in the State of the Art, in many cases, either a model is simple, but not accurate; or it is accurate, but extremely complex and difficult to solve. Therefore, the present work proposes a methodology which aims at balancing both the accuracy and the difficulty in generating a model that not only represents the diversity of the bio-oil molecules and reactions, but also is simple to create.

The following section intends to clarify the strategy of the present work, by giving a detailed insight on its resolution and aims. For this reason, the motive of the present thesis, its main objective, the strategy of the thesis and the accomplishments to be achieved will be presented. As will be shown, the section was divided into three steps: **Why**, **How** and **What**. These will help to get a better understanding of the thesis work and aims.

II.1 Why

The present work is **motivated by the ability of bio-oil to generate fuels and valuable chemical building blocks**. However, it is required to remove the oxygen from the mixture in order to generate these valuable hydrocarbons. Through a hydrotreating process, bio-oils can be converted into non-oxygenated molecules. For this reason, the main **objective is to contribute to the development and understanding of a hydrotreating process dedicated to bio-oils**.

II.2 How

To achieve the main objective, several approaches can be proposed, from experimental studies to the application of these to modeling techniques. Due to the limited time given to the project, the present thesis will focus mostly on the modeling approach. This will be supported by previously obtained experimental data (Ozagac 2016) and a new set of experiments that will be carried out in this work.

The present work will describe the behavior of a bio-oil mixture under hydrotreating conditions through a stochastic molecular reconstruction method (SMR) and a stochastic simulation algorithm (SSA). As pointed out in the State of the Art, these two tools are suitable for the simulation of complex mixtures, as already proven for oil fractions.

How will these two approaches be applied to bio-oils?

As mentioned before, bio-oils are a complex mixture and therefore the scientific community usually starts by analyzing single model compounds, then mixtures of these and only at the end a real bio-oil feed. This strategy was also followed by Ozagac (Ozagac 2016) in his PhD thesis. For these reasons, the present work will be unrolled along the same strategy, but from the modeling point of view.

As shown in Figure II-1, the modeling approach will be based on the experimental data. The thesis will start by the modeling (SSA) of single model compounds (guaiacol and furfural), from where it is intended to retrieve information regarding functional group reactions and kinetics. Although the SSA is not the most adapted approach for model molecules, it is requested from the reader to keep in mind the final aim: bio-oil, which is a complex mixture of many different molecules with various functional groups. The application of the SSA to single molecules can be seen as a step in the adaptation of a complex algorithm to bio-oils. Thanks to the experimental data on single molecules, the algorithm can be gradually evaluated and validated.

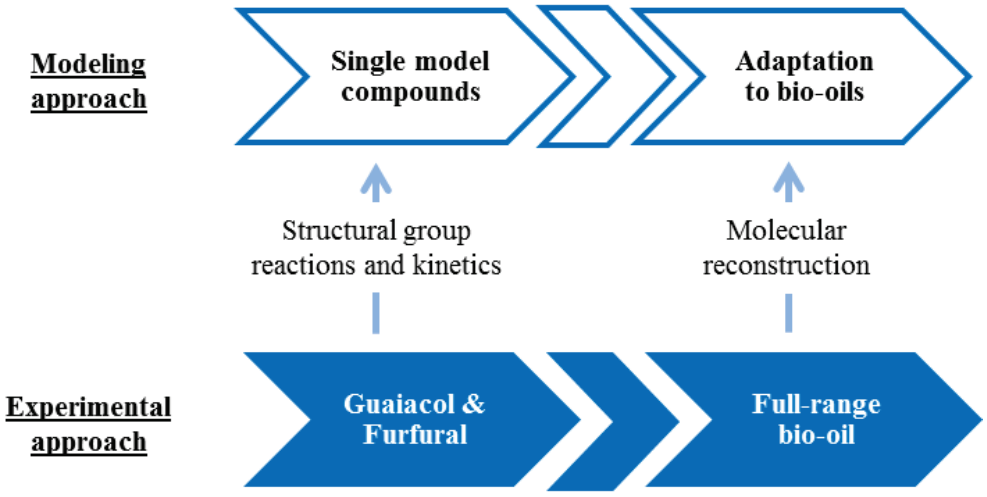


Figure II-1 – Proposed strategy for the present PhD thesis.

The following step of the strategy should be the application of the same strategy to mixtures of model compounds, from where information regarding the interaction of molecules could be extracted. However, due to the limited time available in a PhD thesis, this step will be skipped. A direct application of the information extracted from the single model compounds will be used for a real bio-oil. To apply such a molecular model, a molecular description of the bio-oil mixture is required. Since such a description cannot be provided analytically, it will be modeled by molecular reconstruction (MR) by means of a a synthetic mixture representative of the bio-oil fraction. Later on, this library of molecules will be used as input in the simulation of the actual bio-oil fraction with the kinetics from the single model compounds.

In conclusion, the present thesis will follow the strategy of Figure II-1 via a stochastic simulation procedure, supported by a molecular reconstruction technique. The retrieved outputs will afterwards be confronted with experimental data obtained in this work and in the thesis of Ozagac (Ozagac 2016). **It is not expected that a full agreement between the simulation and the results** can be obtained in a single PhD thesis, as the created tool will not be adjusted to enough experimental data and will not

account for the interaction between various molecules. However, **it is expected that both outputs will follow the same trends**.

II.3 What

By following the strategy in Figure II-1, the main objective should be fulfilled, but various other smaller accomplishments will reached. These actions will be explained in more detail during the manuscript. However, this section will present the crucial points of the thesis, in order to give the reader a clear guideline of the document and its aims.

As mentioned before, the present thesis will also include an experimental part. This will follow the guidelines of Ozagac (Ozagac 2016) and will only be performed for single molecules, mainly guaiacol and furfural. Although it will be explained later in the manuscript, the aim is to extend the previous experimental data set to lower temperatures.

Besides the experimental part, two deterministic models, one for guaiacol and another for furfural, will be developed. As shown in the State of Art, little information exists regarding the kinetics of oxygenated molecules under hydrotreating conditions, in particular over a reduced $NiMo/Al_2O_3$ catalyst. For this reason, a deterministic model will serve as an initial kinetic support for the stochastic model of the single molecules.

Afterwards, the stochastic simulation algorithm will be adapted to the reactions of the single molecules. At the same time, the remaining reactions found in the bio-oil network will also be implemented. It is important to note that the current version of the SSA only simulates monomolecular reactions. Since the hydrotreating process is performed with an excess of hydrogen, the application of the SSA does not present a problem for the hydrotreating reactions themselves. Furthermore, as the mixtures of model compounds will not be simulated, the undesirable reactions, which are mostly bimolecular, will not be added to the simulator either. Nevertheless, a number of undesirable monomolecular reactions, such as dehydration, hydration, decarbonylation, decarboxylation and retro-aldol, will be added. The stochastic simulation algorithm will also include the integration of a temperature profile. Indeed, this is important for batch tests in order to be able to simulate the heating and cooling periods, during which the various reactions are also contributing. To finish, the algorithm will be adapted as well for the reading and printing of reactant/product classes (guaiacol, phenols, etc.) and oxygenated molecules classes (monoalcohol, dialcohol, etc), respectively.

The thesis will continue with the simulation of guaiacol and furfural with the kinetic parameters estimated by the deterministic approach. From here, new parameters will be estimated for both model molecules to better fit the diversity of the products that are generated with the stochastic algorithm

(generation of many products) but are absent in the deterministic model (only products quantified by gas chromatography).

Afterwards, the thesis will focus on the molecular reconstruction of bio-oils. The first half of this method has been set up during an internship. The stochastic reconstruction method was adapted to bio-oil mixtures by addition of new properties (oxygenated molecules classes, size exclusion chromatography, etc.), new molecular structures (atoms, tetrahydropyran, etc.) and finally new probability distributions functions (uniform distribution, normal distribution, etc.) The adaptation of the algorithm will not be presented in the thesis manuscript. For further information please read the reference (Ribeiro 2018). During the internship, ten building diagrams were tested and only four were chosen as potential generators of a bio-oil library. The stochastic reconstruction results will then go through an entropy maximization step. This step will first be adapted to the bio-oils by addition of the new properties in the optimization formula. The new library of molecules will be discretized (match the number of molecules to their composition) later on and fed to the stochastic simulation algorithm for the simulation of a real bio-oil.

Below, a summarized list of the objectives is proposed for the present thesis. Through their fulfillment and following the previously defined strategy, it will be possible to achieve the main objective of this thesis: Developing and understanding the hydrotreating process of bio-oils.

Proposed actions:

1. Experimental data:
   1.1. Guaiacol
   1.2. Furfural
2. Create a deterministic model:
   2.1. Guaiacol
   2.2. Furfural
3. Adaptation of the SSA:
   3.1. Reactions of guaiacol
   3.2. Reactions of furfural
   3.3. Remaining reactions of bio-oil
   3.4. Undesirable reactions
   3.5. Temperature integration
   3.6. Oxygenated compounds classes
   3.7. Reactants classes
4. Create a stochastic model:
   4.1. Guaiacol
   4.2. Furfural
5. Adaptation of the SMR:
   5.1. Properties
   5.2. Molecular structures
   5.3. Probability distribution function
6. Creation of building diagrams
7. Reconstruction of bio-oil
8. Simulation of bio-oil hydrotreating

II.4 Conclusion of Chapter II

In conclusion, the present work intends to develop and understand the hydrotreating process of bio-oils, as these mixtures can be used for the production of fuels and valuable chemical building blocks. To achieve this objective, the thesis will use a stochastic molecular reconstruction method and a stochastic simulation algorithm to handle the bio-oil complexity *in silico*.

PART II: Experimental approach

"Essential information for elucidation of the HDO mechanism has been obtained during studies involving model compounds."
(Furimsky 2000)

Table of Contents

Chapter III. Experimental methodology

In collaboration with Elodie Dusson, Manuel Menard and Yohann Mouillet

When one wants to develop modeling tools, an experimental background of the context under study is always required. The present work will therefore be based on the experimental work of Matthieu Ozagac (Ozagac 2016). Furthermore, additional tests were performed at lower temperatures, in order to achieve a data range from low to high conversions. These new experimental tests were performed according to the procedures of Ozagac (Ozagac 2016). The following chapter intends to summarize the experimental data obtained by Ozagac (Ozagac 2016) and the new data set obtained in the present thesis. Thus, this section will expose not only the pre-requirements for the experimental tests, but also the procedure for the hydroconversion of the model compounds and their analyses.

III.1 Feedstock

During his thesis work, Ozagac (Ozagac 2016) analyzed various model compounds. From individual components to mixtures experiments, the author was able to propose the most probable reaction mechanisms. Furthermore, Ozagac (Ozagac 2016) studied as well the hydroconversion of a sample of bio-oil. In addition to these tests, the author always took under account the formation of undesired compounds and its reactional behavior under the process conditions.

Independently of the type of test, Ozagac (Ozagac 2016) tried to maintain a coherency between his studies and the real bio-oil. As model molecules, the author selected compounds which, besides being majority in bio-oils, are also the most studied in the open literature. Regarding the mixtures of these, the author focused in maintaining the total acid number (TAN) and the ratios between oxygen/carbon and hydrogen/carbon the closest possible to the bio-oils properties.

With the intent of continuing the work of Ozagac (Ozagac 2016), the present work will focus on the same model compounds. However, only two of the compounds will be studied.

III.1.1 Prior experimental tests

As described in the work of Ozagac (Ozagac 2016), several feedstocks were used starting by single model compounds, passing to mixtures of model compounds and ending with a bio-oil sample from the VTT Technical Research Centre of Finland. The increase of the mixture complexity is associated to the comprehension of the reaction mechanisms dedicated to each of the bio-oil typical functional groups. Through this type of strategy, the author was able to first understand the reactional behavior of the model compounds, then the interaction between the different molecules and finally compared these results against the conversion of a bio-oil sample. In the end, Ozagac (Ozagac 2016) concluded that mixtures

of typical molecules present in bio-oil describe approximately well the behavior of this complex feedstock. For all the experimental studies, Ozagac (Ozagac 2016) applied a hydrogen pressure of 13 MPa in a reactor with the feedstock and a basket with reduced $NiMo/Al_2O_3$. The author performed as well several non-isothermal tests (i.e. tests with a residence time of 0 h), designated here by "blank tests".

III.1.1.1 Single model compounds

Regarding the single model compounds, Ozagac (Ozagac 2016) selected guaiacol, furfural, D-glucose and acetic acid. As mentioned before, all compounds are regularly found in bio-oil in considerable amounts. Furthermore, each molecule can be traced to its original fraction of lignocellulosic biomass: guaiacol as a representative of lignin, D-glucose as a product of cellulose degradation, and furfural and acetic acid as exemplars of hemicellulose conversion (Patwardhan 2010).

Besides the presence of water, Ozagac (Ozagac 2016) selected as solvent *n*-hexadecane due to its absence of reactivity during the process and analytical procedures. In the end, each experimental test fulfilled a total mass of 150 g, accounting with the water, the solvent and the model compound. Contrary to the former component, the mass fraction of water and *n*-hexadecane varied according to the experimental test. For each model compound, the author analyzed the effect of the reaction time, temperature and amount of water in the medium. All the tests performed by Ozagac (Ozagac 2016) are described in Table III-1 until Table III-4. For the generation of the kinetic model, the studies of the water impact will be set aside and only the tests present in Table III-10 will be accounted for.

Guaiacol

For guaiacol, Ozagac (Ozagac 2016) tested several solutions, in particular to analyze the impact of water, as shown in Table III-1. For the effect of the reaction time and the temperature, the author also worked also with the solution at bold: water and guaiacol at 30 %w/w and n-hexadecane until the fulfilment of 150 g. The conditions referent to these tests can be found in Table III-10.

Table III-1 - Solutions of guaiacol studied by Ozagac (Ozagac 2016).

Water (%w/w)	Guaiacol (%w/w)	*n*-hexadecane (%w/w)
50	30	20
30	**30**	**40**
20	30	50
1.3	30	68.7
0	30	70

Furfural

With a similar purpose as guaiacol, the author tested several solutions of furfural, as shown in Table III-2. For the reaction time and the temperature, Ozagac (Ozagac 2016) selected the solution in bold:

water at 30 %w/w, furfural at 13 %w/w and n-hexadecane until the fulfilment of 150 g. Once again, the conditions of these tests can be found in Table III-10.

Table III-2 - Solutions of furfural studied by Ozagac (Ozagac 2016).

Water (%w/w)	Furfural (%w/w)	*n*-hexadecane (%w/w)
87	13	0
55	13	32
30	**13**	**57**
15	13	72
8	13	79
1.3	13	85.7
0	13	87

D-Glucose

Analogous to the two previous compounds, the impact of water in the medium was analyzed through the several tests present in Table III-3. For the reaction time and the temperature, Ozagac (Ozagac 2016) used the solution in bold: water at 80 %w/w and D-glucose until the fulfilment of 150 g. For this model compound, the author chose not to add any solvent. The conditions associated to these tests are in Table III-10.

Table III-3 - Solutions of D-glucose studied by Ozagac (Ozagac 2016).

Water (%w/w)	D-Glucose (%w/w)	*n*-hexadecane (%w/w)
80	**20**	**0**
60	20	20
45	20	35
37.5	20	42
30	20	50

Acetic acid

As mentioned before, the addition of the acid to the model compounds groups intends only to get the model compound studies closer towards reality. As shown in Table III-4, only one test was required for Ozagac (Ozagac 2016) to conclude its unfeasibility. Besides difficulties at the experimental procedure level, the post-treatment was likewise complicated.

The acetic acid produces majority acetone and other light compounds. All these products tend to vaporize easily, which causes mass losses (superior to 30 %w/w of carbon) and, consequently, affect the final mass balance. Additionally, the typical post-treatment used by the author included the cleaning

and recuperation of all residues through acetone. As the model compound produces large amounts of acetone, the post-treatment solvent was changed to dichloromethane. However, the separation of dichloromethane from the products is difficult, which injures the post-treatment process and the compounds recuperation. No further tests with acetic acid as an individual molecule were performed.

Table III-4 - Solution of acetic acid studied by Ozagac (Ozagac 2016).

Water (%w/w)	**Acetic acid (%w/w)**	***n*-hexadecane (%w/w)**
30	6.7	63.3

III.1.1.2 Mixtures of model compounds

Concerning the binary and ternary mixtures of model compounds, Ozagac (Ozagac 2016) analyzed several solutions as well. On the one hand, the effect of water was studied thanks to the solutions in italic in Table III-5. On the other hand, the author tried to also understand the interaction between the different functional groups thanks to the solutions in bold in Table III-5.

Table III-5 – Binary and ternary mixtures studied by Ozagac (Ozagac 2016).

Type of mixture	**Water (%w/w)**	**Guaiacol (%w/w)**	**Furfural (%w/w)**	**D-Glucose (%w/w)**	**Acetic acid (%w/w)**	***n*-hexadecane (%w/w)**
Binary	**30**	**0**	**0**	**20**	**7**	**43**
	30	**30**	**0**	**20**	**0**	**20**
	30	**0**	**13**	**20**	**0**	**37**
	30	**30**	**13**	**0**	**0**	**27**
	30	**0**	**13**	**0**	**7**	**50**
	73	*0*	*0*	*20*	*7*	*0*
	50	*30*	*0*	*20*	*0*	*0*
	67	*0*	*13*	*20*	*0*	*0*
	57	*30*	*13*	*0*	*0*	*0*
	43	*0*	*13*	*0*	*7*	*0*
Ternary	**30**	**30**	**0**	**20**	**7**	**13**
	30	**0**	**13**	**20**	**7**	**30**
	30	**30**	**13**	**20**	**0**	**7**
	30	**30**	**13**	**0**	**7**	**20**
	43	*20*	*30*	*0*	*7*	*0*
	60	*20*	*0*	*13*	*7*	*0*
	37	*20*	*30*	*13*	*0*	*0*
	50	*0*	*30*	*13*	*7*	*0*

Regarding the quaternary mixtures, the author aimed not only to study the impact of guaiacol in the reactional medium, as seen in Table III-5, but also to analyze the effect of the reaction time and temperature, as shown in Table III-10.

Table III-6 – Quaternary mixtures studied by Ozagac (Ozagac 2016).

Type of mixture	Water (%w/w)	Guaiacol (%w/w)	Furfural (%w/w)	D-Glucose (%w/w)	Acetic acid (%w/w)	*n*-hexadecane (%w/w)
Quaternary	30	30	13	20	7	0
	30	15	13	20	7	15

III.1.1.3 Bio-oil

Ozagac (Ozagac 2016) used a bio-oil provided by the VTT Technical Research Centre of Finland. The sample was through fast pyrolysis of lignocellulosic biomass (forest residues). The original feed underwent fast pyrolysis at 480 °C with a residence time inferior to 1 s and. Table III-7 summarizes the principal properties of the bio-oil.

Table III-7 – Bio-oil from VTT studied by Ozagac (Ozagac 2016).

Water	Carbon	Hydrogen	Oxygen	Nitrogen	Ash	Viscosity (40 °C) (cSt)	TAN (mg_{KOH}/g_{B-H})	Density (20 °C) (kg/L)
(%w/w)								
26.4	56.4	6.4	36.7	0.0042	0.2	14.9	91	1.2

Regarding the experimental tests, Ozagac (Ozagac 2016) analyzed the impact of the reaction time and temperature with and without the presence of guaiacol (report to Table III-10). Thanks to this compound, the author was able to inhibit the formation of macromolecules and undesired solid residues. Furthermore, Ozagac (Ozagac 2016) studied also the impact of the absence of catalyst and the use of nitrogen. As these experimental tests will not be accounted for the kinetic model, they can be found only in the thesis work of Matthieu Ozagac (Ozagac 2016).

Table III-8 - Solutions of bio-oil studied by Ozagac (Ozagac 2016).

Bio-oil (%w/w)	Guaiacol (%w/w)
100	0
75	25
50	50

III.1.2 New experimental tests

Although Ozagac (Ozagac 2016) performed a vast setup of experimental tests, not all his results can be adapted to the present work. By focusing more in the evolution of the hydroconversion process through the reaction time, the author obtained high conversions (> 40%) for the reactants. With such advanced reactional state, a kinetic study is not feasible. For this reason, new experimental tests were performed at lower temperatures than Ozagac (Ozagac 2016) to allow a better observation of the initial kinetics of the system.

Unlike the author, the new experimental tests aimed only single model compounds, more precisely guaiacol and furfural. Both model compounds were studied in the form of the solutions in Table III-9. As a continuity of the trials performed by Ozagac (Ozagac 2016), it was considered the validation of repeatability tests done by the author. D-glucose has set aside due to its fast conversion and reaction similarity with the furanic compounds.

In total, three new tests were effectuated for both compounds at low temperatures (100 °C and 150 °C), as shown in Table III-10. The experimental setup, as well the protocol applied will be presented in the following section.

Table III-9 - Solutions of the new experimental tests.

Water (%w/w)	**Guaiacol (%w/w)**	**Furfural (%w/w)**	***n*-hexadecane (%w/w)**
30	30	0	40
30	0	13	57

Table III-10 - Prior (✓) and current (●) experimental tests all at 13MPa and in the presence of reduced $NiMo/Al_2O_3$ (Ozagac 2016).

Temperature (°C)		100				150				200				250				300			
Reaction time (h)		0[Δ]	1	3	5	0[Δ]	1	3	5	0[Δ]	1	3	5	0[Δ]	1	3	5	0[Δ]	1	3	5
Single model compounds [*]	Guaiacol (Gua)						●	●	●			✓			✓	✓		✓	✓	✓	✓
	Furfural (Fur)	●	●				●					✓	✓	✓	✓	✓			✓	✓	
	D-glucose (Glu)									✓	✓	✓			✓	✓		✓		✓	
	Acetic acid (Ace)														✓						
Binary mixtures [*]	Fur + Ace														✓						
	Glu + Ace														✓						
	Glu + Fur														✓						
	Gua + Glu														✓						
	Gua + Fur														✓						
Ternary mixtures	Glu + Gua +Fur														✓						
	Glu + Fur + Ace														✓						
	Glu + Gua + Ace														✓						
	Gua + Fur +Ace														✓						
Quaternary mixtures [+]	without solvent										✓			✓	✓				✓		
	with solvent														✓						
Bio-oil [°]	without Guaiacol										✓			✓	✓	✓			✓		
	with Guaiacol										✓			✓	✓	✓			✓		

[Δ] Blank test. [*] Experimental tests with a constant amount of water. [+] All four compounds: guaiacol, furfural, D-glucose and acetic acid. [°] Additional point: 30 min @ 250 °C.

III.2 Equipment

The experimental tests are performed in a pilot unit in *IFP Energies nouvelles*. Although Ozagac (Ozagac 2016) used a setup identified as unit 546 (U546), the new trials will be done in unit 866 (U866). The similarity between both units assures the same behavior of the system.

The unit 866 consists of a batch reactor of 0.5 L, which is placed the middle of the unit (see Figure III-1 a). This reactor must endure the hydroconversion conditions. Therefore, the recipient is made in Inconel 718 and can sustain a temperature and pressure up to 450 °C and 270 bar, respectively.

In addition to the reactor, the unit has a mixer, a cooling system and equipment to measure and control the temperature and the pressure (see Figure III-1 b). Furthermore, the batch has a gas entry and exit, as shown in Figure III-2. In the present work, nitrogen was used for the purging of the reactor and the lines, while hydrogen was used for the reaction.

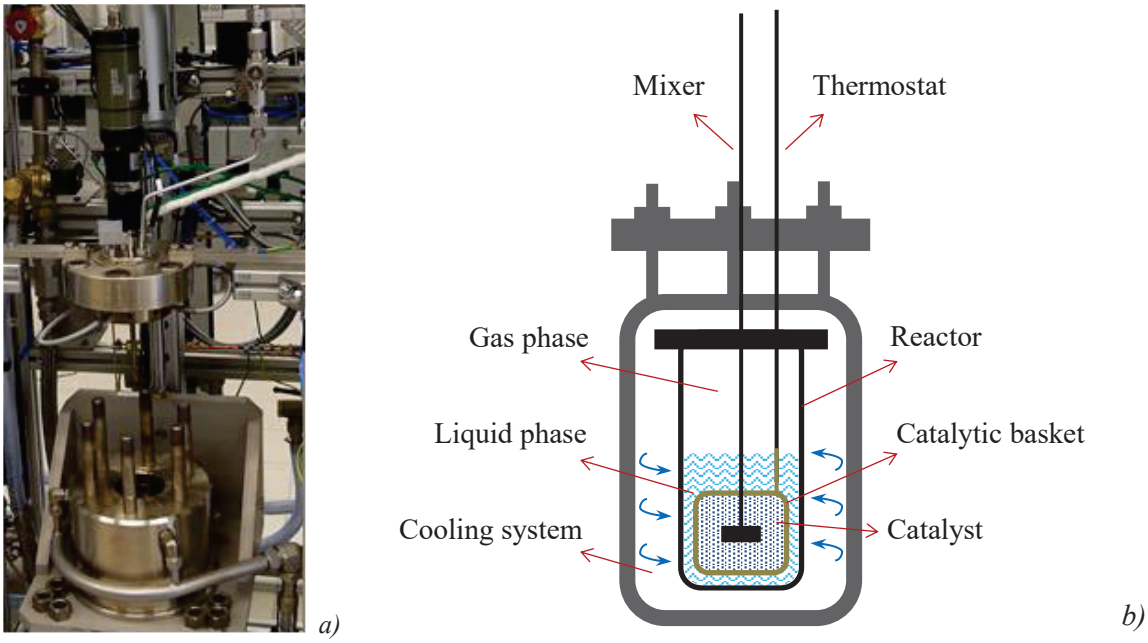


Figure III-1- Description of the unit 866: a) external view, b) internal view with the reactor.

Inside the reactor, there is a liquid feedstock, stirred by a mixer, and the catalyst in solid form. This placed in a basket or in the liquid medium. As seen by Ozagac (Ozagac 2016), although the basket may bring more diffusional obstacles, the free circulation of catalyst leads to its clogging in the bottom of the reactor. As a consequence, the recuperation of the catalyst, for the post-treatment, becomes difficult. The catalytic basket as a donut form, meaning it is a cylinder with a hole in the middle. In there is where the mixer is placed. This piece is mobile as well and rotates thanks to the gas inside the reactor. In the form of a Rushton turbine, the mixer can achieve, through a radical flow agitation, the 1200 rpm. To avoid an overheating of the equipment, the mixer is cooled by cold water at 5 °C.

In the reactor, there are also two thermostats. One placed at the gas phase and another inside the liquid phase through the catalytic basket (see Figure III-1 b). When heating the system, both can achieve the desired temperature with a ramp of 15 °C/min (Ozagac 2016). The system thermodynamic is stable when both temperatures present the same value. The cooling of the reactor is done by means of a compressed air system. Placed around the reactor, the air rotates as a vortex and reduces the temperature at a ramp of 30 °C/min during the first 200 °C of cooling (Ozagac 2016). All the pressure controls are in the gas lines at the entry or exit of the reactor.

Associated to the unit, there is a ballast (1 L), a Grayel bottle (15 L) and the purge lines, as shown in Figure III-2. The hydrogen always comes from the ballast, which is fed through the network gas at 100 bar. The alimentation passes firstly by a compression, which assures the desired pressure inside the ballast. The nitrogen directly comes from the gas network at 20 bar. During the purging of the system, the nitrogen is fed to the reactor. After a given time of stabilization, the gas is removed from the unit directly to the gas treatment area. When testing the reactor isolation and heating the reactor, hydrogen exit likewise through the gas treatment line. After the reaction, the hydrogen and the generated gases are stored in a Grayel bottle.

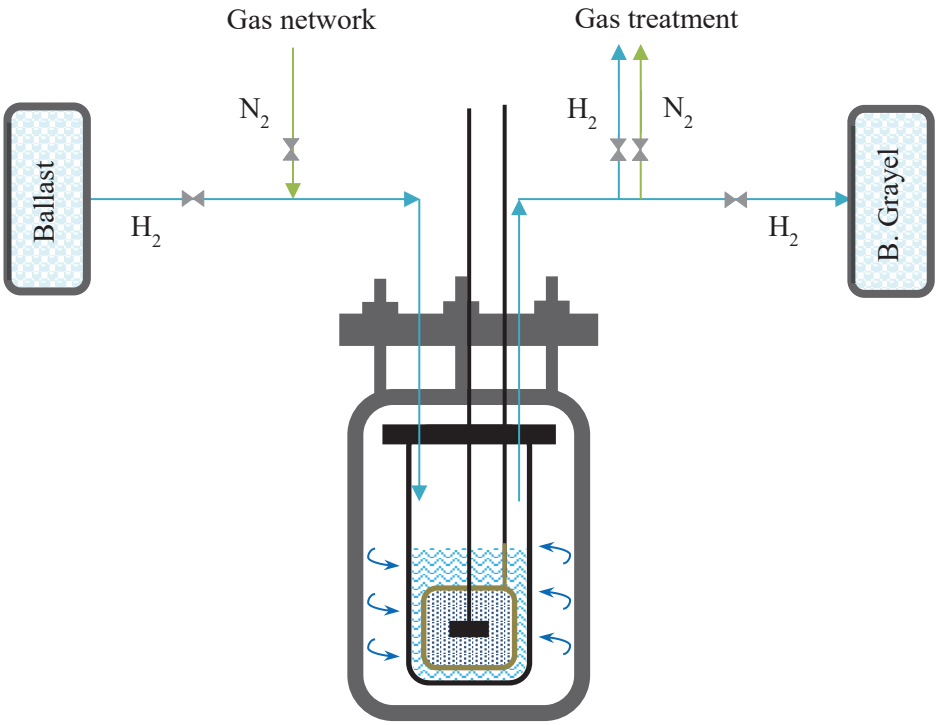


Figure III-2 – Schematic process flow of unit 866.

For more details regarding the pilot units, please, consult the thesis work of Matthieu Ozagac (Ozagac 2016).

III.3 Experimental protocol

The present section intends to describe the experimental protocol used in the obtainment of the new tests. The approach is divided in four parties: the catalyst preparation, the feedstock preparation, the follow up of an experimental test and the consequent post-treatment. As this protocol is based on the experimental approach of Ozagac (Ozagac 2016), any further details can be found in his work.

III.3.1 Catalyst preparation

As mentioned at the beginning of this chapter, the experimental tests are done in the presence of reduced $NiMo/Al_2O_3$, which is typical of the hydroconversion processes. The catalyst is initially in the form of spherical particles with diameters between 2 and 4 mm. The amount of nickel is 9.2 %w/w, while the presence of molybdenum is 5.4 %w/w. The catalytic support is alumina with a specific surface area of 135 m^2/g.

As studied by Ozagac (Ozagac 2016), the original particles result in a considerable limitation to the molecules diffusion. For this reason, the author proposed the reduction of the grains diameter. From several experimental tests, the author concluded that a particle size between 1 and 2 mm was the ideal size.

Unlike the typical $NiMo/Al_2O_3$ used in the hydrodesulfurization, the catalyst used in the present trials was reduced in the presence of hydrogen. In a glass cell, the catalyst is activated when traversed by a hydrogen flux fixed at 2.7 $NLH_2/h/g_{catalyst}$. After 2 h of reduction at 400 °C, the catalyst is shifted into the basket and then into the reactor already charged with the feedstock. In order to avoid the oxidation of the nickel, the transference of catalyst is done is a sealed pouch in an inert argon atmosphere. According to the Ozagac (Ozagac 2016), the hydroconversion tests require 15 g of reduced $NiMo/Al_2O_3$. Furthermore, with this amount it is possible to assure the desired reactions and the reduction of the diffusion limitations.

III.3.2 Feedstock preparation

According to the experimental tests (see Table III-10), the feedstock can be of guaiacol or furfural. The fractions of the components in the mixtures are present in Table III-9, while Table III-11 gives the source of the compounds.

In the making of the feedstock, it was used a precision balance with an uncertainty of 0.3 g. The final feedstock contains the model molecule, demineralized water and the solvent in a total of 150 g. Once in the reactor, the catalytic basket in inserted (under an argon atmosphere) and all the setup in placed inside the unit, which is then closed.

Table III-11 – Commercial references of the studied compounds.

Compound	CAS number	Purity (%w/w)	Source reference
Guaiacol	90-05-1	98+	Sigma-Aldrich
Furfural	98-01-1	99	Sigma-Aldrich
n-Hexadecane	544-76-3	99	Alfa Aesar

III.3.3 Experimental test

After closing the unit, the reactor is purged with nitrogen to remove all the air enclosed as well. Afterwards, the reactor is purged with hydrogen to eliminate all the nitrogen. The purge process is repeated three times for each gas with a pressure around about 0.3 MPa. In the end, the reactor is filled with hydrogen until a pressure around about 13 MPa. For a time equivalent to the reaction time of the test, a sealing test is performed by closing both the gas entry and exits of the batch reactor. The objective of this procedure is to confirm the non-existence of hydrogen leaks. If the loss of hydrogen pressure during the test is negligible comparing with test pressure (< 0.05 MPa/h), as seen in Figure III-3 step *1*, then the hydroconversion treatment can finally start.

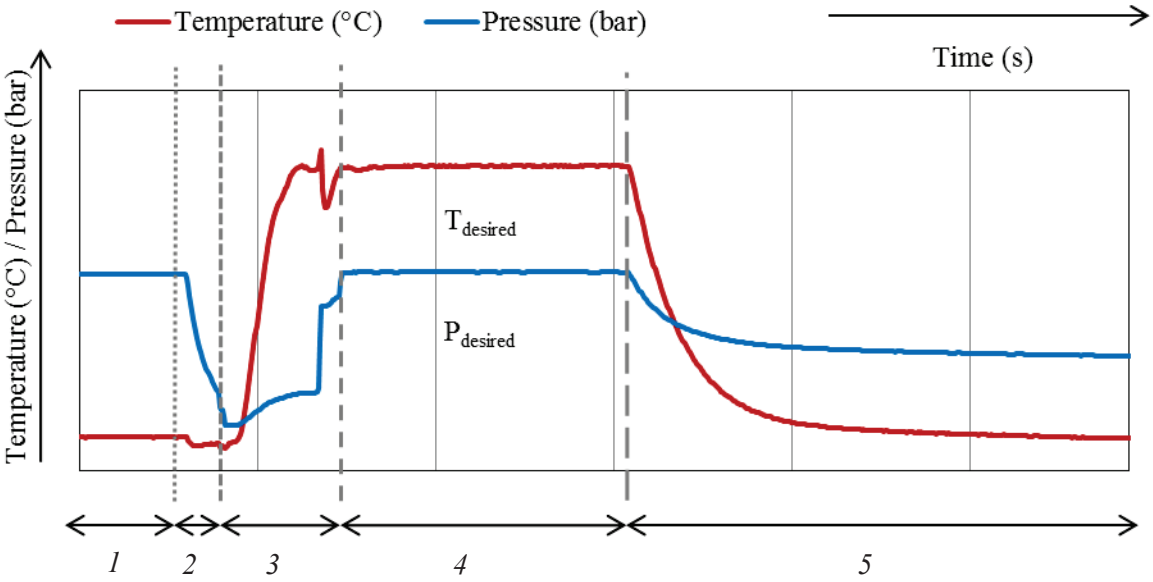


Figure III-3 –Dynamic thermal behavior of the reactor during the experimental tests.

Before heating the batch, the pressure inside of the reactor is dropped until 30 bar to avoid overpressure the reactor during the temperature increment. As shown in Figure III-3 step *2*, the reduction of the pressure reflects a reduction of the temperature as well. Afterwards, both the gas entry and exist are closed and the temperature is set to the desired value (see Figure III-3 step *3*). When the temperature hits the value expected for the experimental test, the pressure is rapidly increased until 13 MPa by injecting hydrogen in the reactor. It is important to take into account the temperature stabilization between the gas and liquid in the reactor. The reaction time is accounted when the desired conditions

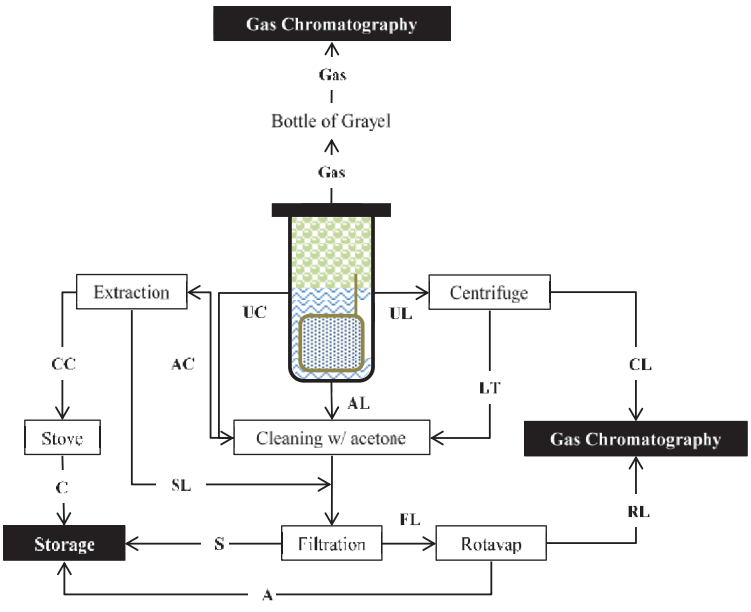
are accomplished as shown in Figure III-3 step *4*. During the process, the fed of hydrogen is kept always at 130 bar and the mixer, which was set at the beginning of the heating, at 1200 rpm.

After finishing the reaction, the hydrogen feed is stopped and the temperature is reduced to around about 15 °C thanks to the cooling system. As demonstrated by step *5* in Figure III-3, the reduction of the temperature is reflected in a decrease of pressure, as expected. When the system is finally thermally stable, the gas is removed from the reactor until the pressure hits 0 bar (relative pressure). The gas is stored in a Grayel bottle. The liquid is retrieved after confirming that is safe to open the batch.

In the case of a non-isothermal test, or blank test, step *4* of Figure III-3 is not performed. Therefore, a blank test corresponds to the temperature dynamic periods of the reactor, meaning the heating and the cooling of the system without the isothermal period in between. These tests are performed for the same mixtures mention in section III.1.1 and with the presence of the catalyst. The final aim of these blank tests is to understand the impact of the dynamic periods of the reactor in the reaction behavior of the compounds.

III.3.4 Post-treatment

At the end of the reaction, the batch contains a gas phase, a liquid phase and the catalyst, as seen in Figure III-4. At the bottom of the reactor, solid residues may exist. However, their petty quantity makes this solid phase negligible and easily separated from the liquid.



Acetone – A; **Acetone Catalyst** – AC; **Acetone Liquid** – AL; **Catalyst** – C; **Centrifuge Liquid** – CL; **Clean Catalyst** – CC; **Filtration Liquid** – FL; **Liquid Traces** – LT; **Rotavap Liquid** – RL; **Solids** – S; **Soiled Liquid** – SL; **Unrefined Catalyst** – UC; **Unrefined Liquid** – UL.

Figure III-4 - Representative diagram of the post-treatment of the reaction.

When both phases are discharged, the thermostat and the mixer, attached to the top of the equipment, are cleaned with acetone. This new liquid is stored and later treated with the all the liquid generated during the cleaning of the reactor, basket, catalyst, and used auxiliary material.

III.3.4.1 Gas phase

As mentioned previously, the gas phase is transfer to a Grayel bottle. The bottle is initially purged with nitrogen and after put under vacuum conditions. The expansion of the gas from a small volume (the reactor top) to a 15 L bottle reduces the pressure from around about 30 bar to 2 bar. The transfer stops when all the gas inside of the reactor disappears, meaning a pressure balance between the batch and the bottle. Once inside the bottle, the gas stays there until a new test.

To analyze the components in the gas, a polypropylene pouch with 1 L and under vacuum conditions, is filled with the gas. After injected into a gas chromatograph, it is possible to unveil the gas phase compounds thanks to analysis method created by Ozagac (Ozagac 2016).

III.3.4.2 Catalyst

The catalyst is retrieved from the catalytic basket emerged in the liquid solution. Then the whole equipment is clean with acetone. The originated liquid is stored in a recipient. Afterwards, the catalyst is extracted from the basket and cleaned with acetone. The remaining liquid joins the previous acetone effluent in the recipient.

The catalyst, now placed in a tray, is dried for approximately 5 min in a stove (510 °C) to help evaporate any remaining compounds. When the particles are no longer humid, the catalyst is transferred into an extraction cell for a more complete cleaning. The extraction is performed in an equipment called DIONEX ASE 150. This is used to clean the catalyst by passing acetone. The DIONEX takes around about 1 h, at 100 °C and 100 bar, to process the catalyst.

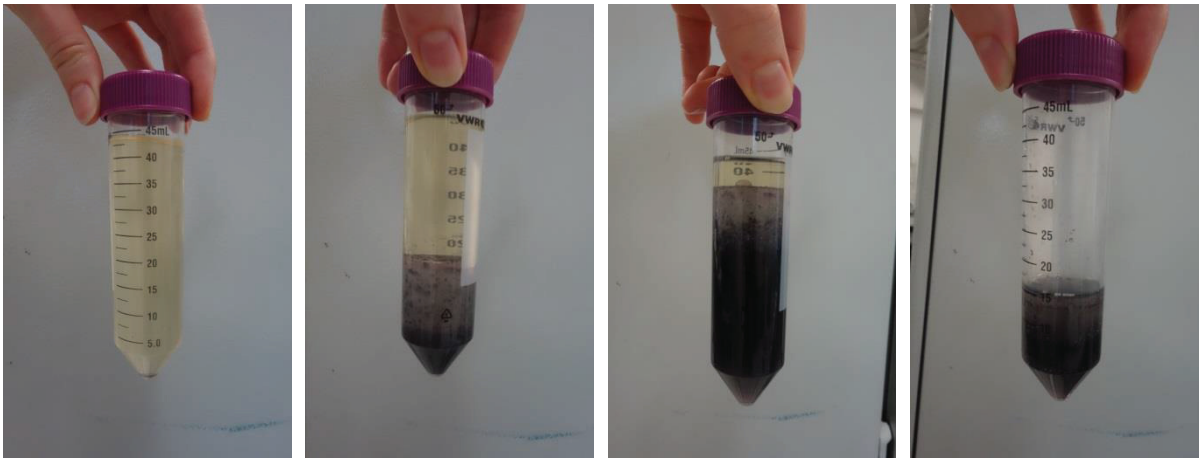
In the end, the cleaned catalyst goes to the stove (510 °C) once again to evaporate the remaining acetone (≈ 2 h). The recovered liquid is joined with the rest of the clean liquid in the recipient. After being dried, the catalyst is stored. Any analytical procedures are performed, as these are nor a part of the aims of the present work.

III.3.4.3 Liquid phase

The liquid phase is divided into two effluents: the main liquid, which has the unconverted reactants, the products, the water and the solvent, and the cleaning liquid, which contains mostly acetone.

The first is transferred to 50 ml centrifuge tubes, as shown in Figure III-5. Due to its multiphasic state, the main liquid is centrifuged during 20 min at 4000 rpm. After a clear separation of the three phases, these are separated thanks to a pipette to their respective flasks.

A cleaning liquid is generated from the washing with acetone of the reactor, the mixer, the thermostat and all the equipment used during the post-treatment. This is stored in addition to the cleaning liquid of the catalyst and its basket, following a vacuum filtration step. Due to the thermal sensibility of the products, this process is done with carefulness to avoid the fusion of the liquid into a white and undesirable solid. The solid residues and remaining catalyst particles retained in the filter are dried in a stove (510 °C) for around about 2 h and later on stored. The clarified liquid is transferred to a rotavap, where the acetone is boiled in vacuum, at 70 °C and 90 rpm, from the mixture and imminently condensed into another recipient. When the condensation of acetone is over, a two phase liquid is obtained. Thanks to a pipette the two liquid products and the acetone are separated.



a) *b)* *c)* *d)*

Figure III-5 – Guaiacol liquid effluent from the centrifuge (CL) after 3 h of reaction at 150 °C: a) lightest phase, b)and c) lightest, intermedium and heavier phase, d) intermedium and heavier phases.

In summary, an experimental test results in three different product, where the liquid represent the majority (see Figure III-6). This output consists into five different samples: three from the centrifuge and two from the rotavap. For the centrifuge, it was assumed that the superior phase was a solvent (n-hexadecane) phase, the intermedium phase was an aqueous phase and the inferior phase was an organic phase. For the rotavap, it was equally assumed that the inferior phase was an organic phase, while the superior phase could be either a solvent or an aqueous phase. Although a phase association can be made between the rotavap and the centrifuge samples, it was decided to analyze all the samples individually.

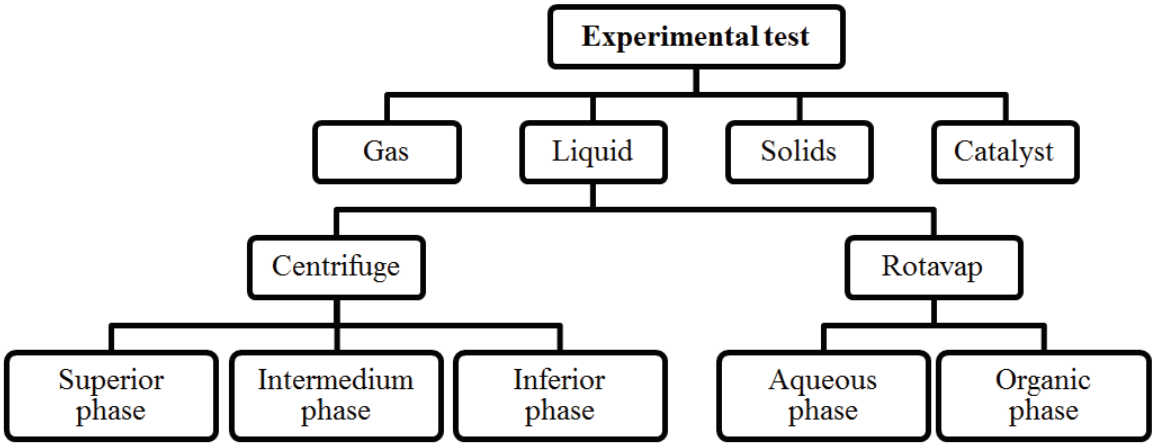


Figure III-6 - Scheme of the effluents for each experimental test.

III.3.4.4 Global Mass balance

To assured that none inconvenient has been found during the test, a balance is done to account the conservation of mass. According to Lavoisier, the mass balance can be represented by Equation III-1.

$$Input = Output + Acumulation \qquad \textit{Equation III-1}$$

Since the reactor and all the equipment were cleaned with acetone, which enters into the mass balance, the accumulation term will be considered equal to zero (no traces were left in the equipment). The *Inputs* present in Equation III-1 consist in the reactant mixture (water, guaiacol and n-hexadecane), the catalyst and the hydrogen. The last, not only accounts with the injected hydrogen, but also with the gas already inside the reactor at the end of the sealing test (see Figure III-3 step *2* and *3*). The *Outputs* are all the recovered liquid phases, except the acetone, the cleaned catalyst, the solids and the gas in the Grayel bottle. The liquid phases included all the effluents who exist the centrifuge and the rotavap. As final balance, Equation III-1 can be rewritten in the form of Equation III-2.

$$m_{mixture} + m_{catalyst} + \left(m_{reactor} + \left[m_0 - m_f\right]_{ballast}\right)_{H_2} = m_{liq\ centri} + m_{liq\ rota} + m_{dried\ catalyst} + m_{gases} + m_{solids} \qquad \textit{Equation III-2}$$

As associated to any experimental procedure, the balance may be affected by mass losses. These can be represented by the evaporation of light products in the stove or remains in the filtration recipient. Although atypical, a mass gain is equally possible if the separation between the effluents and the cleaning liquid (acetone) is not perfect. This feature can only be proved through the analytical results, as the experimental procedure does not include a mass balance to the acetone. For all experimental tests, a percentage error, given by Equation III-3, was applied to the mass balances.

$$Global\ Mass\ Balance\ (\%) = \frac{Output}{Input} \cdot 100 \qquad \textit{Equation III-3}$$

The results are in Table III-12, together with the balance of the experimental data obtained by Ozagac (Ozagac 2016) (only the tests further applied in the present work). The results for the **furfural** tests are in red, while for **guaiacol** tests, these are in blue, as shown in Table III-12. In average, the global mass balance is around about 101 % with a 2 % of standard deviation. For this reason, it was concluded that the experimental tests were valid for further applications.

Table III-12 - Results of the global mass balance for the new and the old set of experimental data.

Global Mass Balance (%)				
Reaction temperature (°C)	**Reaction time (h)**			
	0 h	**1 h**	**3 h**	**5 h**
100 °C	101 %	100 %	/	/
150 °C	/	100 % / 102 %	102 %	102 %
200 °C	/	/	98 %	/
250 °C	106 %	104 %	99 %	/
300 °C	99 %	101 %	101 %	101 %

III.4 Effluents analyses

In the present work, a special attention was paid to the chemical composition of the gas and liquid products in order to achieve key data required for the modeling techniques. The gas products were analyzed by gas chromatography (GC) coupled with a flame ionization detector (FID) and a thermal conductivity detector (GC-FID/TCD) while the liquid products were analyzed by GC coupled with mass spectrometry (GC/MS) and GC-FID for the species identification and quantification respectively.

In a few words, the GC technique consists in the separation for molecular species by affinity with the columns inside the instrument. In practice, the gas phase is injected into the apparatus and carried through the column thanks two carrier gases (mobile phase), helium and argon. Then, the molecules are separated according to their affinity with the column filling (stationary phase). While leaving at different retention times, the molecules can be identified by MS or quantified by FID or TCD. On one hand, the MS technique promotes the ionization of the molecules. After divided into several fragments, these are detected using their mass-to-charge ratio and consequently lead to the identification of the molecules. On the other one hand, the FID recognizes the ions generated from the combustion of the sample gas with hydrogen, while the TCD detects the molecules in the gas sample through the changes in the thermal conductivity of the column. In the present work, this detector is applied to all non-hydrocarbon

molecules, as hydrogen, carbon monoxide, carbon dioxide, nitrogen and oxygen, while the former is particular to hydrocarbons.

In the present work, the GC instruments were equipped with a PONA column with 0.2 mm of diameter, 50 m of length and 0.5 μm of film thickness. For the gas analysis, two additional columns for the non-hydrocarbons were used: a HP-Plot 5A column with 0.32 mm of diameter, 30 m of length and 1 μm of film thickness and a HP-Plot Q column with 0.32 mm of diameter, 30 m of length and 20 μm of film thickness. The follow section will present a brief description of the operating conditions and the calibration procedures. The GC/MS analysis was conducted with a Thermo Trace GC/MS equipped with the same columns and program than the GC/FID analysis. The mass analyzer was a simple quadrupole with an electron ionization ion source at 70 eV. The products were identified using the NIST library.

III.4.1 Gas phase

The gas phase, stored in a polypropylene pouch, was submitted to a GC/MS and GC-FID/TCD analyses for the species identification and quantification, respectively. Similar to Ozagac (Ozagac 2016) procedure, the gas sample, after injected, passed by a 7 min step at 30 °C. Then this follows a heating process until 200 °C according to a 20 °C/min ramp. After 3.5 min at 200 °C, the column was cooled down to the initial temperature. In the present work, the samples were only analyzed once.

The calibration of the GC instrument is done thanks to two different gas bottles. Both the gases and the compositions are known and close to what could be expected, as shown in Table III-13. Any reproducibility trials were performed, being the incertitude obtained by Ozagac (Ozagac 2016) for the same gas phases assumed valid for the new tests: ± 0.015 mol and ± 0.002 mol for hydrogen and hydrocarbons respectively. The identification of the pics is done according to retention times, as the compounds expected exist in the calibration bottles.

Table III-13 – Gases used in the calibration of the GC-FID/TCD analysis of the gas phase.

Bottle number	Compounds (% mol/mol)									
	CH_4	C_2H_6	C_3H_8	C_4H_{10}	C_5H_{12}	CO	CO_2	N_2	H_2	O_2
1	10.00	1.99	1.99	1.99	2.00	1.99	2.00	2.00	76.00	-
2	0.50	0.50	0.50	0.50	0.50	5.00	5.00	85.50	1.00	1.00

III.4.2 Liquid phase

Similar to the work of Ozagac (Ozagac 2016), the liquid effluent was composed by five samples, as shown in Figure III-6. Each sampled was submitted to a GC-FID analysis for the species quantification, except for the organic phase of the rotavap as it was unstable and kept generating an additional phase when sampled. For the compounds identification, the superior and inferior phases of the centrifuge

samples originated by the trials at 1 h and 150 °C for guaiacol and 1 h and 100 °C were analyzed through GC-MS. For the intermedium and rotavap phases, it was assumed the existence of a mix of molecules of the superior and inferior phases.

In the GC instrument, and similar to the work of Ozagac (Ozagac 2016), the liquid samples pass by a 10 min step at 100 °C. Then the samples follow a heating process until 250 °C according to a 5 °C/min ramp. After 10 min at 250 °C, the samples are cooled to the initial temperature. In the present work, the samples were only analyzed once.

The GC instrument was regulated with an external calibration technique. This approach involves the addition of an external compound (non-existent in the samples) to the solutions used for the calibration and thus the identification of the species in the mixtures through it. In the present work, the calibration was achieved thanks to four solutions composed by guaiacol, furfural, propylene glycol and butanol.

An Effective Carbon Number (ECN) method was used to estimate the FID response factor of the oxygenated species. This approach can be used for calculating relative response factors in cases where pure materials are not available for detector calibration, as the case of some oxygenated mixtures (Scanlon and Willis 1985). In summary, the quantification of each species in the samples can be obtained according to Equation III-4:

$$C_i = A_i \cdot \frac{C_{PG}}{A_{PG}} \cdot \frac{k_i}{k_b} \cdot \left(\frac{k_{PG}}{k_b}\right)^{-1} \qquad \textit{Equation III-4}$$

Where C_i and C_{PG} are the concentration of the species i and the propylene glycol, respectively, in g/L, A_i and A_{PG} are the GC-FID peak area of the species i and propylene glycol, respectively, in mm^2, C_{PG}/A_{PG} is the slope of calibration curve of propylene glycol given by Figure III-7 and finally k_i/k_b and k_{PG}/k_b are the relative response factors for species i and propylene glycol, respectively, calculated relative to benzene.

The area of species i comes for the GC-FID, while the ratio between propylene glycol concentration and area is generated by the external calibration curve shown in Figure III-7. The relative response factors are calculated through the ECN method assuming as reference compound benzene. A detailed explanation of this quantification approach can be found in the work of Ozagac (Ozagac 2016).

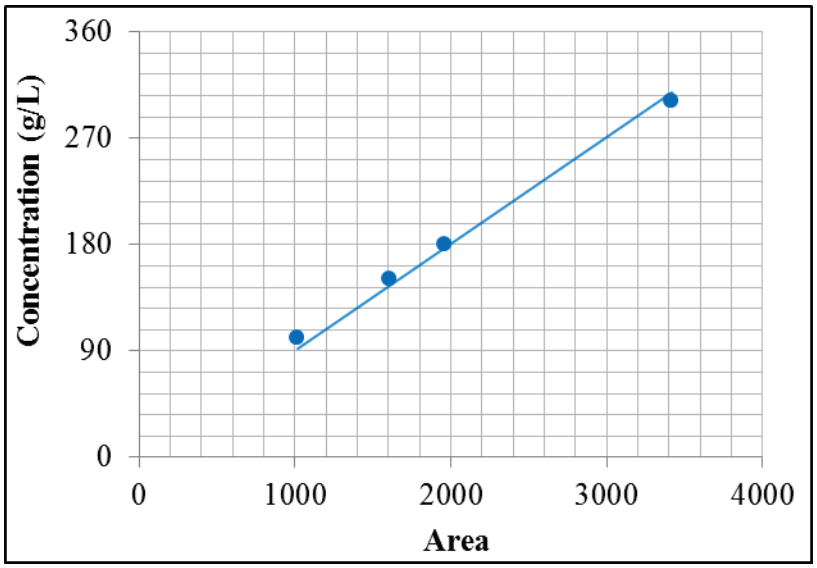


Figure III-7 - Calibration curve for propylene glycol.

III.5 Conclusion of Chapter III

The present chapter summarized the experimental tests performed by Ozagac (Ozagac 2016) at high conversion and presented the new set of conditions tested in the present work at lower conversion. These tests focused on two model compounds: guaiacol and furfural. In order to follow the strategy proposed for the present thesis, the hydrotreating tests performed on both single model compounds and on a bio-oil sample will be used in the present work. In this chapter, the experimental methodology was also presented, together with the analytical approach applied to the new set of experiences. Both methodologies follow the approach applied by Ozagac (Ozagac 2016).

In conclusion, the experimental tests which will be used in the present work were performed at 13 MPa over a reduced $NiMo/Al_2O_3$ catalyst at temperatures ranging from 100 °C to 300 °C and with a reaction time interval of 1 h to 5 h. Three blank tests were also performed (one to guaiacol and two to furfural). In the end, the global mass balance closed very well for all tests, with an average value of 101 %.

Chapter IV. Experimental results

In collaboration with Elodie Dusson, Manuel Menard and Yohann Mouillet

The present chapter will detail the experimental data obtained by Ozagac (Ozagac 2016) and the results obtained for the new set of experimental conditions. From the latter set of results, only the tests that will be used later on in the kinetic models will be analyzed. All the experimental tests discussed in this section are summarized in Table III-12.

The present chapter is divided into four sections: results summary, carbon mass balance, hydrogen consumption and hydrotreating of guaiacol and furfural. These last two sections will highlight the main components in the process and their behavior, in order to later extrapolate these observations for the creation of the kinetic models.

IV.1 Summary of the experimental tests

From the five samples generated per test, only four were analyzed by GC-FID. As it happens, the organic phase of the rotavap was instable, leading to the constant generation of a two-phase sample, even after being sampled. For this reason, it was not possible to analyze this fraction. However, and as mentioned before, it is possible to associate, and thus mix, the rotavap and the centrifuge samples. Therefore, the mass of the non-analyzed sample (organic phase of the rotavap) was added to the centrifuge samples. In the case of guaiacol, the organic phase of the rotavap was joined with the inferior phase of the centrifuge. On the contrary, in the case of furfural, all the organic phases of the rotavap were added to the superior phase of the centrifuge. Indeed, after a successful analysis of one of these phases, the data revealed more similarities with the superior phase than the inferior. As will be presented later, the results extracted from each experimental test (and sample) have a similar appearance to what can be expected for hydrotreating tests. For this reason, it was considered that the assumptions presented here were correct, and thus the results could be used for further applications.

After analyzing the samples, the concentration of each compound, retrieved from Equation III-4, was transformed into mass by means of the density of each major compound in the sample, as previously done by Ozagac (Ozagac 2016). This means that the density of hexadecane, water and guaiacol were used as density of the superior, intermedium and inferior phase, respectively. For the aqueous phase of the rotavap, the density of water was used.

Although the following section will detail the analytical results, it is important to note that for all experimental tests, acetone was found in the samples (around about ≤ 7 %w/w_{feed}), especially in the rotavap effluents. This may reflect an inefficient separation of acetone in the rotavap, justifying the minor differences found in the global mass balance. Furthermore, traces of compounds originated from

acetone (methylpentatone and methylhydroxypentanone) were equally found. These products may be generated by the aldol and condensation reactions. Besides acetone, all experimental tests had hexadecane, the solvent of the feed. However, there were also other compounds originated from the solvent in the samples, as pentadecane, butadecane, etc. These may be generated by the degradation of the solvent or simply be impurities in the initial fraction of hexadecane. In all cases, the presence of these was extremely low and was not quantified. To conclude, it is also important to highlight that although the characterization of the samples was difficult, the best was done to ensure a proper quantification of the compounds.

IV.2 Carbon mass balance

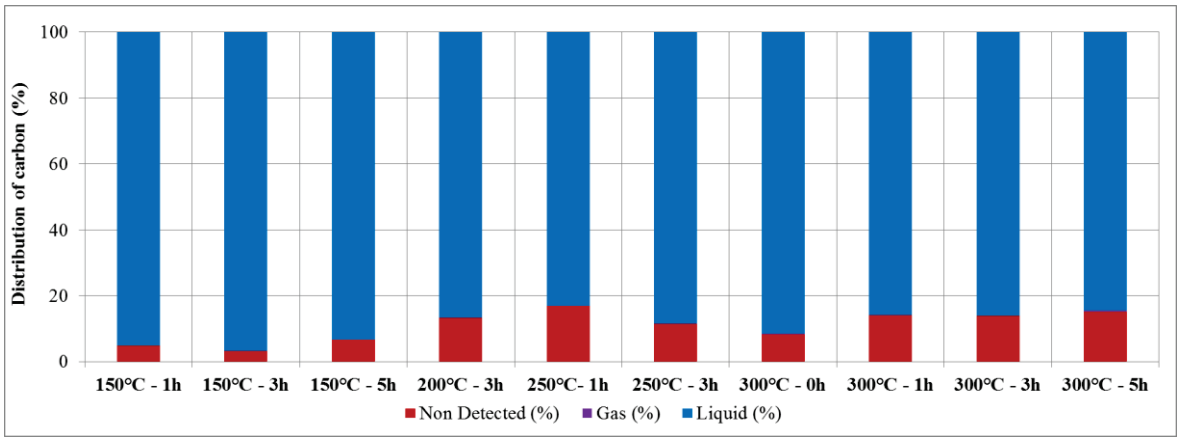
A carbon mass balance was done for all experimental tests. The balance did not account with the solvent (hexadecane) and the undesirable washing liquid, acetone. Moreover, the balance did not account with any compounds originated from any of the two molecules. Equation IV-1 represents the carbon mass balance, where $mC_{guaiacol,0}$ is the initial carbon mass of guaiacol, $mC_{guaiacol,f}$ is the final carbon mass of guaiacol and mC_i is the carbon mass of compound i. The summation rewritten in the expression includes all products from the liquid and gas phases.

$$mC_{guaiacol,0} = mC_{guaiacol,f} + \sum_{i}^{All\ products} mC_i \qquad \textit{Equation IV-1}$$

The previous expression allows the analysis of the distribution of carbon in the all the experimental tests. Figure IV-1 summarizes the results obtained for guaiacol (*a)*) and furfural (*b)*) for all the experimental tests concerned in the present thesis. In the case of guaiacol, between 85 % and 97 % of the carbon in the mixture can be identified, as seen in Figure IV-1 *a).* In the case of furfural, around about 80 % of the carbon can be identified for less severe experimental tests, as seen in Figure IV-1 *b)*. For the trials at 250 °C, around about 60 % of the carbon can be identified. The difference seen in the both blank tests (100 °C and 250 °C) can be justified by the increase of the temperature. In the case of thermosensitive molecules, as furfural, this may lead to undesirable reactions (e.g. polymerization) and thus the production of molecules difficult to analyze. In a more global perspective, the non-detected fractions may be associated to several factors as the undesirable evaporation of products in the rotavap, in addition to the acetone or the incorrect application of the pure components density, used for the conversion of the compounds concentration into mass. Regarding the gas, and as shown in Figure IV-1, this barely has any impact on the effluent. Although there are non-detected molecules, the majority of the experimental tests have over 80 % of carbon in the effluents identified. For this reason, the results here presented were considered suitable for the continuation of the present thesis.

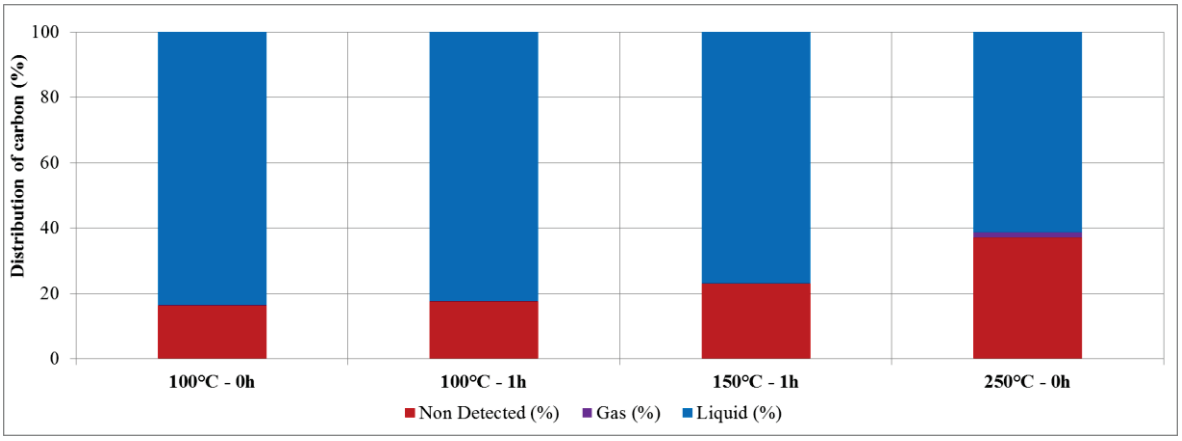
Carbon mass balance

Guaiacol



a)

Furfural



b)

Figure IV-1 – Carbon balance for all experimental tests of: a)Guaiacol; b) Furfural.

The evolution of the reactants and their products, through the temperature and time, is given by the conversion (Equation IV-2) and selectivity (Equation IV-3) of the compounds. The conversion of a given reactant i is obtained by the initial mass weight ($m_{0,i}$) (or concentration) and the molar fraction of the same compound at the end of the reaction ($m_{t,i}$).

$$Conversion_i\ (\%) = \frac{m_{0,i} - m_{t,i}}{m_{0,i}} \cdot 100 \qquad \textit{Equation IV-2}$$

The selectivity of a given product i is calculated thanks to its mass weight (m_i) and the sum of all products (M) detected by GC-FID/MS. By using selectivities, it is considered that the composition of the non-detected compounds is similar to the composition of those detected.

$$Selectivity_i\ (\%) = \frac{m_i}{\sum_i^M m_i} \cdot 100 \qquad \textit{Equation IV-3}$$

IV.3 Hydrogen consumption

In this experimental work, hydrogen is fed to the reactor in a continuous flow, which is stopped at the end of the reaction before the cooling of the system. The consumption of hydrogen is almost constant, reflecting the excess of hydrogen as industrially expected However, during the tests is possible to denoted a slight consume of hydrogen. The consumption of hydrogen is given by Equation IV-4.

$$Hydrogen\ consumption = Hydrogen\ input - Hydrogen\ output \qquad \textit{Equation IV-4}$$

While the output of hydrogen is retrieved from the GC-FID analysis, the input can be calculated through the pressure differences in the system. Besides the hydrogen trapped inside the reactor after the sealing test ($n_{reactor\ at\ t=0}$), the input accounts with the hydrogen fed to the system, by the ballast, between the beginning of the test ($n_{ballast\ at\ t=0}$) and end of the reaction ($n_{ballast\ at\ t=reaction\ time}$). In molar quantities, the hydrogen input is described by Equation IV-5.

$$Hydrogen\ input = n_{reactor\ at\ t=0} + (n_{at\ t=0} - n_{at\ t=reaction\ time})_{ballast} \qquad \textit{Equation IV-5}$$

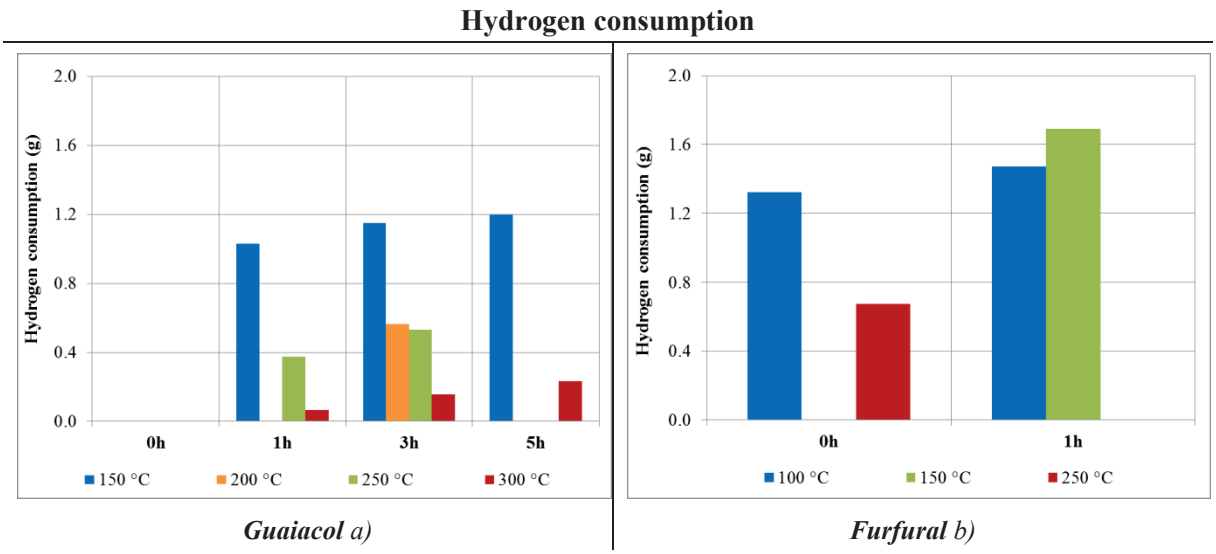
The molar quantities can be obtained thanks to the pressure (P) and temperature (T) data retrieved from the pilot unit digital interface and the ideal gas law showed in Equation IV-6.

$$P \cdot V = n \cdot R \cdot T \qquad \textit{Equation IV-6}$$

The pressure is in Pa, the volume (V) in m^3, the molar quantity (n) in mol, the ideal gas constant (R) in $J/(mol \cdot K)$ and the temperature in K.

The hydrogen consumption for all experimental tests is summarized in Figure IV-2. For the hydrotreating of guaiacol, the consumption of hydrogen increases slightly with the reaction time, but decreases with the temperature. The increase of the reaction time provides the opportunity for the propagation of reactions, as HDT reactions. Therefore, the hydrogen consumption increases as well. Concerning the temperature, higher temperatures lead either to non-HDT reactions, such as transalkylation, or to inverse reactions, such as aromatization of rings, which generate hydrogen. In the case of guaiacol, the last hypothesis is the most probable.

For the hydrotreating of furfural, and similarly to guaiacol, the consumption of hydrogen increases slightly with the reaction time, as shown in Figure IV-2 *b)*. This is probably due to the increase of HDT reactions in the medium. As for the temperature, in the blank tests, from 100 °C to 250 °C, the hydrogen consumption reduces. This may also be due to the increase of the aromatization reactions. However, at 1 h, from 100 °C to 150 °C, the consumption of hydrogen increases. One reason could be, that non-HDT reactions (Ozagac 2016) or aromatization require higher temperatures than 150 °C to occur. Therefore, the increase of the temperature, in this case, leads to the increase of the reactions rate (report to Arrhenius Law) and thus to a higher consumption of hydrogen.



***Guaiacol** a)* ***Furfural** b)*

Figure IV-2 - Hydrogen consumption for all experimental tests of: a) Guaiacol and b) Furfural.

IV.4 Hydrotreating of guaiacol

As mentioned previously, three guaiacol hydrotreating tests were done in the present thesis (report to Table III-10). This section will present the results (reactant and products) for each test, in addition to seven other experimental tests for the hydrotreating of guaiacol performed by Ozagac (Ozagac 2016) in his thesis work. These are all the experimental data of the author concerning the hydrotreating of guaiacol at a constant water content (30 %w/w).

IV.4.1 Conversion of guaiacol

As shown in Figure IV-3, the conversion of guaiacol increases with the temperature, as expected, and in general, with the reaction time. At 3 h, the conversion of guaiacol at 200 ºC is almost four times higher than at 150 ºC. Then, the conversion increases around 5 % more at 250 ºC. At last, the conversion is around 33 % at 300 ºC, meaning that it increased around about 12 % in comparison with the experimental point at 250 °C. In summary, although guaiacol is a not an extremely thermosensitive molecules, the variation in the temperature still affects its conversion. From Figure IV-3, it is possible to conclude that the conversion of guaiacol is quite similar for temperature between 200ºC and 250ºC. However, its conversion drops heavily or increases at lower or higher temperatures, respectively.

Regarding the impact of the reaction time, the conversion of guaiacol has two trends. At higher temperatures (300 ºC), the consumption of guaiacol increases with the time, as shown in Figure IV-3. However, at lower temperatures, the conversion of guaiacol is similar for shorter reaction times, as 1 h and 3 h, and slightly greater for longer reaction times (5 h). This stable baseline of conversion may be due to the competition of guaiacol with its derivatives for the catalytic sites (Wang et al. 2013). If the derivatives of guaiacol are preferred, guaiacol stops being consumed and thus its conversion is constant.

In summary, the conversion of guaiacol during hydrotreating is favored not only at higher temperatures, but also at higher reaction times.

To conclude this section, it is important to highlight that for the blank test, guaiacol converted around about 11 %. This reveals the importance of the dynamic periods of the reactor: heating and cooling. For this reason, it is important to account with both periods in any modeling tool created for the experimental tests here presented.

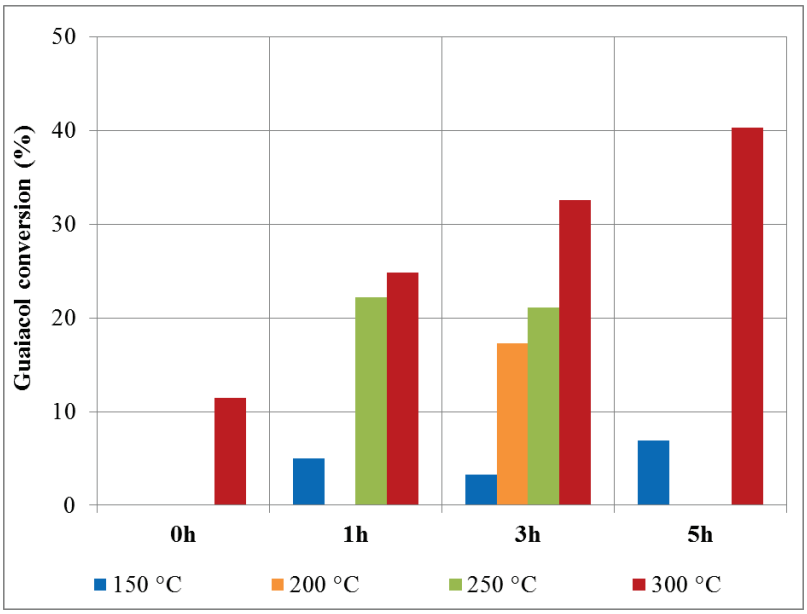


Figure IV-3 - Conversion of guaiacol (%) for all experimental tests.

IV.4.2 Products distribution

The distribution of the produced detected by GC-FID/MS was analyzed according to the reaction temperature and time. For a better understanding of the results, it is recommended to the reader to follow together with the present section the reaction network of guaiacol proposed in Figure V-1 in the Deterministic modeling methodology chapter. This scheme only contains the main compounds.

IV.4.2.1 Effect of the reactor temperature

At lower temperatures, as 150 ºC, guaiacol is mainly converted into methanol, methoxycyclohexane, phenol, methylphenol and dimethycyclopentanone, as shown in Figure IV-4 *a)*. For 1 h, methanol and methylphenol are the main products, followed by dimethycyclopentanone and then phenol and finally methoxycyclohexane. In the increase of the reaction time by two hours (3 h), methanol is the main product with a selectivity of around about 80 %. The remaining compounds are methoxycyclohexane, then phenol and finally methylphenol. Dimethycyclopentanone was not detected at 3 h of reaction. With another increase of two hours in the reaction time (5 h), the products pattern is similar to the trend found at 1 h, with exception that methoxycyclohexane has a higher selectivity than phenol. From the results at

all reaction times, it is possible to conclude that demethoxylation, hydrodeoxygenation, saturation and transalkylation may be the main reactions for the conversion of guaiacol and its derivatives at lower temperatures. However, it is important to remark that the high selectivity of methanol should be accompanied with a high selectivity of phenol or a derivate of phenol as benzene, cyclohexanol or cyclohexane (to keep the mass balance). Since such it is not observed, it could be assumed that not only a given fraction of methylphenol may be originated by the methylation of phenol (Otyuskaya et al. 2017), but also that dimethycyclopentanone may be derivate from phenol. The pathway or ensemble of routes for this last reaction is unknown and inexistent in the open literature.

At medium temperatures (200 ºC and 250 ºC), guaiacol is converted into methycyclohexanediol, methanol, cyclohexanol, methoxybenzene, hydroxycyclohexanone, methoxycyclohexanone, phenol, cyclohexanediol, methylphenol and benzenediol, as shown in Figure IV-4 *b)* and *c)*. The figures also show that methycyclohexanediol is the main compound, followed by methanol, benzenediol and cyclohexanol. At 250 ºC, from 1 h to 3 h of reaction the pattern of products is similar, with exception to cyclohexanediol. The selectivity of this compound increases probably due to the conversion of benzenediol by saturation or of hydroxycyclohexanone, methoxycyclohexanone by hydroxyl-carbonyl equilibrium and demethylation, in the case of the last compound. In summary, at intermedium temperatures, guaiacol and its derivatives convert mainly by demethoxylation, demethylation, hydrodeoxygenation, saturation and transalkylation. At these temperatures, the equilibrium OH↔CO is more predominant.

At higher temperatures (300 ºC), guaiacol is converted into methycyclohexanediol, methanol, cyclohexanol, cyclohexanone, methoxybenzene, hydroxycyclohexanone, methoxycyclohexanone, methylphenol, cyclohexanediol, phenol and benzenediol, as shown in Figure IV-4 *d)*. For all reaction times, the products pattern is similar, with exception to methycyclohexanediol. The decrease of this compound is due to its conversion though by aromatization into methylbenzenediol, which increases with the reaction time according to Figure IV-4 *d)*. Furthermore, methycyclohexanediol may also be transformed into cyclohexanediol by dealkylation of the methyl group. Although cracking reactions are not common for the hydrotreating of guaiacol (Wang et al. 2013), these may start appearing at 300 °C. Besides methylcyclohexanediol, methanol and benzenediol are the main products, followed by cyclohexanediol and phenol, at 300 ºC. The high presence of benzenediol at high temperature and high reaction time may be due to the aromatization of cyclohexanediol. The same reaction may justify the increase of the selectivity of phenol. Furthermore, at 300 h, cyclohexanone was found, meaning that at this temperature the saturated alcohol cycles are still being oxidized. In summary, at high temperatures, guaiacol and its derivatives also convert by demethoxylation, demethylation, hydrodeoxygenation, saturation, aromatization and transalkylation. Moreover, the equilibrium OH↔CO is still present in the mixture and accompanied by dealkylation of the methyl groups.

IV.4.2.2 Effect of the reaction time

For short reaction times, as the blank test and 1 h, guaiacol is mostly converted into methycyclohexanediol, methanol, cyclohexanol, cyclohexanone, methylphenol, cyclohexanediol, phenol, benzenediol and dimethycyclopentanone as shown in Figure IV-5 *a)* and *b)*. With the increase of the temperature, this last compound disappears completely probably by conversion into one or more of the products presented in in Figure IV-5 *a).* As mentioned before the pathway or ensemble of routes for these last reactions are unknown and inexistent in the open literature. The content of methylphenol also reduces with the increase of the temperature. At higher temperatures, this molecule may be transformed into phenol or cyclohexanol by methyl dealkylation and later on saturation. The first reaction may also be the cause of the decrease of methylcyclohexanediol from 250 °C and 300 °C. On the contrary, benzenediol increases with the increase of the temperature, while cyclohexanediol decreases. This may indicate the presence of aromatization reactions, as mentioned before. In summary, at short reaction times, guaiacol and its derivatives also converted by demethoxylation, demethylation, hydrodeoxygenation, saturation, aromatization and transalkylation. Moreover, the equilibrium OH↔CO is still present in the mixture and accompanied by dealkylation of the methyl groups.

For larger reaction times, as 3 h and 5 h, the products behavior is similar to short reaction times, except at 150 °C. At lower temperatures there is mainly methanol, methoxycyclohexane, phenol and methylphenol, potentially derivate from demethoxylation, hydrodeoxygenation, saturation and transalkylation. At higher temperatures, methylhexanediol, cyclohexanol, cyclohexanediol and benzenediol are the compounds with a higher impact. As seen before, the content of methylcyclohexanediol decreases with the temperature. At higher temperatures, this molecule may be transformed into cyclohexanediol or cyclohexanol by methyl dealkylation and later on hydrodeoxygenation. These can also undergo aromatization at high temperatures promoting the generation of phenol, benzenediol and methylbenzenediol (in the case of methylcyclohexanediol), as exposed before. In summary, at short reaction times as at larger reaction times, guaiacol and its derivatives also converted by demethoxylation, demethylation, hydrodeoxygenation, saturation, aromatization and transalkylation. Dealkylation is also possible, as the equilibrium OH↔CO.

To conclude, the reactivity of guaiacol is extremely impacted by the temperature, as the trends seen at 150 °C are not similar to the trends at other temperatures. For this temperature, it can be assumed that demethoxylation hydrodeoxygenation and transalkylation are the main routes, while for the higher temperatures other reactions, as demethylation, methylation, saturation and dealkylation, are also possible. The unalike trends may be caused by the reactivity of the guaiacol derivatives, which could justified the low conversions of guaiacol (> 10 %), together with the high diversity of products at higher temperatures (200 °C – 300 °C).

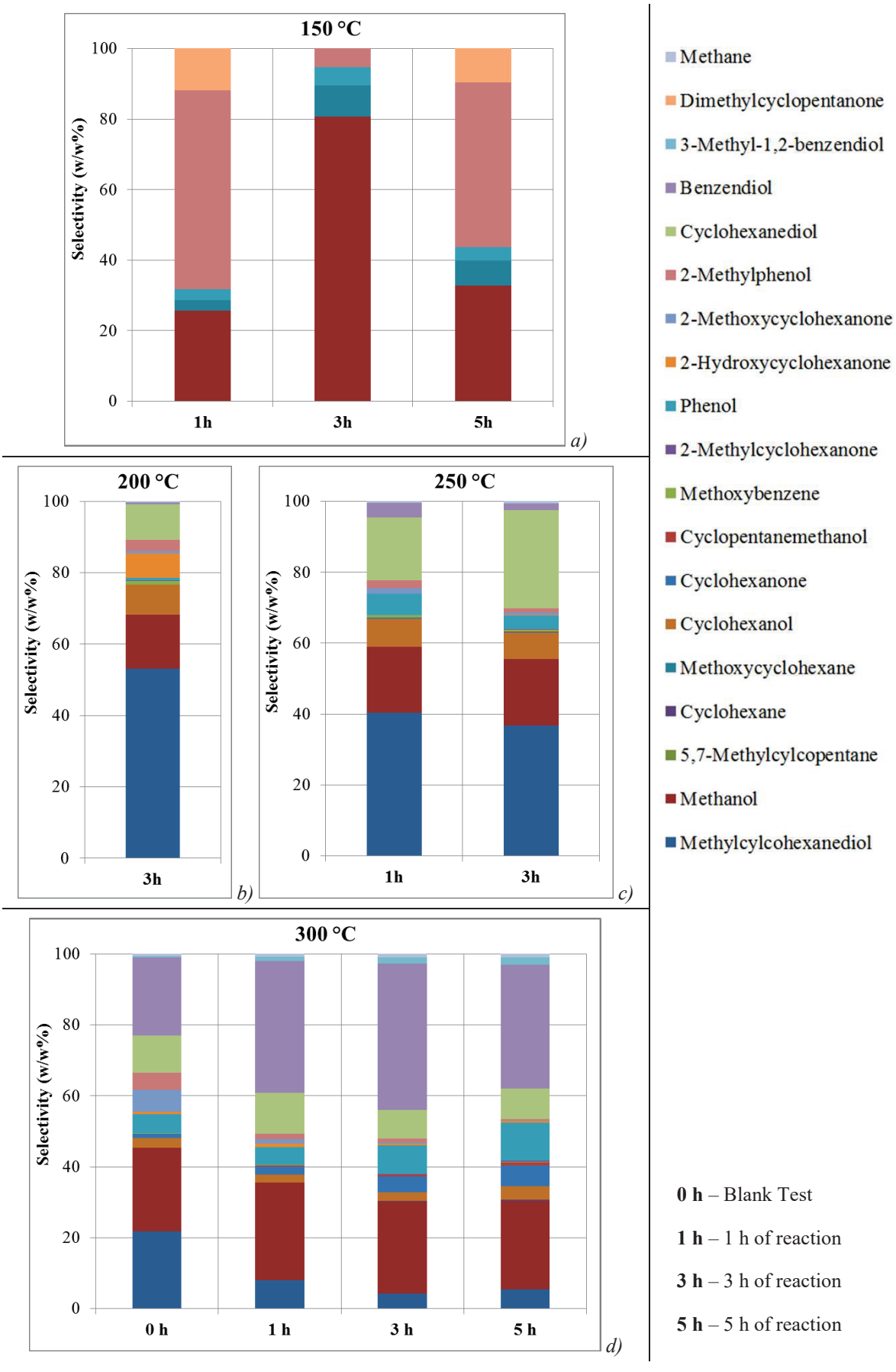


Figure IV-4 - Results of guaiacol HDT at different temperatures: a) 150 °C; b) 200 °C; c) 250 °C and d) 300 °C.

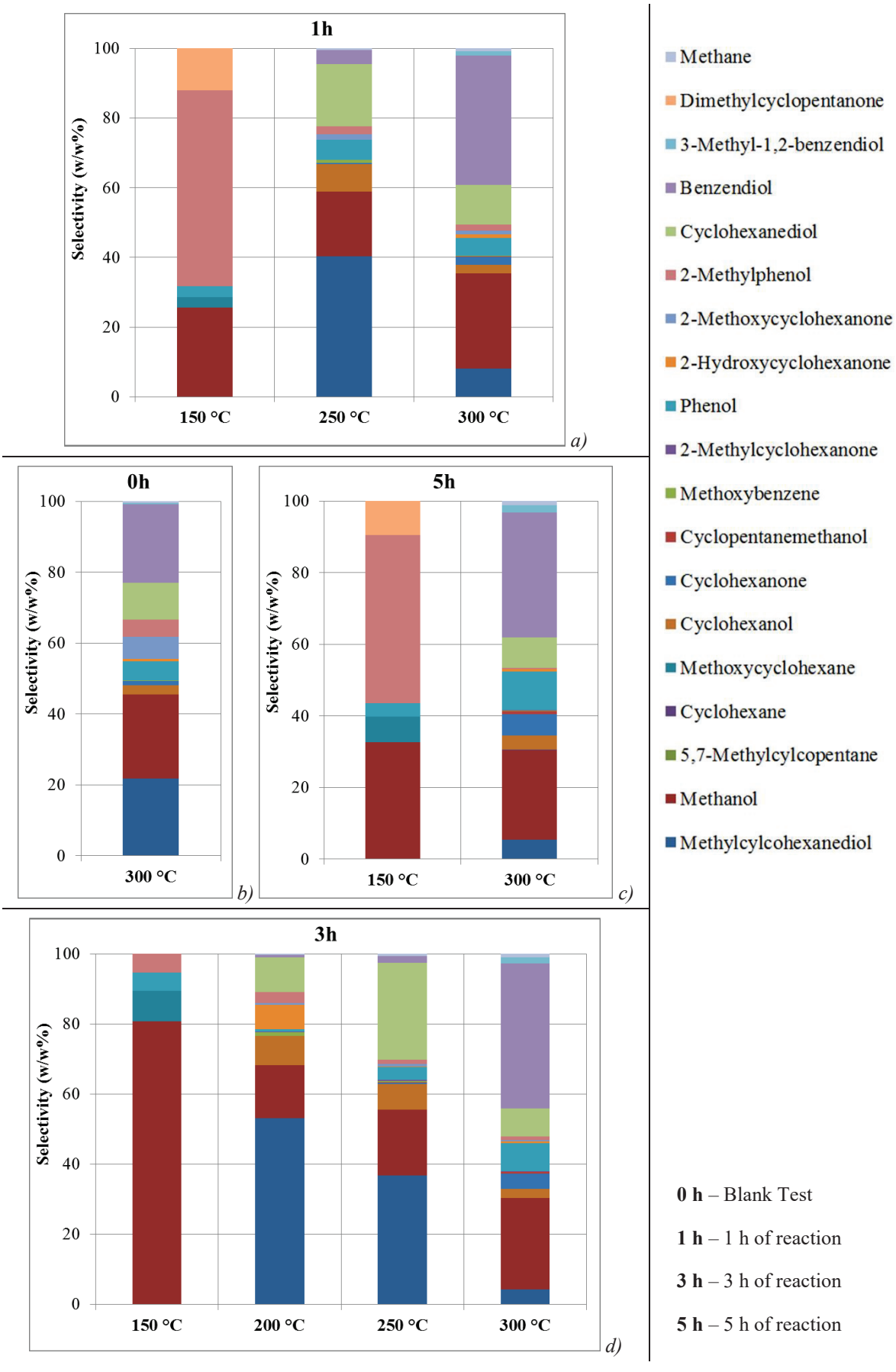


Figure IV-5 - Results of guaiacol HDT at different reaction times: a)1 h; b)Blank test; c) 5 h and d) 3 h.

IV.5 Hydrotreating of furfural

As mentioned previously three furfural hydrotreating tests were done in the present thesis (report to Table III-10). This section presents the results (reactant and products) for each test, in addition to one experimental test for the hydrotreating of guaiacol performed by Ozagac (Ozagac 2016) in his thesis work. The author did several other experimental points for the hydrotreating of furfural. Unfortunately, these were performed at higher conversion (> 40 %) and, thus are not suitable for the application intent in the present work: kinetic models.

IV.5.1 Conversion of furfural

As shown in Figure IV-6, the conversion of furfural increases with the temperature, as expected, and with the reaction time. For the blank test, the conversion increases around about 25 % since the 100 ºC to the 250 ºC. A similar trend can be seen at 1 h of reaction, where the consumption of guaiacol almost doubles after increasing 50 ºC in the reaction temperature. For the reaction time, the conversion of furfural also increases. Since the blank test to a reaction of 1 h, the conversion increases around about 20 %. In summary, and similar to guaiacol, the conversion of furfural is favored by higher temperatures and reaction times.

Concerning the blank tests, furfural converted more than 20 % at 100 ºC and more than 50 % at 250 ºC. Once again, this reveals the importance of the dynamic periods and the importance of accounting them in any modeling tool created for the experimental tests here presented.

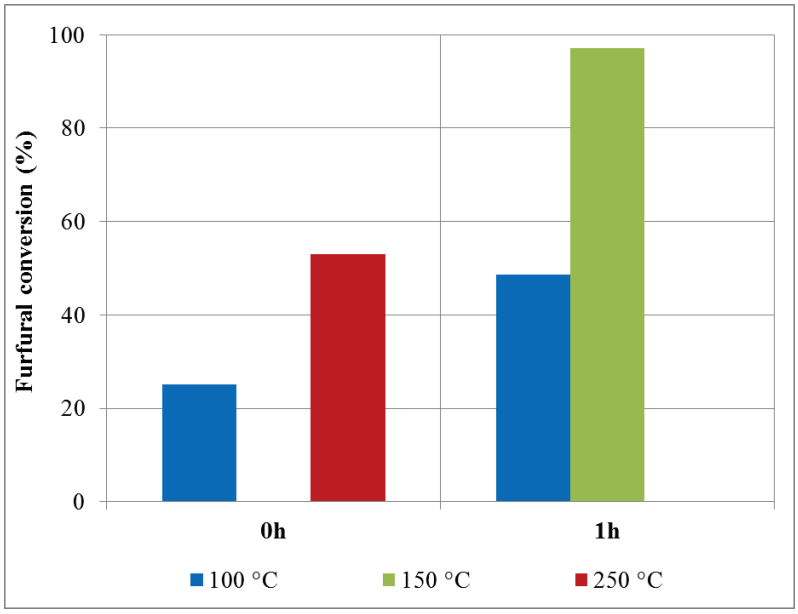


Figure IV-6 - Conversion of furfural (%) for all experimental tests.

IV.5.2 Products distribution

The distribution of the produced detected by GC-FID/MS was analyzed according to the reaction temperature and time. For a better understanding of the results, it is recommended to the reader to follow together with the present section the reaction network of furfural proposed in Figure V-3 of the Deterministic modeling methodology chapter. This scheme only contains the main compounds.

IV.5.2.1 Effect of the reactor temperature

At lower temperatures, as 100 °C and 150 °C, furfural converts mainly to furfuryl alcohol, as shown in Figure IV-7 *a)* and *b)*. Besides furfuryl alcohol, there are also traces of furan, tetrahydrofuran (THF) alcohol, methylfuran, pentanone and butanal, as shown in Figure IV-7 *d)* and *e)*. These two last compounds are more present in the blank test, together with acetylfuran and methoxyphenol. Pentanone and butanal are formed via the opening of aromatic rings, methylfuran and furan, respectively, while acetylfuran may be due to the methylation of the carbonyl in the furfural. Methoxyphenol may be formed through several steps of methylation and dehydration/hydration. With the increase of the reaction time, the content of certain compounds, as THF alcohol, methylfuran and pentanone, varies. However, this variation is extremely small, as shown in the scale of Figure IV-7 *d)* and *e).* This difference is potentially due to the experimental error and thus any conclusion can be extracted from these deviations. In summary, at low temperatures, furfural converts mainly by hydrogenation of the carbonyl into furfuryl alcohol. The remaining molecule traces may probably be originated by the saturation, hydrodeoxygenation and opening of rings.

At higher temperatures, as 250 °C, furfural converts mainly into cyclopentanol, and then furan, methylfuran, butan-2-one, furfuryl alcohol, pentanediol, methoxyphenol and carbon dioxide. This last compound may be originated from the water-gas-swift reaction of monoxide of carbon released during the decarbonylation of furfural. However cyclopentanol, butan-2-one and methoxyphenol are unusual products of the hydrotreating of furfural. These compounds may be due to hydration reactions, as the two first molecules, or to several steps of methylation and dehydration/hydration, as methoxyphenol. The remaining products are typical of the conversion of furfural (maybe of furfuryl alcohol). In summary, at high temperatures, furfural converts mainly by hydration, even if there are still hydrogenation reactions, as hydrogenation of the carbonyl, saturation, hydrodeoxygenation and opening of rings. Furthermore, furfural may also undergo decarboxylation.

IV.5.2.2 Effect of the reaction time

As shown in Figure IV-8 *a)*, the blank test at 250 °C has a higher diversity of products than the blank test at 100 °C. As seen before, these species are mostly derivate from hydration, although other reactions as hydrogenation of the carbonyl, saturation, hydrodeoxygenation and opening of rings can occur as well. In summary, the reactivity of the mixtures increases with the temperature, starting with

hydrogenation reactions, as desired, but quickly evolving to undesired reactions as hydration. This is constraint in the use of a batch reactor, since to achieve high reaction temperature, the mixture must pass by a severe heating period, leading to its conversion.

At 1 h of reaction, the products distribution is similar and leaded by furfuryl alcohol, as shown in Figure IV-8 *b)*. The increase of 50 °C promotes the consumption of the furfuryl alcohol by saturation and hydrodeoxygenation into THF alcohol and methylfuran respectively, as shown in Figure IV-8 *e*). In summary, at low temperatures and 1 h, furfural reacts mainly by hydrogenation of the carbonyl. The remaining molecule traces may probably be originated by the saturation, hydrodeoxygenation and opening of rings.

To conclude, the reactivity of furfural is extremely impacted by high temperatures, as the trends seen at 100 °C and 150 °C are not similar to the trends at 250 °C. For the low temperatures, it can be assumed that hydrogenation of the carbonyl, saturation, hydrodeoxygenation and opening of rings are the main routes, while for the higher temperatures other reactions, as decarbonylation, hydration/dehydration and methylation, are also possible. The presence of these reason may be the reason behind the diversity of products at 250 °C.

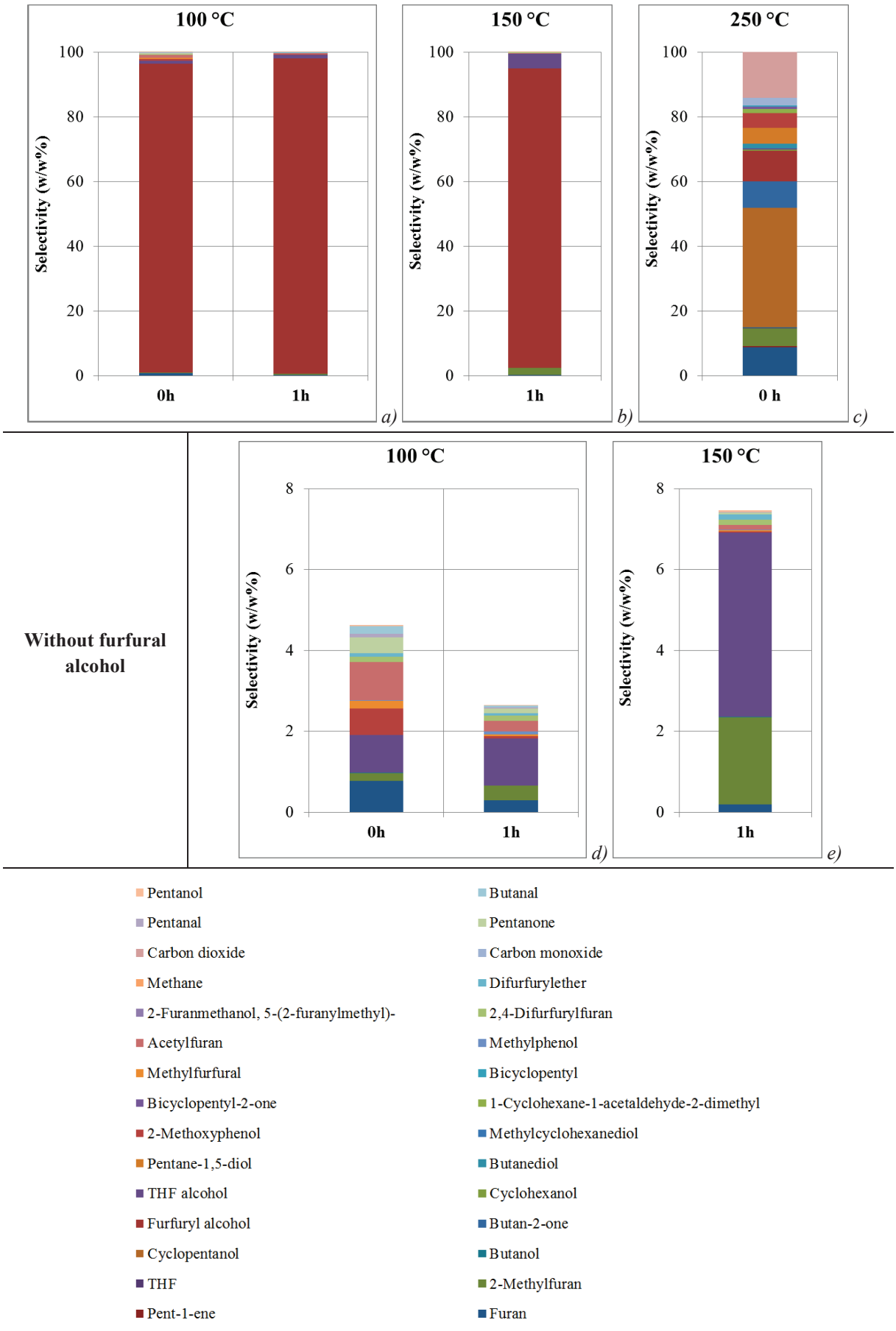


Figure IV-7 - Results of furfural HDT at different temperatures: a) 100 °C; b) 150 °C and c) 200 °C.

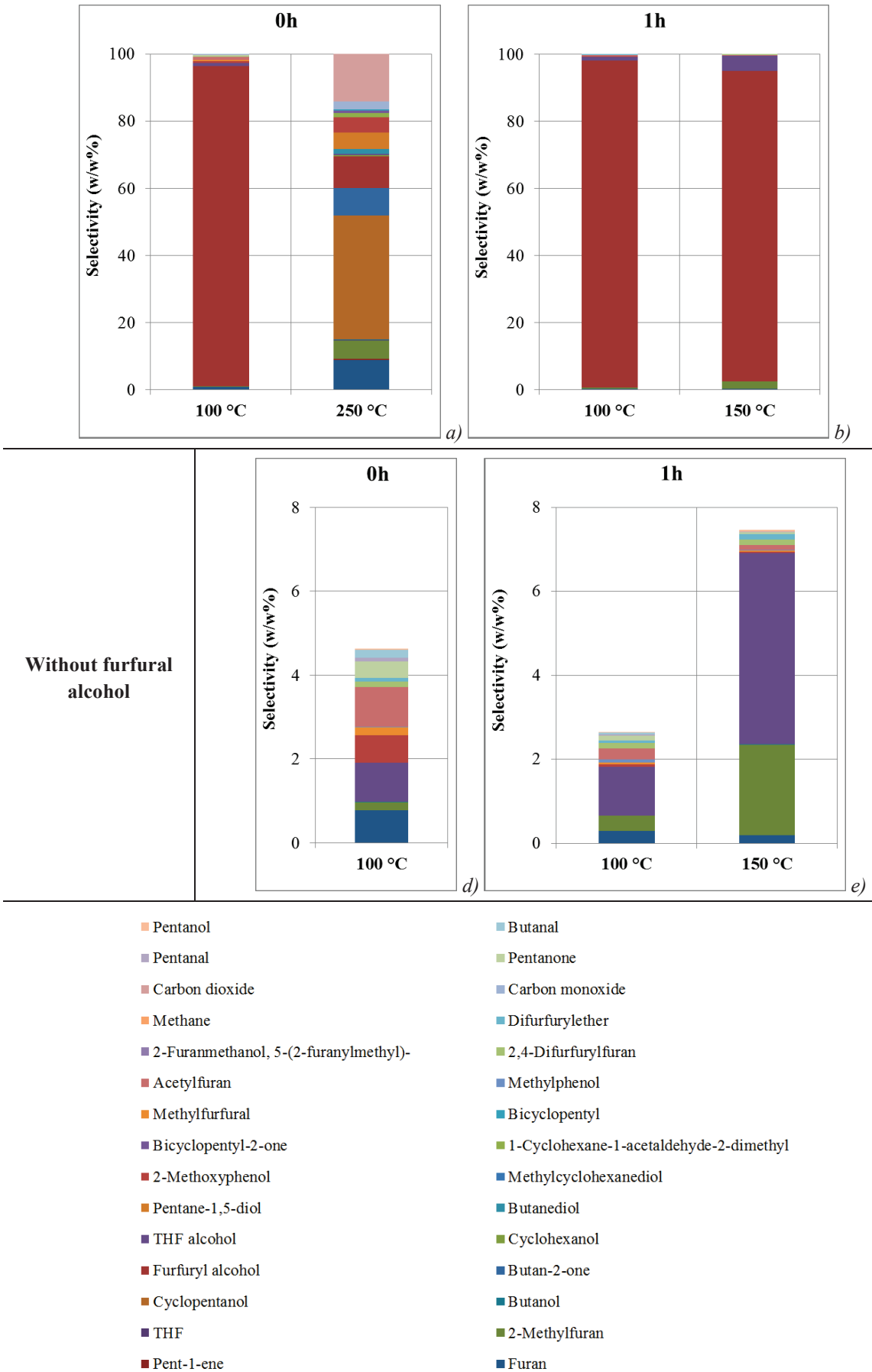


Figure IV-8 - Results of furfural HDT at different reaction time: a) 0 h and b) 1 h.

IV.6 Conclusion of Chapter IV

The experimental tests performed in the present work and by Ozagac (Ozagac 2016) account for more than 80 % of the quantifiable carbon (60 % for the trial at 250 °C). As mentioned before, the characterization of the products was difficult and the best was done to achieve the maximum quantification possible. Be that as it may, with over 80 % quantified, it was assumed that the results were suitable for the application of the strategy of this thesis.

For these experimental tests, the consumption of hydrogen was almost constant, as the hydrotreating tests are performed with an excess of hydrogen. However, it increases with the reaction time, suggesting an increase in the hydrotreating reactions, and it decreases with increasing temperature, indicating that at high temperature hydrogen starts to be produced.

Regarding the conversion of guaiacol and furfural, both increase with the reaction time and temperature as expected. However, a baseline in the conversion of guaiacol between 1 h and 3 h of reaction can be seen. This may be justified by the competition between guaiacol and its derivatives for the catalytic sites (Wang et al. 2013).

Concerning the product distribution, it is quite difficult to identify the reaction pathways of each compound, since these may come from different reactions and several reactants. Be that as it may, the experimental results of both model compounds show that the product distribution is highly affected by the temperature. In the case of the hydrotreating of guaiacol, methanol, methycyclohexanediol, phenol, cyclohexanol, cyclohexanediol, benzenediol and methyphenol are the main products. All these compounds indicate that demethylation, demethoxylation, hydrodeoxygenation, saturation and transalkylation are the main reactions for the hydrotreating of guaiacol. In the case of furfural, furfuryl alcohol is the main product. However, other products such as THF alcohol, furan and methylfuran can also be accounted as secondary products. All these compounds indicate that hydrogenation of the carbonyl, hydrodeoxygenation, saturation, ring opening and decarbonylation are the main reactions for the hydrotreating of furfural. The compounds and reactions here presented will be considered for the creation of a kinetic model for the hydrotreating of guaiacol and furfural.

PART III: Kinetic modeling of model molecules

"Modeling, in its simplest definition, is the process of representing something by using something else in its place."
(Bennett 2009)

Table of Contents

Chapter V. Deterministic modeling methodology

As presented in the State of the Art chapter, there is little information concerning the kinetic parameters of model compounds over reduced NiMo/γ-Al_2O_3. Furthermore, the selected simulation tool, Stochastic Simulation Algorithm (SSA), does currently not include an optimizer, which hinders the parameter estimation of a generated kinetic model. When merged, both increase the complexity to create a kinetic model. For this reason, the model compounds were first modeled through a deterministic approach. The deterministic kinetic model was generated only with the molecules quantified by GC-FID. This approach intends to acquire a range for the values of the kinetic parameters and accelerate the parameter estimation for a larger reaction network when the SSA is used.

The following chapter includes the generation of two kinetic models for two different model molecules: guaiacol and furfural. For both compounds, the chapter presents the proposed reaction network and the kinetic model generation. For the generation of the kinetic model, more details will be given in the guaiacol section and then applied for furfural.

V.1 Hydrotreating of guaiacol

As presented in the Experimental approach section, ten experimental tests representative of the hydrotreating of guaiacol were performed. From these tests, a set of compounds were identified and quantified by GC-MS and GC-FID, respectively.

Based on the results obtained, a deterministic model was generated for the guaiacol hydrotreating. The aim is to understand and estimate the kinetic parameters for the main reactions that guaiacol undergoes during hydrotreating. The development of the model is exemplified in the following section. A brief description of the reaction network and the model creation is included.

V.1.1 Reaction Network

The hydrotreating of guaiacol (G) over a reduced NiMo/γ-Al_2O_3 under 150 °C-300 °C at 13 MPa generates mostly benzenediol (B), phenol (P), methylphenol (ML), methanol (M), cyclohexanediol (CY), cyclohexanol (CL) and methylcyclohexanediol (MY). However other molecules, as methylbenzenediol (MBL), methoxybenzene (MB), methoxycyclohexane (MX), cyclohexane (C) and methane (M4) may also be accounted in order to define a logical reaction network. Certain of these molecules do not have a high selectivity, but all appear in at least one of the ten experimental tests, however. All the compounds are shown in Table V-1.

When displaying all molecules in a network, five reactions can be defined: demethoxylation, demethylation, transalkylation, hydrodeoxygenation and saturation. These reactions match the main

reactions proposed for the hydrotreating of guaiacol over the same catalyst (He and Wang 2012; Wang et al. 2013). As shown in Figure V-1, these five reactions not only represent the guaiacol conversion, but also the transformation of its derivatives. From all the reactions, only the saturation represents the equilibrium between the aromatic reactant and the saturated product.

According to the proposed network, guaiacol may convert through demethoxylation (DEMO), demethylation (DEM), transalkylation (TRANS) and hydrodeoxygenation (HDO). Its derivatives will undergo saturation (SAT) and HDO, except methoxycyclohexane which can only undergo DEMO and DEM. A part from these five reactions, others could be considered, as gas phase reactions or hydroxyl oxidation and carbonyl reduction (reversible reaction). However, the products of such reactions appear in low quantity, making them neglected for the model. Furthermore, according to Wang (Wang et al. 2013), HDO is not an usual reaction for the conversion of guaiacol. However, since the experimental data contradicts such statement, this reaction was accounted as a possible pathway for guaiacol hydrotreating.

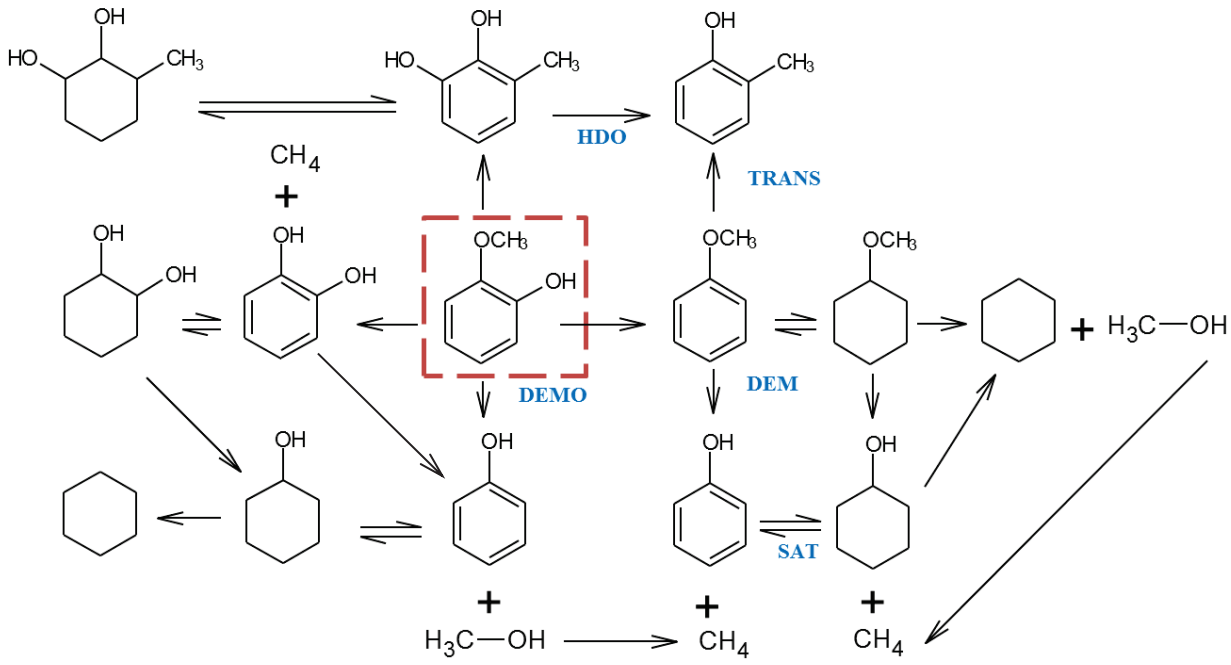


Figure V-1 – Proposed reaction network for the hydrotreating of guaiacol over NiMo/γ-Al2O3 catalyst (the molecule names can the found in Table V-1).

V.1.2 Kinetic model generation

The generation of the kinetic model to guaiacol hydrotreating will be divided in three steps: the model core (how the model is constructed), the evolution of the system (how does the simulation evolve), and the parameter estimation (which involves several simulations to test several parameters until a low value for the objective function is obtained).

V.1.2.1 Model core

The experimental tests were done in a batch reactor with excess of hydrogen. The reactor was under agitation since the start of the heating process until the beginning of the cooling process. For this reason, it is possible to consider a well-mixed and uniform composition at all times. For this case, the material balance over the whole reactor for any given reactant A can be given by Equation III-1.

$$Input = Output + Accumulation + Disappearance \implies 0 = 0 + \frac{dN_A}{dt} + (-r_A) \cdot \rho_B \cdot V \qquad \textit{Equation V-1}$$

Where N_A is the number of moles of compound A, t is the time, dN_A/dt is the variation of the number of moles with the time, $(-r_A)$ is the rate of disappearance of compound A, ρ_B is the bed density (= amount of catalyst per unit of reactor volume), and V is the volume of the reactor. In all experiments, the same weight of catalyst was used and the bed density will therefore be integrated into the kinetic parameters. Moreover, for a constant volume, Equation V-1 is transformed in Equation V-2, where C_A is the concentration of compound A and dC_A/dt is the variation of this concentration over time.

$$-\frac{dC_A}{dt} = (-r_A) \qquad \textit{Equation V-2}$$

As shown in Equation V-2, in an ideal batch reactor, the variation of a given product with time is equal to the disappearance/appearance rate of the same. For a first order reaction (A→B), this parameter, also designated as reaction rate (r_A), can be given by the concentration of compound A (C_A) and the kinetic constant of the reaction (k_v) occurring to the compound inside the reactor, as seen in Equation V-3.

$$r_A = k_v \cdot C_A \qquad \textit{Equation V-3}$$

The kinetic constant can be represented through several correlations. Usually, and as in the present thesis, the constant will be implemented according to the Arrhenius law, given in Equation V-4.

$$k_v = k_{0,v} \cdot exp(-Ea_v/RT) \qquad \textit{Equation V-4}$$

Where k_v is the kinetic constant of reaction v, $k_{0,v}$ is the pre-exponential factor of reaction v, Ea_v is the activation energy of reaction v, R is the ideal gas constant and T is the temperature.

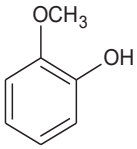
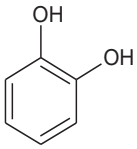
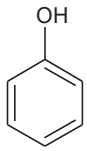
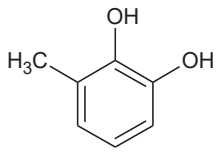
In the present work, all reactions are considered pseudo-first order monomolecular reactions, as hydrogen is in excess. For this reason, all reaction rates for all compounds are represented by Equation V-3. However, since there are several reactions in parallel, the global reaction rate (r_{Global}) of a compound is given by the sum of all reaction rates for each parallel reaction (r_v). While the global reaction rate is given by Equation V-5, the rate for each reaction is given by Equation V-3.

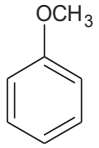
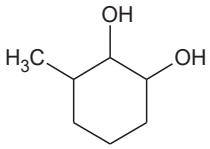
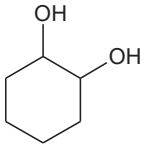
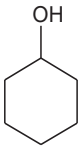
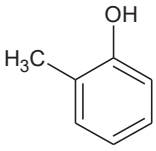
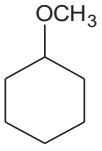
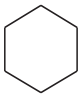
$$r_{Global} = \sum_{\upsilon}^{Q} r_{\upsilon} \qquad \textit{Equation V-5}$$

Where Q is the number of reactions and υ is the reaction in question. It is important to remark that in Equation V-5, each reaction rate must appear with its respective sign. This means that for a disappearance rate, a negative sign (representative of the dissipation of the reactant) must be added.

The application of Equation V-5 to the compounds and reactions present in Figure V-1, generates the expressions in Table V-1. The kinetic model to guaiacol is then described by the Equation V-6 until Equation V-18, where all the kinetic constants are given by Equation V-4. As for SAT, the equilibrium constant (K_{eq}) is calculated by Quantitative Structure Reactivity Correlations (QSRC), as will be shown in the following section. This approach avoids the estimation of the pre-exponential factor and the activation energy of the aromatization, replacing two parameters of the model by one K_{eq}. As a final remark, as hydrogen is in excess, the deterministic model does not account for the effect of hydrogen on the reaction rate.

Table V-1 – Reaction rates for the reactions of guaiacol and its derivatives.

Guaiacol (G) 	$\frac{dC_G}{dt} = -(k_{DEM} + k_{DEMO} + k_{TRANS} + k_{HDO}) \cdot C_G$	*Equation V-6*
Benzenediol (B) 	$\frac{dC_B}{dt} = k_{DEM} \cdot C_G + \frac{k_{SAT}}{K_{SAT}} \cdot C_{CY} - (k_{HDO} + k_{SAT}) \cdot C_B$	*Equation V-7*
Phenol (P) 	$\frac{dC_P}{dt} = k_{DEMO} \cdot C_G + k_{DEM} \cdot C_{MB} + k_{HDO} \cdot C_B + \frac{k_{SAT}}{K_{SAT}} \cdot C_{CL} - k_{SAT} \cdot C_P$	*Equation V-8*
Methylbenzenediol (MBL) 	$\frac{dC_{MBL}}{dt} = k_{TRANS} \cdot C_G - (k_{HDO} - k_{SAT}) \cdot C_{MBL} + \frac{k_{SAT}}{K_{SAT}} \cdot C_{MY}$	*Equation V-9*

Methoxybenzene (MB) 	$\frac{dC_{MB}}{dt} = k_{HDO} \cdot C_G + \frac{k_{SAT}}{K_{SAT}} \cdot C_{MX} - (k_{TRANS} + k_{SAT}) \cdot C_{MB}$	*Equation V-10*
Methyl cyclohexanediol (MY) 	$\frac{dC_{MY}}{dt} = k_{SAT} \cdot C_{MBL} - \frac{k_{SAT}}{K_{SAT}} \cdot C_{MY}$	*Equation V-11*
Cyclohexanediol (CY) 	$\frac{dC_{CY}}{dt} = k_{SAT} \cdot C_B - \left(k_{HDO} + \frac{k_{SAT}}{K_{SAT}}\right) \cdot C_{CY}$	*Equation V-12*
Cyclohexanol (CL) 	$\frac{dC_{CL}}{dt} = k_{SAT} \cdot C_P + k_{HDO} \cdot C_{CY} + k_{DEM} \cdot C_{MX} - \left(k_{HDO} + \frac{k_{SAT}}{K_{SAT}}\right) \cdot C_{CL}$	*Equation V-13*
Methylphenol (ML) 	$\frac{dC_{ML}}{dt} = k_{HDO} \cdot C_{MBL} + k_{TRANS} \cdot C_{MB}$	*Equation V-14*
Methoxycyclohexane (MX) 	$\frac{dC_{MX}}{dt} = k_{SAT} \cdot C_{MB} - \left(k_{DEMO} + k_{DEM} + \frac{k_{SAT}}{K_{SAT}}\right) \cdot C_{MX}$	*Equation V-15*
Cyclohexane (C) 	$\frac{dC_C}{dt} = k_{HDO} \cdot C_{CL} + k_{DEMO} \cdot C_{MX}$	*Equation V-16*
Methanol (M) $H_3C—OH$	$\frac{dC_M}{dt} = k_{DEMO} \cdot (C_G + C_{MX}) - k_{HDO} \cdot C_M$	*Equation V-17*
Methane (M4) CH_4	$\frac{dC_{M4}}{dt} = k_{DEM} \cdot (C_G + C_{MB} + C_{MOX}) + k_{HDO} \cdot C_M$	*Equation V-18*

V.1.2.2 Evolution of the system

After defining the reaction network and the respective equations for each compound, it is possible to construct the simulation for the hydrotreating process through the integration of the concentration in time. The evolution of the chemical system for each compound in given by Equation V-19, where $C_{A,t+\Delta t}$ and $C_{A,t}$ are the concentration of compound A in at $t + \Delta t$ and t respectively and Δt is the interval of time. For each compound, the concentration will follow the previous equation (Equation V-19), where the variation of the concentration with time (dC_A/dt) is given by the Equation V-6 until Equation V-18 for each compound.

$$C_{A,t+\Delta t} = C_{A,t} + \frac{dC_A}{dt} \cdot \Delta t \qquad \textit{Equation V-19}$$

Besides having a periodic evolution, the concentration of each compound depends likewise of the temperature. As seen in the Experimental results chapter, the blank test of guaiacol resulted in around about 10 % conversion of the reactant, revealing the importance of the dynamic period of the temperature in the chemical system. For this reason, the deterministic model was extended to the dynamic period of the temperature: heating and cooling of the reactor.

The inclusion of the temperature behavior is done through the temperature dependent parameters: the kinetic constant and the equilibrium constant. On the one hand, the kinetic constant is related with the temperature thanks to the Arrhenius Law, given in Equation V-4. On the other one hand, the equilibrium constant can be calculated through QSRC (Korre 1994) for a specific temperature and reaction. By repeating the calculation for the set temperatures covered by the temperature profile of experimental tests (≈ 20 °C to 300 °C), a relation between the temperature and the equilibrium constant can be created as presented in Figure V-2. In the case of guaiacol, saturation is the only reversible reaction. As shown in Figure V-1, the reaction network of guaiacol and its derivatives presents four distinct saturation reactions, due to the different functional groups. In order to reduce the complexity of the fitting, it was assumed the same correlation for all reactions. Since benzenediol and cyclohexanediol are the compounds with higher content amongst those who undergo saturation, the correlation considered was from the reverse reaction of benzenediol↔cyclohexanediol.

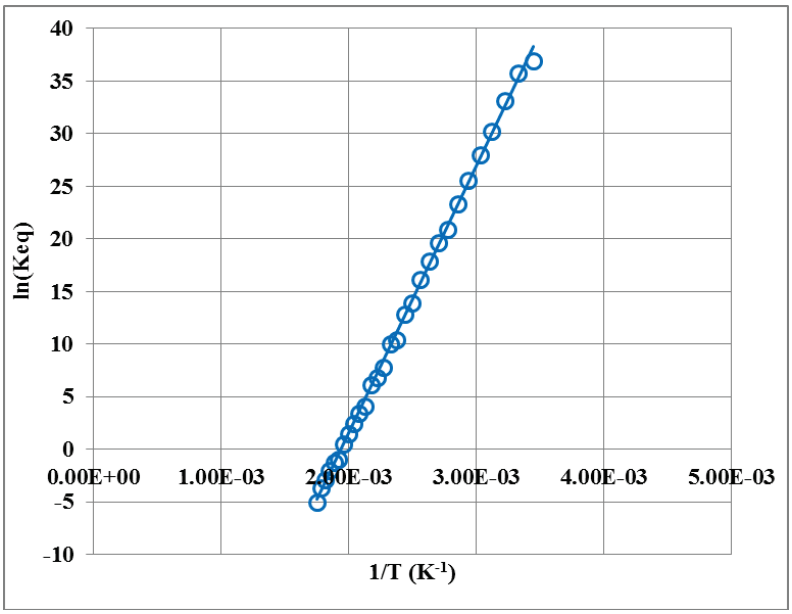
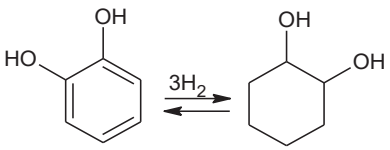


Figure V-2 – Correlation between the ln() of equilibrium constant and the temperature for guaiacol derivatives.

Figure V-2 represents the correlation between the natural logarithm of the equilibrium constant and the inverse of the temperature. Such relation converges into an expression given by Equation V-20.

$$\ln(K_{eq}) = \frac{a}{T} + b \qquad \textit{Equation V-20}$$

This equation is dependent of two parameters: a and b. Both are obtained through the linear regression of the point in Figure V-2. These two values vary according to the data set under analysis, which are intrinsically connected to the reaction type. For the reaction of benzenediol↔cyclohexanediol, the expression for the equilibrium constant is given by Equation V-21.



$$\ln(K_{eq}) = \frac{25\,403.33}{T\ (K)} - 49.33 \qquad \textit{Equation V-21}$$

In the end, the concentration of each compound in the medium will be a function of time and temperature. The evolution of these concentrations will be given by a time and temperature profile obtained from the pilot unit, plus Equation V-4 and from Equation V-6 to Equation V-21. Furthermore, besides guaiacol that has an initial concentration, all the other compounds will have an initial concentration of zero.

V.2 Hydrotreating of furfural

As presented in the Experimental approach section, four experimental tests representative of the hydrotreating of furfural were performed. From these tests, a set of compounds was identified and quantified by GC-MS and GC-FID, respectively.

Similarly, to guaiacol, a deterministic model was generated for the furfural hydrotreating. The aim is to understand and estimate the kinetic parameters for the main reactions underwent by furfural in the process under question. The development of the model was already exemplified in section V.1. However, the following section will describe further details, mostly related to furfural.

V.2.1 Reaction Network

The hydrotreating of furfural (F) over a reduced NiMo/γ-Al_2O_3 under 100 °C-250 °C at 13 MPa generate mainly furfuryl alcohol (FA). However other molecules, as furan (FN), methylfuran (MF), tetrahydrofuran alcohol (TA), tetrahydrofuran (T), pentanediol (PY), pentanone (PO), pentanal (PA), pentanol (PL), butanal (BA), butanol (B) and carbon monoxide (CO) may be accounted as well in order to defined a logical reaction network. Certain molecules amongst these do not have a high selectivity. However, all appear at least in one of the ten experimental tests. All the compounds are shown in Table V-2.

When displaying all molecules in a network, five reactions can be defined: Decarbonylation, carbonyl hydrogenation, ring opening, hydrodeoxygenation and saturation. These match the main reactions proposed for the hydrotreating of furfural over the same type of catalyst (Bunch et al. 2007; He and Wang 2012; Wang et al. 2013). As shown in Figure V-1, these five reactions not only represent the fufural conversion, but also the transformation of its derivatives. In State of the Art chapter, direct hydrogenolysis was presented as a possible pathway for aldehydes. Be that as it may, this reaction was not considered for the network and the model. Firstly, in order to reduce the kinetic parameters to estimate and secondly because the main product of furfural is furfuryl alcohol, followed by furan, which make methylfuran negligible. Contrary of guaiacol, the present reaction network not only accounts with the equilibrium between saturation↔aromatization, but also with the equilibrium between carbonyl reduction (hydrogenation)↔hydroxyl oxidation.

According to proposed network, furfural may convert through decarbonylation (CO) or carbonyl hydrogenation (COH). Its derivatives may afterwards undergo ring opening (RO), hydrodeoxygenation (HDO) and saturation (SAT). Furthermore, if a furanic ring opens, a carbonyl is generated, meaning that decarbonylation (CO) and carbonyl hydrogenation (COH) are also two possible reactions. A part from these five reactions, others could be considered, as gases reactions or hydration. However, the products of such reactions appear in low quantity, making them neglected for the model.

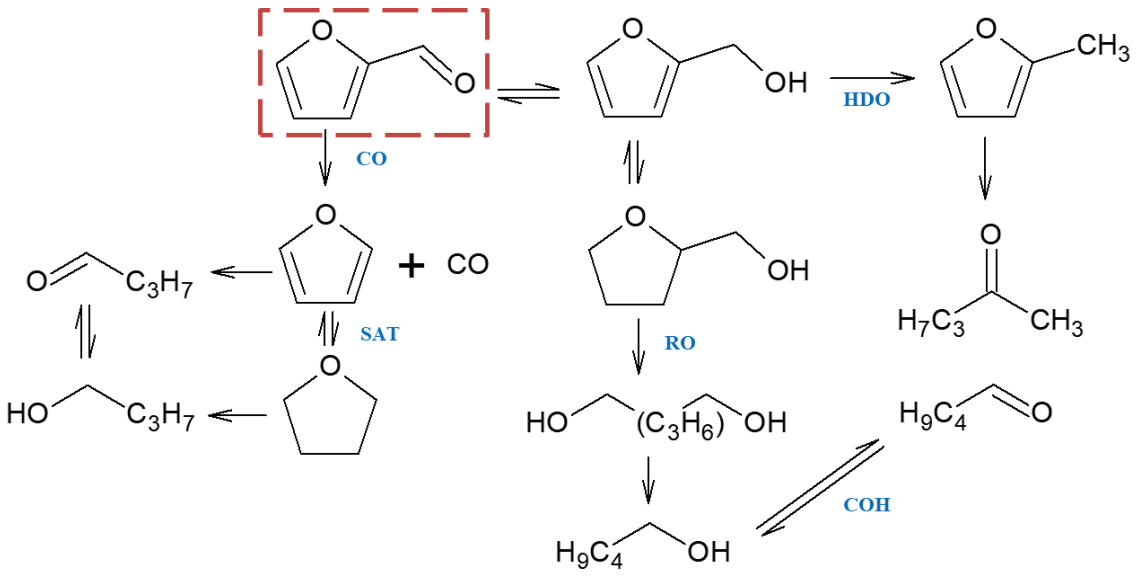


Figure V-3 – Proposed reaction network for the hydrotreating of furfural over NiMo/Al_2O_3 catalyst (the molecule names can the found in Table V-2)

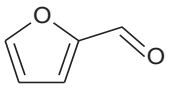
V.2.2 Kinetic model generation

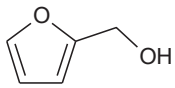
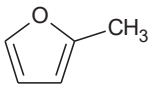
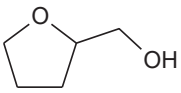
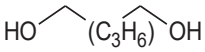
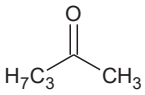
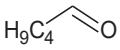
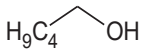
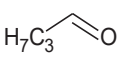
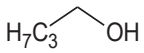
The generation of the kinetic model to furfural hydrotreating will be divided into three steps: the model core, how the model is constructed, the evolution of the system, how does the simulation evolve, and finally the parameters estimation, which involves several simulations to test several parameters until a small objective function is obtained.

V.2.2.1 Model core

The generation of the furfural model is exactly the same as explained in section V.1.2.1. For the case of furfural, the application of Equation V-5 to the compounds and reactions present in Figure V-3, generates the expressions in Table V-2. The kinetic model to furfural is then described by the Equation V-22 until Equation V-34, where all the kinetic constants are given by Equation V-4. As for SAT and carbonyl hydrogenation (COH), the equilibrium constant (K_{eq}) is calculated by Quantitative Structure Reactivity Correlations (QSRC), as it will be shown in the following section. Similar to guaiacol, this approach avoids the estimation of the pre-exponential factor and the activation energy of the aromatization and hydroxyl oxidation, replacing four parameters of the model by two equilibrium constants K_{eq}.

Table V-2 – Reaction rates for the reactions of furfural and its derivatives.

Furfural (F) 	$\frac{dC_F}{dt} = k_{COH} \cdot K_{COH} \cdot C_{FA} - (k_{CO} + k_{COH}) \cdot C_F$	*Equation V-22*

Furfuryl alcohol (FA) 	$\frac{dC_{FA}}{dt} = k_{COH} \cdot C_F + \frac{k_{SAT}}{K_{SAT}} \cdot C_{TA} - (k_{HDO} + k_{SAT} + k_{COH} \cdot K_{COH}) \cdot C_{FA}$	*Equation V-23*
Furan (FN) 	$\frac{dC_{FN}}{dt} = k_{CO} \cdot C_F + \frac{k_{SAT}}{K_{SAT}} \cdot C_T - (k_{RO} + k_{SAT}) \cdot C_{FN}$	*Equation V-24*
Methylfuran (MF) 	$\frac{dC_{MF}}{dt} = k_{HDO} \cdot C_{FA} - k_{RO} \cdot C_{MF}$	*Equation V-25*
Tetrahydrofuran Alcohol (TA) 	$\frac{dC_{TA}}{dt} = k_{SAT} \cdot C_{FA} - \left(k_{RO} + \frac{k_{SAT}}{K_{SAT}}\right) \cdot C_{TA}$	*Equation V-26*
Tetrahydrofuran (T) 	$\frac{dC_T}{dt} = k_{SAT} \cdot C_{FN} - \left(k_{RO} + \frac{k_{SAT}}{K_{SAT}}\right) \cdot C_T$	*Equation V-27*
Pentanediol (PY) 	$\frac{dC_{PY}}{dt} = k_{RO} \cdot C_{TA} - k_{HDO} \cdot C_{PY}$	*Equation V-28*
Pentanone (PO) 	$\frac{dC_{PO}}{dt} = k_{RO} \cdot C_{MF}$	*Equation V-29*
Pentanal (PA) 	$\frac{dC_{PA}}{dt} = k_{RO} \cdot C_{MF} + k_{COH} \cdot K_{COH} \cdot C_{PL} - k_{COH} \cdot C_{PA}$	*Equation V-30*
Pentanol (PL) 	$\frac{dC_{PL}}{dt} = k_{HDO} \cdot C_{PY} + k_{COH} \cdot C_{PA} - k_{COH} \cdot K_{COH} \cdot C_{PL}$	*Equation V-31*
Butanal (BA) 	$\frac{dC_{BA}}{dt} = k_{RO} \cdot C_{FN} + k_{COH} \cdot K_{COH} \cdot C_B - k_{COH} \cdot C_{BA}$	*Equation V-32*
Butanol (B) 	$\frac{dC_B}{dt} = k_{RO} \cdot C_T + k_{COH} \cdot C_{BA} - k_{COH} \cdot K_{COH} \cdot C_B$	*Equation V-33*
Carbon monoxide (CO)	$\frac{dC_{CO}}{dt} = k_{CO} \cdot C_F$	*Equation V-34*

V.2.2.2 Evolution of the system

For furfural, the evolution of the system is exactly the same as presented in section V.1.2.2. As previously seen, the kinetic constant is related with the temperature thanks to the Arrhenius Law, as posed in Equation V-4, while the equilibrium reactions may be described by an equilibrium constant, which is calculated through QSRC for a specific temperature. When plotting several temperatures with their respective constant, as shown in Figure V-4, it is possible to extract a correlation in the form of Equation V-20. For the reactivity of furfural, Equation V-21 was also considered for saturation. The intent is to estimate the parameters of the reactivity of furfural, while assuming that similar reactions, as HDO and saturation, have the same kinetic parameters. Regarding the equilibrium OH↔CO, there are three distinct reactions, due to the different functional groups. Therefore, and in order to reduce the complexity of the fitting, it was assumed the same correlation for all reactions. Since furfural and furfuryl alcohol are the compounds with higher content which undergo HDT of the carbonyl, it was considered the correlations of the reverse reaction of furfural↔furfuryl alcohol.

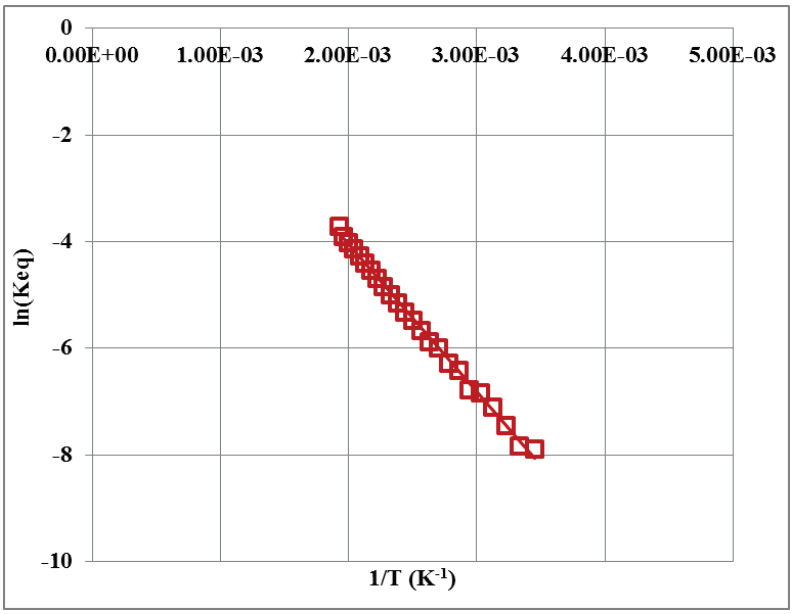
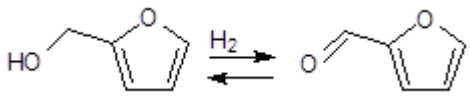


Figure V-4 – Correlation between the ln() of equilibrium constant and the temperature for furfural derivatives.

Figure V-4 represents the correlation between the natural logarithm of the equilibrium constant and the inverse of the temperature given by Equation V-35 for the HDT of the carbonyl.



$$\ln(K_{eq}) = \frac{-2\,774.87}{T\ (K)} + 1.50 \qquad \textit{Equation V-35}$$

In the end, the concentration of each compound in the medium will be a function of time and temperature. The evolution of these concentrations will be given by a time and temperature profile

obtained from the pilot unit, plus Equation V-4 and from Equation V-22 to Equation V-35. Furthermore, besides guaiacol that has an initial concentration, all the other compounds will have an initial concentration of zero.

V.3 Conclusion of Chapter V

The chapter presented the creation of two deterministic models to describe the conversion of the model compounds used in the present thesis: guaiacol and furfural. These models are based on experimental tests of both model compounds and mass balances for each compound in a batch reactor.

The reaction network of each model compound is generated from the consumption of the reactants and the compounds with higher selectivity. Furthermore, the networks also include other molecules which are required to define a logical structure for the reaction network. All molecules used in the models appear at least in one of the experimental tests. The reaction network of guaiacol accounts for demethoxylation, demethylation, transalkylation, hydrodeoxygenation and saturation reactions, while the reaction network of furfural accounts not only for hydrodeoxygenation and saturation reactions, but also for decarbonylation, carbonyl hydrogenation and ring opening.

A kinetic model was generated based on these reaction networks. Through the mass balance of a batch reactor, several equations representative of the reactivity of each compound were written. These expressions include not only the kinetic term, but also the behavior of the compounds in the reactor. Moreover, besides accounting for the evolution of all major compounds during the reaction, the reactor models also include the dynamic periods of the reactor. These were considered in the kinetic constant and in the equilibrium constants used for the carbonyl hydrogenation and saturation reactions. These last constants were extracted from the open literature and assumed equal for all reactions of the same family (all SAT have the same constants, as do all COH), similar to the kinetic parameters.

Chapter VI. Deterministic results

This chapter will present the results of the estimation of the kinetic parameters for both model compounds. The estimated parameters will then serve as initial parameters for the stochastic models, as mentioned previously. Therefore, the aim of the deterministic approach is to acquire an idea of the kinetic parameters range, which is rather reduced in the open literature, and then extrapolate it for the stochastic models of the model compounds.

The following chapter includes the results of the two kinetic models. Here, the experimental data used, the estimation procedure and a discussion will be presented for both model compounds.

VI.1 Hydrotreating of guaiacol

As seen in Experimental approach section, the hydrotreating of guaiacol was performed over a reduced NiMo/γ-Al_2O_3 and at 13 MPa ten times at different reaction times and temperatures. Based on the results obtained from these tests and the reactor model created in chapter V.1.1, it is possible to estimate the kinetic parameters of the reaction families considered on the reaction network. These optimized parameters and theirs results are presented here, together with the experimental data used.

VI.1.1 Experimental data

As mentioned before, the kinetic model was obtained using ten experimental tests of the hydrotreating of guaiacol (at constant water content). These were obtained in a batch reactor under 13 MPa of hydrogen over a reduced NiMo/γ-Al_2O_3 at a range of temperature between 150 °C to 300 °C and a range of reaction time between 1 h to 5 h. A blank test (0 h) at 300 °C was also done.

To fit with the set of expression defined for each compound in chapter V.1.2, the experimental results were transformed from mass (original form) into molar concentration. This conversion was done for all compounds included in the reaction network by Equation VI-1.

$$C_i = \frac{m_i}{M_{w,i} \cdot V} \qquad \textit{Equation VI-1}$$

Where C_i is the molar concentration of compound i, m_i is the mass of compound i, $M_{w,i}$ is the molecular weight of compound i and V is the volume of the reactor, which is 0.5 L. The experimental concentrations used for the estimation of the kinetic parameters are summarized in Table VI-1.

To conclude, and as mentioned before, the models also account for the dynamic periods of the batch reactor. For each experimental test, the profile of the temperature during the trial was inserted in the algorithm. These profiles are extracted from the software behind the pilot unit, as mentioned before.

Table VI-1 - Experimental concentration used in the estimation of the kinetic parameters of the hydrotreating of guaiacol.

Temperature (°C)	**150**			**200**		**250**			**300**	
Reaction time (h)	**1**	**3**	**5**	**3**	**1**	**3**	**0**	**1**	**3**	**5**
Compounds	**Concentration (mol/L)**									
Guaiacol initial	0.727	0.726	0.728	0.727	0.726	0.725	0.727	0.726	0.726	0.726
Guaiacol final	0.690	0.702	0.678	0.601	0.565	0.572	0.644	0.545	0.489	0.434
Benzenediol	0.000	0.000	0.000	0.001	0.007	0.003	0.021	0.075	0.110	0.115
Cyclohexane	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
Cyclohexanediol	0.000	0.000	0.000	0.013	0.031	0.045	0.009	0.022	0.020	0.027
Cyclohexanol	0.000	0.000	0.000	0.013	0.016	0.014	0.003	0.005	0.007	0.014
Methane	0.000	0.000	0.000	0.003	0.005	0.006	0.003	0.012	0.017	0.024
Methanol	0.036	0.075	0.064	0.073	0.117	0.112	0.076	0.192	0.238	0.283
Methylbenzenediol	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.004	0.006
Methylcyclohexanediol	0.000	0.000	0.000	0.064	0.062	0.054	0.017	0.014	0.010	0.015
Methylphenol	0.023	0.001	0.027	0.004	0.004	0.002	0.005	0.004	0.003	0.001
Methoxybenzene	0.000	0.000	0.000	0.001	0.001	0.001	0.000	0.000	0.000	0.000
Methoxycyclohexane	0.001	0.002	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Phenol	0.002	0.002	0.003	0.001	0.013	0.008	0.006	0.012	0.025	0.041

VI.1.2 Estimation of parameters

Since the kinetic constants for all the reactions follow the Arrhenius law, ten kinetic parameters (pre-exponential factor and activation energy) may be accounted for the five reactions proposed in the guaiacol reaction network. The number of parameters already considers the replacement of the parameters of the aromatization by the equilibrium constant, calculated by QSRC.

The parameters were estimated based on the experimental data. For the estimation, a sum of the squared errors criterion was considered and is given in Equation VI-2, where $C_{i,j}^{exp}$ and $C_{i,j}^{cal}$ are the experimental and predicted values, respectively, for the concentration of compound i in the experimental test j, Z is the number of experimental tests, N is the number of compounds in the model and w_i is the weight given to each compound.

$$SSQ = \sum_{j}^{Z} \sum_{i}^{N} \left[\left(C_{i,j}^{exp} - C_{i,j}^{cal} \right)^2 \cdot w_i \right] \qquad \textit{Equation VI-2}$$

The objective is to reduce at the maximum the SSQ by modifying the kinetic parameters, which varies the values of the calculated concentration. The set of parameters which allows the smallest value of SSQ is the optimum for a correct model and, consequently, a suitable simulation of the hydrotreating of

guaiacol over reduced NiMo/γ-Al_2O_3. To achieve this optimum, an initial set of kinetic parameters based on open literature (Otyuskaya et al. 2017) (Grilc et al. 2014) was used. These values were then modified to improve the adjustment to the experimental data, either manually, or via the Generalized Reduced Gradient (GRG) algorithm.

Besides the kinetic parameters, the model has a weight parameter associated to each compound. Thanks to these weights, the model is able to obtain a better refinement of its results and increase its precision concerning certain desired characteristics. The Table VI-2 presents the values for the weights applied in the parameter estimation. As shown, the compounds at higher quantity in the mixture have a higher weight in order to be better predicted.

Table VI-2 - Weights for the guaiacol and its derivatives.

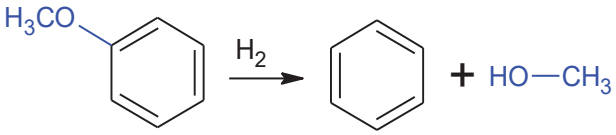
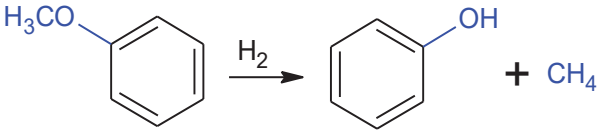
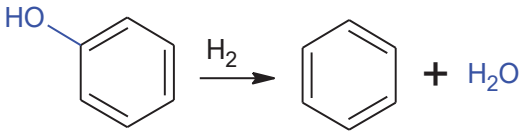
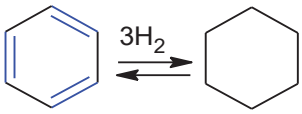
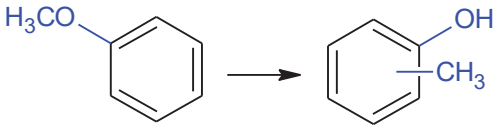
Temperature (°C)	**150**			**200**		**250**			**300**	
Reaction time (h)	**1**	**3**	**5**	**3**	**1**	**3**	**0**	**1**	**3**	**5**
Compounds	**Weight**									
Guaiacol initial	100	100	100	100	100	100	100	100	100	100
Guaiacol final	1	1	1	100	100	100	100	100	100	100
Benzenediol	1	1	1	1	1	1	1	1	1	1
Cyclohexane	1	1	1	100	100	100	100	100	100	100
Cyclohexanediol	1	1	1	1	1	1	1	1	1	1
Cyclohexanol	1	1	1	1	1	1	1	1	1	1
Methane	100	100	100	100	100	100	100	100	100	100
Methanol	1	1	1	1	1	1	1	1	1	1
Methylbenzenediol	1	1	1	100	100	100	100	100	100	100
Methylcyclohexanediol	1	1	1	1	1	1	1	1	1	1
Methylphenol	1	1	1	1	1	1	1	1	1	1
Methoxybenzene	1	1	1	1	1	1	1	1	1	1
Methoxycyclohexane	100	100	100	100	100	100	100	100	100	100
Phenol	100	100	100	100	100	100	100	100	100	100

VI.1.2.1 Results

Based on the experimental information, the model structure previously created and the procedure of estimation of parameters, it was possible to adjust the predicted concentrations to those of the experimental data. As mentioned before, the adjustment is done thanks to the fitting of the kinetic parameters. In the present case, the number of parameters to be estimated is ten, two per reaction, and they are summarized in Table VI-3. For the estimation of the parameters, only the molecules with higher selectivity and the reactant were accounted for. However, it is important to note that certain compounds

with higher selectivity, such as phenol, cyclohexanediol and methylphenol, could not be correctly fitted. The best was done to ensure a proper fitting of the remaining compounds, in particular of the major effluent: guaiacol. In summary, and considering that the aim of the deterministic model is to formulate a range for the kinetic parameters, the procedure applied here was considered suitable.

Table VI-3 – Estimated kinetic parameters for the HDT of guaiacol at 13 MPa over a reduced NiMo/γ-Al_2O_3 at a range of temperature between 150 °C to 300 °C and a range of reaction time between 0 h to 5 h.

Name	Reaction	k_0 (/s)	Ea (kJ/mol)
Demethoxylation (DEMO)		1 300	75
Demethylation (DEM)		1 500	82
Hydrodeoxygenation (HDO)		100	80
Saturation (SAT)		130	60
Transalkylation (TRANS)		3 000	75

When faced with the data of the open literature for CoMo and NiMo/Al_2O_3 catalyst, the estimated activation energies are within the range of obtained by Laurent and Delmon (Laurent and Delmon 1994) for the conversion of guaiacol (59 kJ/mol - 111 kJ/mol for a sulfided NiMo and 71 kJ/mol - 113 kJ/mol for a sulfided CoMo). When compared with the parameters obtained by Otyuskaya et al. (Otyuskaya et al. 2017) for anisole, the parameters of demethylation and transalkylation have a lower activation energy, although hydrodeoxygenation has a activation energy within the range saw by the authors (74 kJ/mol – 82 kJ/mol for a reduced CoMo). The deviations between the parameters may be due to the solvent and mainly the molecules structure.

In Figure VI-1, there are the results from each compound used in the fitting. As illustrated, guaiacol, methanol, benzenediol, methane, methylcyclohexanediol and cyclohexanol were the effluents used for the estimation of the kinetic parameters. In a general perspective, certain compounds and experimental

points fit reasonably well, while others are overestimated or underestimated. To help analyzing the results, and due to the uncertainty of the experimental results, a range of values (minus and plus 10 % of the values of the full line in grey) was defined and is given by the dash lines in grey.

Guaiacol is the major compound in the process effluent and therefore considered the most important in the estimation. For this compound, the model predicts reasonably well the concentration of guaiacol at lower temperatures (150 ºC and 200 ºC) and also for the blank test at 300 ºC. The remaining experimental points (250 ºC and 300º C) are underestimated, as shown in Figure VI-1 *a)*. This indicates that the model converts more guaiacol than it should at higher temperatures. The cause of such differences is not trivial in the case of guaiacol. Besides the fact that this compound may undergo all defined reactions, the absence of other reactions, such as methylation or dealkylation (report to Experimental results chapter) or even an unsuitable equilibrium constant and approximation may also lead to deviations. Even more importantly, catalyst deactivation may also play a significant role, especially since deviations mainly occur at higher temperatures and longer residence times.

Concerning the effluents, methanol experimental points are all underestimated, although the data at 200 ºC, 250 ºC and 300 ºC (except at 5 h) is inside or at the marge of the 10 % deviation. The deviations presented in as shown in Figure VI-1 *b)* may originate from the kinetic parameters of demethoxylation or hydrodeoxygenation. Methanol can also be impacted by the lack of gas phase reactions, as decomposition or steam reforming of methanol (Ozagac 2016) or other bimolecular reactions (J. P. Diebold 2000).

Methane is also impacted by the absence of gas reactions, as methanation of carbon monoxide or dioxide (Ozagac 2016). These compounds are not originated from guaiacol, but from gas reactions involving methanol. In the present case, the concentration of methane is overestimated by the model, as shown in Figure VI-1 *d)*. At 150 ºC, methane was not detected in the effluents and for these the model predicts concentration values small enough to be considered as zero (report to appendix B). Besides the lack of the equilibrium reaction of methanation, other reactions as methylation may also cause such deviations. Moreover, unsuitable kinetic parameters for the hydrodeoxygenation and demethylation reactions could equally impact the predicted concentration.

Concerning benzenediol, the model overestimates the experimental points at lower temperature (200 ºC and 250 ºC) and underestimates those at higher temperature (300 ºC), as seen in Figure VI-1 *c)*. At 150 ºC, benzenediol was not detected in the effluents and for these the model predicts concentration values small enough to be considered as zero (report to appendix B). As mentioned before, the kinetic parameters of demethylation may not be the most suitable for the mixture under question. This may be the reason for the deviations in the concentration of benzenediol at lower temperatures, as for methane. The deviations at higher temperatures may be promoted by aromatization reactions, which as seen in the Experimental results chapter, increase at 300 ºC. Although these reactions were considered by the

implementation of the equilibrium constant, an unsuitable constant and respective approximation (note that the same constant was assume for all saturation reactions) can explain such deviations.

Contrary to benzenediol, the model underestimates the concentration of methylcyclohexanediol at lower temperatures (200 ºC and 250 ºC) and overestimates those at higher temperature (300 ºC), as seen in Figure VI-1 *e)*. At 150 ºC, methylcyclohexanediol was not detected in the effluents and for these the model predicts concentration values small enough to be considered as zero (report to appendix B). Similar to benzenediol, the deviations related to this compound may be related to aromatization and hydrodeoxygenation reactions. Moreover, unsuitable kinetic parameters for transalkylation or even the absence of methylation may impact the concentration of methylcyclohexanediol.

Finally, the model predicts reasonably well the concentration of cyclohexanol. Although certain experimental points are overestimated or underestimated, all are near or within the 10 % deviation area. Regarding the experimental data at 150 ºC, cyclohexanediol was not detected in the effluents and for these the model predicts concentration values small enough to be considered as zero (report to appendix B).

In summary, the proposed kinetic model can predict the concentration of certain compounds at given experimental conditions, while the remaining components have deviations to the experimental values. As mentioned in the discussion above, catalyst deactivation, unsuitable kinetic parameters, absence of gas reactions or bimolecular reactions, as methylation, or even an unsuitable equilibrium constant and respective approximation may lead to the differences presented in Figure VI-1. Be that as it may, it is important to reference that the estimation of the kinetic parameters associated to the reaction network presented is not simple. As discussed in the State of the Art, the reactivity of the bio-oil compounds is affected by the catalyst, the temperature, the solvent and the molecule structure. In the present case, catalytic sites and the molecule structure are not considered in the kinetic model, meaning that the competition between the compounds for the catalytic sites is neglected, alongside the different reactivities associated to the different molecules. As consequence, both reasons are intrinsically connected to the shortcomings in the estimation of the parameters and the fitting of the concentrations.

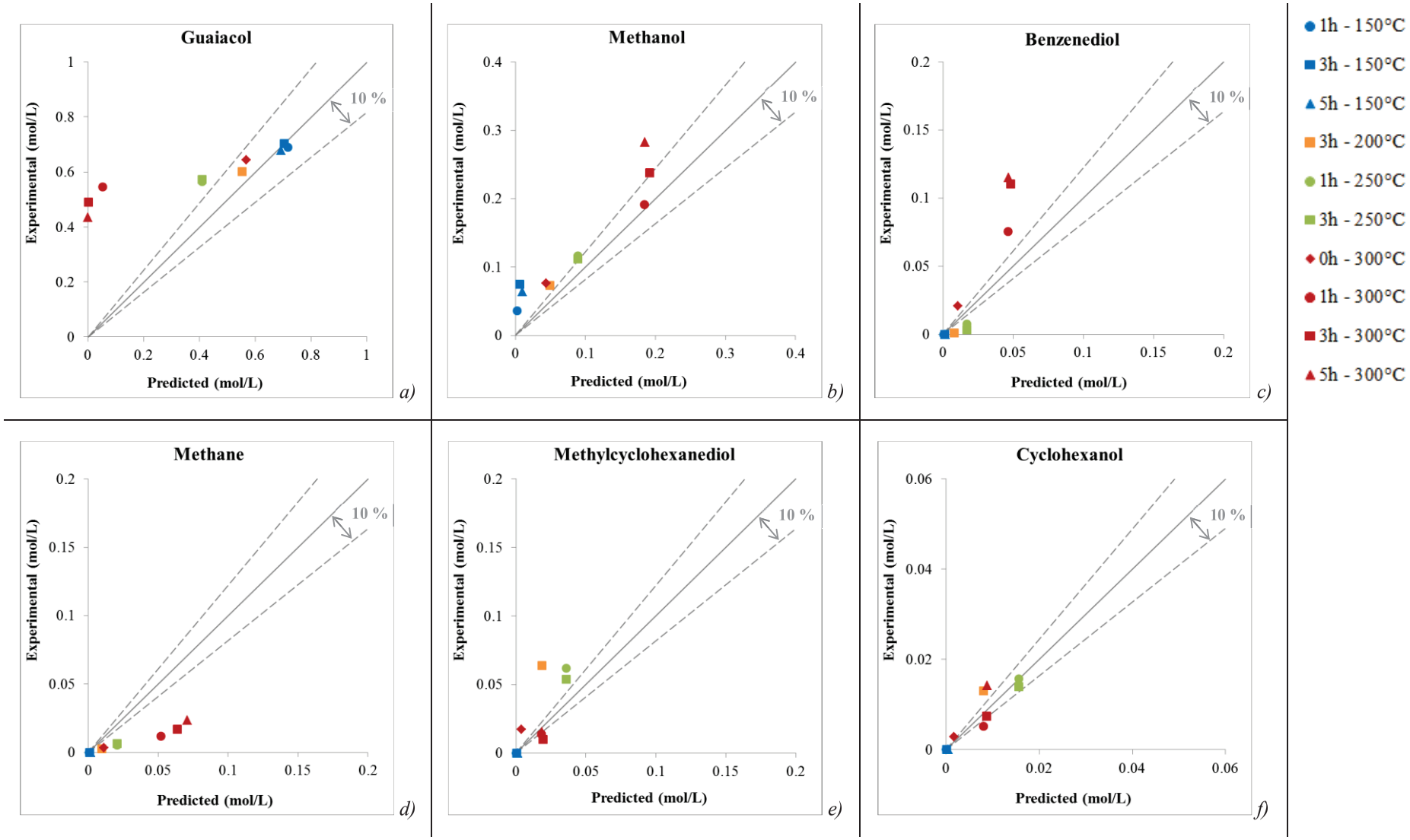


Figure VI-1 – Parity plots of the deterministic model of guaiacol HDT: a) Guaiacol, b) Methanol, c) Benzenediol, d) Methane, e) Methycyclohexanediol and f) Cyclohexanol.

VI.2 Hydrotreating of furfural

As seen in the Experimental approach section, the hydrotreating of furfural was performed over reduced NiMo/γ-Al_2O_3 and at 13 MPa at different reaction times and temperatures. Thanks to the results exploited from these tests and the reactor model created in chapter V.2.1, it is possible to estimate the kinetic parameters of the reaction families considered on the reaction network. These optimized parameters and their results are presented here, together with the experimental data used.

VI.2.1 Experimental data

As mentioned before, the kinetic model was obtained thanks to the four experimental tests of the hydrotreating of furfural (at constant water content). These were obtained in a batch reactor under 13 MPa of hydrogen over a reduced NiMo/γ-Al_2O_3 at a range of temperature between 100 °C to 150 °C for 1 h. Two blank tests (0 h) at 100 °C and 250 °C were also performed.

Similar to guaiacol, to fit with the set of expression defined for each compound in chapter V.2.2, the experimental results were transformed from mass (original form) into molar concentration. This conversion was done for all compounds included in the reaction network by Equation VI-1, previously presented. The experimental concentrations used for the estimation of the kinetic parameters are summarized in Table VI-1.

Table VI-4 - Experimental concentration used in the estimation of the kinetic parameters of the hydrotreating of furfural.

Temperature (°C)	**100**		**150**	**250**
Reaction time (h)	**0**	**1**	**1**	**0**
Compounds	**Concentration (mol/L)**			
Furfural initial	4.2E-03	4.3E-03	4.3E-03	4.3E-03
Furfural final	3.0E-01	2.1E-01	1.2E-02	2.0E-01
Butanal	2.5E-04	1.3E-04	0.0E+00	0.0E+00
Butanol	1.6E-05	8.8E-06	9.4E-05	7.8E-04
Carbon Monoxide	0.0E+00	0.0E+00	0.0E+00	5.1E-03
Furan	1.1E-03	8.1E-04	1.1E-03	2.7E-02
Furfuryl alcohol	9.6E-02	1.9E-01	3.6E-01	2.0E-02
Methylfuran	2.3E-04	8.7E-04	1.0E-02	1.4E-02
Pentanal	1.0E-04	6.1E-05	4.8E-05	0.0E+00
Pentanediol	0.0E+00	0.0E+00	0.0E+00	1.0E-02
Pentanol	3.4E-05	2.5E-05	1.8E-04	0.0E+00
Pentanone	4.5E-04	2.7E-04	1.8E-04	0.0E+00
Tetrahydrofuran	0.0E+00	0.0E+00	0.0E+00	1.3E-04
Tetrahydrofuran alcohol	8.9E-04	2.2E-03	1.7E-02	1.0E-03

To conclude, and as mentioned before, the models also account for the dynamic periods of the batch reactor. For each experimental test, the profile of the temperature during the trial was inserted in the algorithm. These profiles are extracted from the software behind the pilot unit, as mentioned before.

VI.2.2 Estimation of parameters

Similarly to guaiacol, the kinetic constants for all the reactions follow the Arrhenius law, meaning ten kinetic parameters (pre-exponential factor and activation energy) may be accounted for the five reactions proposed in the furfural reaction network. The number of parameters already considers the use of the equilibrium constants for aromatization and oxidation of the hydroxyl. These are calculated by QSRC, as previously seen.

In the case of furfural, the parameters were once again estimated using the experimental data. For the estimation, the sum of the squared errors, described in Equation VI-2, was once again considered. As explained before, the set of parameters which allows the smallest value of SSQ is the optimum for a correct model and, consequently, a suitable simulation of the hydrotreating of furfural over reduced NiMo/γ-Al_2O_3. To achieve this optimum, an initial set of kinetic parameters based on open literature (Otyuskaya et al. 2017) (Grilc et al. 2014) was used. These were then modified in order to adjust the experimental profiles to the experimental data, either manually or via the Generalized Reduced Gradient (GRG) algorithm.

Table VI-5 – Weights for the furfural and its derivatives.

Temperature (°C)	**100**		**150**	**250**
Reaction time (h)	**0**	**1**	**1**	**0**
Compounds	**Weight**			
Furfural initial	100	100	100	100
Furfural final	1	1	1	1
Butanal	1	1	1	1
Butanol	1	1	1	1
Carbon Monoxide	10	10	10	10
Furan	100	100	100	100
Furfuryl alcohol	10	10	10	10
Methylfuran	1	1	1	1
Pentanal	1	1	1	1
Pentanediol	1	1	1	1
Pentanol	1	1	1	1
Pentanone	1	1	1	1
THF	1	1	1	1
THF alcohol	100	100	100	100

Also, similarly to guaiacol, the model for furfural has a weight parameter associated to each compound. In Table VI-5 is shown the weights applied in the parameter estimation for the furfural reaction network. Once again, the compounds with higher impact on the mixture are associated with to a higher weight value and vice-versa.

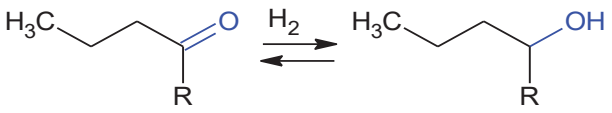
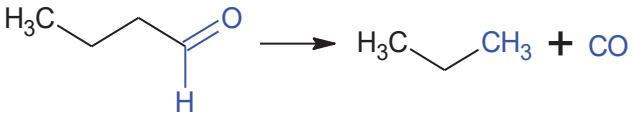
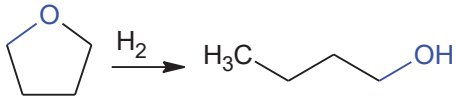
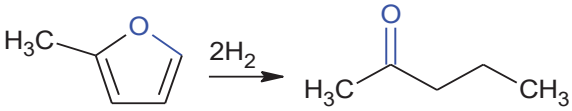
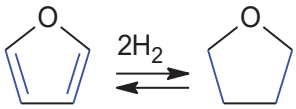
VI.2.2.1 Results

Based on the experimental information, the generated core of the model and the estimation procedure, it was possible to adjust the predicted concentration to those of the experimental data. As mentioned before, the adjustment is done thanks to the fitting of the kinetic parameters. In the present case, ten parameters are estimated (two per reaction) and are summarized in Table I-18. For the estimation of the parameters not only furfuryl alcohol (molecule with higher selectivity) and furfural were accounted for. Other smaller molecules, as THF alcohol, methylfuran, furan and butanol, were included in the estimation as well. This approach allows the estimation of all the kinetic parameters, instead of only the parameters of the main reaction: furfural into furfuryl alcohol by carbonyl hydrogenation. However, it is important to note that not all compounds (e.g. furan) could not be fitted correctly. The best was done to ensure a proper fitting of the remaining compounds, in particular of the major effluent molecules: furfuryl alcohol. In summary, and considering that the aim of the deterministic model is to formulate a range for the kinetic parameters, the procedure applied here was considered suitable.

When faced with the data of the open literature for CoMo and NiMo/Al_2O_3 catalyst, the estimated activation energies are similar to those obtained by Laurent and Delmon (Laurent and Delmon 1994) for the conversion of methylacetophenone (73 kJ/mol for a sulfided NiMo and 50 kJ/mol for a sulfided CoMo). When compared with the parameters obtained by Bunch et al (Bunch et al. 2007) for benzofuran, the estimated activation energies are also similar (66 kJ/mol for a reduced NiMo). The deviations between the parameters may be due to the solvent and mainly the molecules structure.

In Figure VI-2, the results from each compound used in the fitting are presented. As illustrated, furfural, furfuryl alcohol, tetrahydrofuran (THF) alcohol, furan, methylfuran and butanol were the effluents used for the estimation of the kinetic parameters. In a general perspective, the model can fit reasonably well the experimental point at lower temperatures (100 ºC and 150 ºC), while at 250 ºC the concentrations of most of the compounds are overestimated or underestimated. Similar to guaiacol, a 10 % deviation area has added to parity plots.

Table VI-6 – Estimated kinetic parameters for the HDT of furfural at 13 MPa over a reduced NiMo/γ-Al_2O_3 at a range of temperature between 100 °C to 250 °C and a range of reaction time between 0 h to 1 h.

Name	Reaction	k_0 (/s)	Ea (kJ/mol)
Carbonyl Hydrogenation (COH)		10 000	56
Decarbonylation (CO)		500	70
Hydrodeoxygenation (HDO)		25 000	75
Ring Opening (RO)	 	6 000	60
Saturation (SAT)		500	60

The kinetic model can predict within the error area the concentration for furfural at 100 ºC and 150 ºC, as shown in Figure VI-2 *a)*. For a higher temperature, 250 ºC, the model converts all furfural, which is far from the desired. This shortcoming may be associated to the kinetic parameters, however, and as seen in the Experimental results chapter, the range of products at 250 ºC is higher than at other temperatures. This reflects more compounds and, so a more complex reaction network and reactivity. Furthermore, the experimental tests used in the estimation of parameters are not homogenous, meaning that between the test at 150 ºC and the test at 250 ºC there is a gap of 100 ºC. Since furfural is an extremely thermosensitive molecule, this gap of temperature is sufficient to cause impact on the reactivity, as shown by the blank test at 100 ºC (report to Experimental results chapter). The 100 ºC gap is then reflected in the model, which is return cannot properly account with the concentration at 250 ºC, at the same time as considers the experimental information of the points at 100ºC and 150 ºC. Also, as for guaiacol, catalyst deactivation may play a significant role, especially at higher temperatures and longer residence times.

Furfuryl alcohol is the major products of furfural and therefore, the most important in the estimation. For this compound, the model predicts reasonably well the concentration at 100 ºC and 15 ºC. At 250 ºC, the model overestimates the concentration of furfuryl alcohol for almost five times more, as shown in

Figure VI-2 *b)*. This means that at 250 ºC, furfuryl alcohol should have already converted. According to Ozagac (Ozagac 2016) and the experimental concentrations at 250 ºC, furfural alcohol should be transformed into cyclopentanol by hydration. However, since this reaction is not accounted in the proposed reaction network, the model cannot correctly predict the concentration.

Similar to furfural alcohol, the model can predict the concentration of THF alcohol at 100 ºC and 15 ºC, but overestimate it at 250 ºC by a factor of around about five as well, as shown in Figure VI-2 *c)*. The lack of conversion of this product at high temperatures may be related to unsuitable kinetic parameters for the opening of the ring reaction or saturation. Regarding the latest, it is important to remark that the correlation here implemented is the same applied for the reaction network of guaiacol. This approximation may cause deviations in the results, in particular in at 250 ºC were the equilibrium constant is more affected. Furthermore, the absence of undesirable reactions, as hydration, can also be connected to the overestimation of the concentration of THF alcohol at 250 ºC.

Concerning furan, the model predicts extremely low values at 100 ºC and 150 ºC, which can be considered as zero (report to appendix B), as shown in Figure VI-2 *d)*Although the experimental values for both temperatures is also low, the model should be able to fit them as for the THF alcohol (same scale in the axis). For the concentration at 250 ºC, the model also predicts an extremely low value, even when the experimental concentration is higher than those at 100 ºC and 150 ºC. For the lower temperatures, unsuitable kinetic parameters for the opening of the ring, saturation and decarbonylation, together with an inadequate equilibrium constant may be the cause of the problem. However, and more probable at 250 ºC, undesirable reactions as hydration, diel-alder and aldol reactions may be the main cause for the unsuitable fit (Ozagac 2016) (J. P. Diebold 2000).

Regarding methylfuran, the same behavior seen in furan occurs at 100 ºC, as shown in Figure VI-2 *e)*. The model predicts extremely diluted concentrations which are practically zero. Once again, although the experimental values at 100 ºC also low, the model should be able to fit them as for the THF alcohol (same scale in the axis). At 250 ºC, the model overestimates the concentration of methylfuran. This may be originated from unsuitable parameters for the hydrodeoxygenation or the opening of the ring. However, and more probable at 250 ºC, undesirable reactions as hydration, diel-alder and aldol reactions may be the main cause for the unsuitable fit, as seen by Ozagac (Ozagac 2016). Be that as it may, the mode predicts reasonably well the experimental concentration at 150 ºC.

Finally, the model predicts all the experimental points of butanol within or at the margin of the 10 % error area, as shown in Figure VI-2 *f)*. Contrary to the other compounds, the model can predict approximately well the concentration at 250 ºC. One of the reasons may be related with the reactivity of butanol, which is smaller than the others compounds and thus, has less possible reactions, in particular those who are undesirable. Be that as it may, it is important to highlight the scale of the concentration. The values are extremely low, so even if there are deviations between the experimental and predicted

concentration (report to Table VI-4 and appendix B), these will also be small and hardly detectable in the figure.

In summary, the proposed kinetic model can predict the concentration of all compounds at 100 ºC and 150 ºC, but not at 250ºC, with exception of butanol. As mentioned, catalyst deactivation, unsuitable kinetic parameters, absence of undesirable reactions, such as hydration, or even an unsuitable equilibrium constant and respective approximation may lead to the differences presented in Figure VI-2. These deviations, in particular at 250 ºC, can also be related to the non-homogeneity of the experimental data. Nevertheless, and similar to guaiacol, the estimation of the kinetic parameters associated to the reaction network presented is not simple. As mentioned before, the models do not account with the impact of the catalytic sites and the molecule structure. For these reasons, the competition between the compounds for the catalytic sites is neglected, together with the different reactivities associated to the different molecules. As consequence, both reasons are intrinsically connected to the shortcomings in the estimation of the parameters and the fitting of the concentrations. Additionally, the proposed reaction network requires the estimation of ten kinetic parameters (two per reaction). However, there are only four experimental tests available for the hydrotreating of furfural. As a consequence, the kinetic parameters are associated with certain shortcomings, which could be eliminated with additional experimental tests.

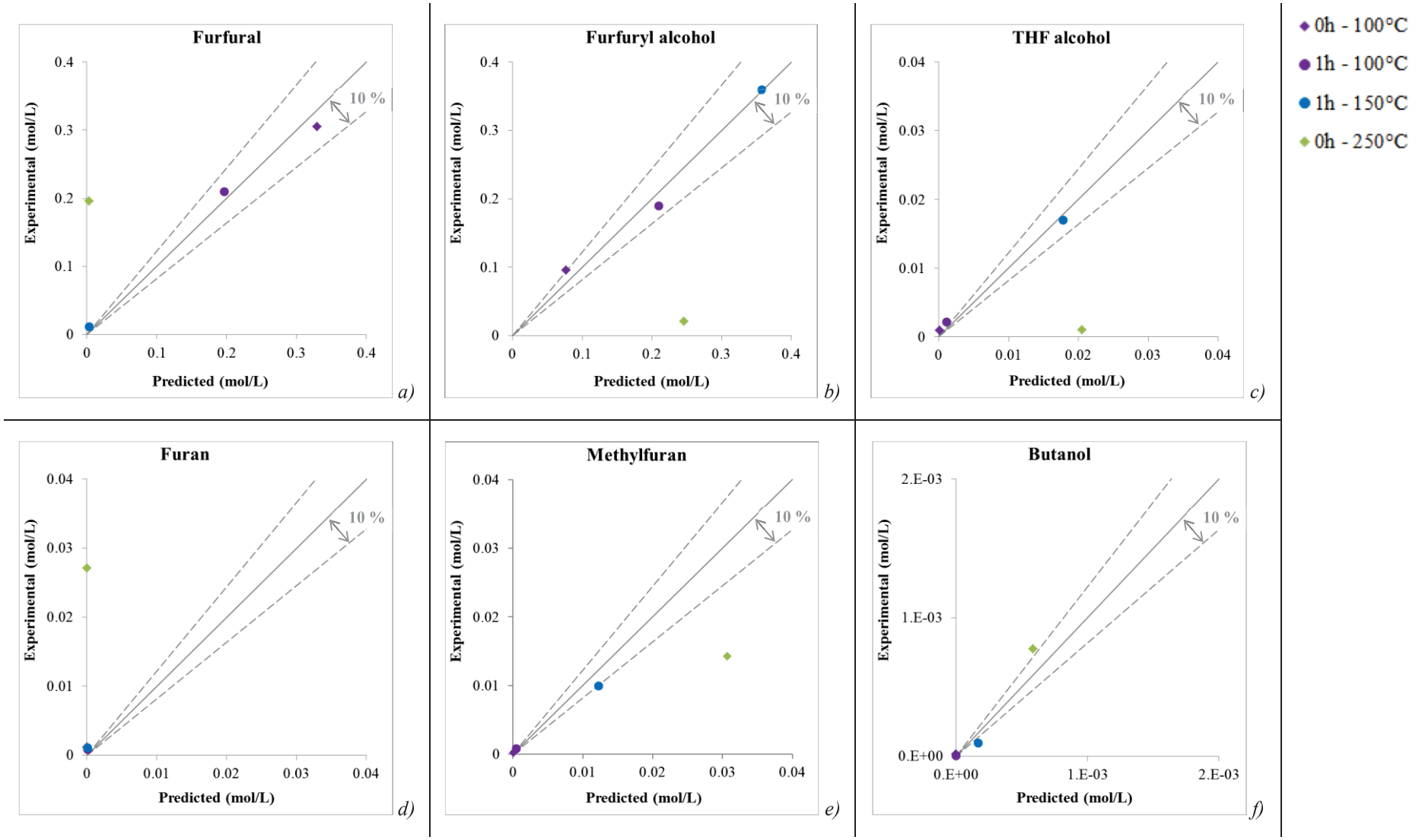


Figure VI-2 – Parity plots of the deterministic model of furfural HDT: a) Furfural, b) Furfuryl alcohol, c) THF alcohol, d) Furan, e) Methylfuran and f) Butanol.

VI.3 Conclusion of Chapter VI

In the present chapter, the results of the deterministic models were presented and discussed for both model compounds: guaiacol and furfural. These models were based on the experimental data available, the model structure previously created and the procedure of estimation of parameters.

For guaiacol, the model predicted reasonably well the concentration of certain compounds and experimental conditions, while others are overestimated or underestimated. The deviations found in the model may be linked to catalyst deactivation, unsuitable parameters, the absence of reactions and the disregard of the impact of the molecule structure and the catalytic sites in the reactivity. The kinetic parameters found for the proposed reaction network follow the order describe in Equation VI-3 and Equation VI-4 for the pre-exponential factor and the activation energy respectively.

$$k_{0,TRANS} > k_{0,DEM} > k_{0,DEMO} > k_{0,SAT} > k_{0,HDO} \qquad \textit{Equation VI-3}$$

$$E_{DEM} > E_{HDO} > E_{TRANS} = E_{DEMO} > E_{SAT} \qquad \textit{Equation VI-4}$$

For furfural, the model predicts reasonably well the concentration of all compounds for the lower temperatures (100 ºC and 150 ºC), but overestimated or underestimated at 250 ºC. As for guaiacol, the deviations found in the model may be justified by catalyst deactivation, unsuitable parameters, the absence of reactions and the disregard of the impact of the molecule structure and the catalytic sites in the reactivity. Furthermore, the lack of experimental data plays an important role in the estimation of parameters. The kinetic parameters found for the proposed reaction network follow the order describe in Equation VI-5 and Equation VI-6 for the pre-exponential factor and the activation energy respectively.

$$k_{0,HDO} > k_{0,COH} > k_{0,RO} > k_{0,CO} = k_{0,SAT} \qquad \textit{Equation VI-5}$$

$$E_{HDO} > E_{CO} > E_{RO} > E_{COH} = E_{SAT} \qquad \textit{Equation VI-6}$$

In conclusion, although both models present certain shortcomings, it is important to highlight the difficulty associated to the fitting of both. Even when faced with lack of experimental information and intrinsic information about the kinetic behavior of the compounds, both models were able to predict reasonably well most experimental tests and their trends within or at the limit of the 10 % error area.

Chapter VII. Stochastic modeling methodology

The SSA was the tool selected to reproduce the evolution of the hydrotreating of bio-oil model compounds and the hydrotreating of an industrial feed. In the following chapter, a brief description of the principle behind the stochastic simulation approach, its functionality, precision and limitations will be discussed. Furthermore, the chapter includes the extension of the SSA to the thermal behavior of the pilot unit, and the adaptation to bio-oil model compounds reactions and their subsequent stochastic simulation results. For more information about the SSA, the used model compounds and their reactivity, the reader is also referred to the following references: (Gillespie 1976), (Gillespie 1992), (Oliveira 2013) and (Ozagac 2016).

VII.1 Stochastic Simulation Algorithm

As discussed in the State of the Art chapter, a network of reactions can be described through a deterministic or a stochastic approach. In a complex molecular system, deterministic simulation is not the most suitable method to keep the molecular detail. Furthermore, this procedure implies a pre-definition of the reaction pathways, which in the case of complex mixtures is close to impossible.

On the contrary, the stochastic simulation is able to describe a reaction network molecule by molecule through a discrete evolution in time, while keeping track of the molecular detail. This feature gives the method the power of generating the reaction network while the simulation is unrolling, thereby avoiding the pre-definition of the reaction pathways.

VII.1.1 Theoretical aspects

As stated by the collision theory, for a reaction to occur particles must collide with the correct orientation and with sufficient energy. Such collision can be seen as a random event, as it requires that the molecules get close enough to collide and thus react. The stochastic approach is based on this idea, which suggests that each reaction has a certain probability of occurring.

As previous discussed, the stochastic approach can be described by the CME (Chemical Master Equation). This analytical equation (see Equation I-2) describes the evolution of time of a given chemical system through a single probability function. Briefly, this equation declares that for a given time, there is the probability of the chemical system being at a given state. For example, at beginning of a reactional process (initial time) there is no products, just reactants. However, in the following time, the chemical system may be represented by several reactants and one product, or several reactants and another type of product, or several reactants and more than one product, or even just reactants (no reaction happened). Be that as it may, this equation is extremely complex to solve, either analytically or numerically, for large chemical systems. For this reason, several authors (Gillespie 1976, 1992;

McDermott et al. 1990) proposed similar stochastic approach that can be solve through Monte Carlo methods (random sampling approaches).

One of these methods is the Stochastic Simulation Algorithm (SSA). Although work exactly in the same way, Gillespie (Gillespie 1976, 1992) proposed the description of the chemical system by a probability density function, $P(\Delta t, \mu)$, as shown in Equation VII-1.

$$P(\Delta t, \mu) = h_v \cdot c_v \cdot exp\left(-\sum_{\mu=1}^{M} h_v \cdot c_v \cdot \Delta t\right) \qquad \textit{Equation VII-1}$$

Where c_v is the stochastic for reaction v, h_v is the combinatorial factor that represents all the possible ways that the reactant molecules may collide and generate reaction v, Δt is the time interval and M is the total number of reactions for a given molecule.

The stochastic simulation of reactions consists in the integration of Equation VII-1 by a Monte Carlo approach. As such, there is the need to decompose Equation VII-1 through the formula of Bayes (Equation VII-2). In the end, the probability density function, $P(\Delta t, \mu)$, can be replaced by two other simple probabilities, as discussed by Gillespie (Gillespie 1976, 1992) and shown in Equation VII-3 and Equation VII-4. Where $P_1(\Delta t)$ is the probability for the next reaction to happen in a given interval of time (Δt) and $P_2(\mu|\Delta t)$ is the probability of reaction μ is the next reaction to happen, knowing that it will occur in the interval of time (Δt) previously defined.

$$P(\Delta t, \mu) = P_1(\Delta t) \cdot P_2(\mu|\Delta t) \qquad \textit{Equation VII-2}$$

$$P_1(\Delta t) = \sum_{\mu=1}^{M} h_v \cdot c_v \cdot exp\left(-\sum_{\mu=1}^{M} h_v \cdot c_v \cdot \Delta t\right) \qquad \textit{Equation VII-3}$$

$$P_2(\mu|\Delta t) = h_v \cdot c_v \Big/ \sum_{\mu=1}^{M} h_v \cdot c_v \qquad \textit{Equation VII-4}$$

In order to perform a Monte Carlo integration, both Δt and μ are calculated from Equation VII-3 and Equation VII-4, respectively, through the generation of two random numbers. The first random number (RN_1) allows the calculation of the interval of time (Δt) as shown in Equation VII-5, the invers of Equation VII-3. The second random number (RN_2) allows the sampling of the cumulative version of $P_2(\mu|\Delta t)$ in order to obtain the next reaction, as shown in Equation VII-6. $P_2(\mu - 1|\Delta t)$ is the probability of reaction $(\mu - 1)$ is the next reaction to happen, knowing that it will occur in the interval of time (Δt) previously defined.

$$\Delta t = -\ln(RN_1) \Big/ \sum_{\mu=1}^{M} h_v \cdot c_v \qquad \textit{Equation VII-5}$$

$$P_2(\mu - 1|\Delta t) < RN_2 < P_2(\mu|\Delta t) \qquad \textit{Equation VII-6}$$

VII.1.2 Methodology description

As shown in Figure V-1, the stochastic simulation algorithm requires as input the initial conditions of the process and simulation, the reactant molecules, reaction rules and their rate constants. The first parameters will shape the simulation and the kinetics of the reactions. The reaction rules will define the pathways through which the reactant molecules may undergo. Finally, the rate constants, adjusted to the process conditions, will establish the reactivities of each reaction and consequently of each molecule.

After having defined all the inputs, the process starts by calculating all the reactivities for all the molecules. For such, first the structures of the molecules are swept in search of the active zones. For each zone, one reaction can occur and, as consequence, a reactivity must be calculated. Afterwards, these reactivities are organized, and via the calculation of the probabilities, a PDF for each molecule is obtained. This PDF is a histogram where each column displays the probability of a given reaction. To correctly sample the reaction, the PDF is transformed in a cumulative function. Via a Monte Carlo method, a reaction is sampled for the PDF, as well the time that will elapse until this next event occurs. After updating the reaction system by specifying which molecule and reaction will happen, this process is repeated until the simulation arrives to its final time, which is specified by the user. It is important to highlight that this process will not only transform the reactants molecules, but also their products, if these can undergo any reaction.

To help observe the evaluation of the system, an extra temporal parameter (registration time) was added, which stops the simulation in factions of time and prints the system properties. For a better understanding of the SSA method, this section was divided into several steps corresponding to the order seen in Figure V-1.

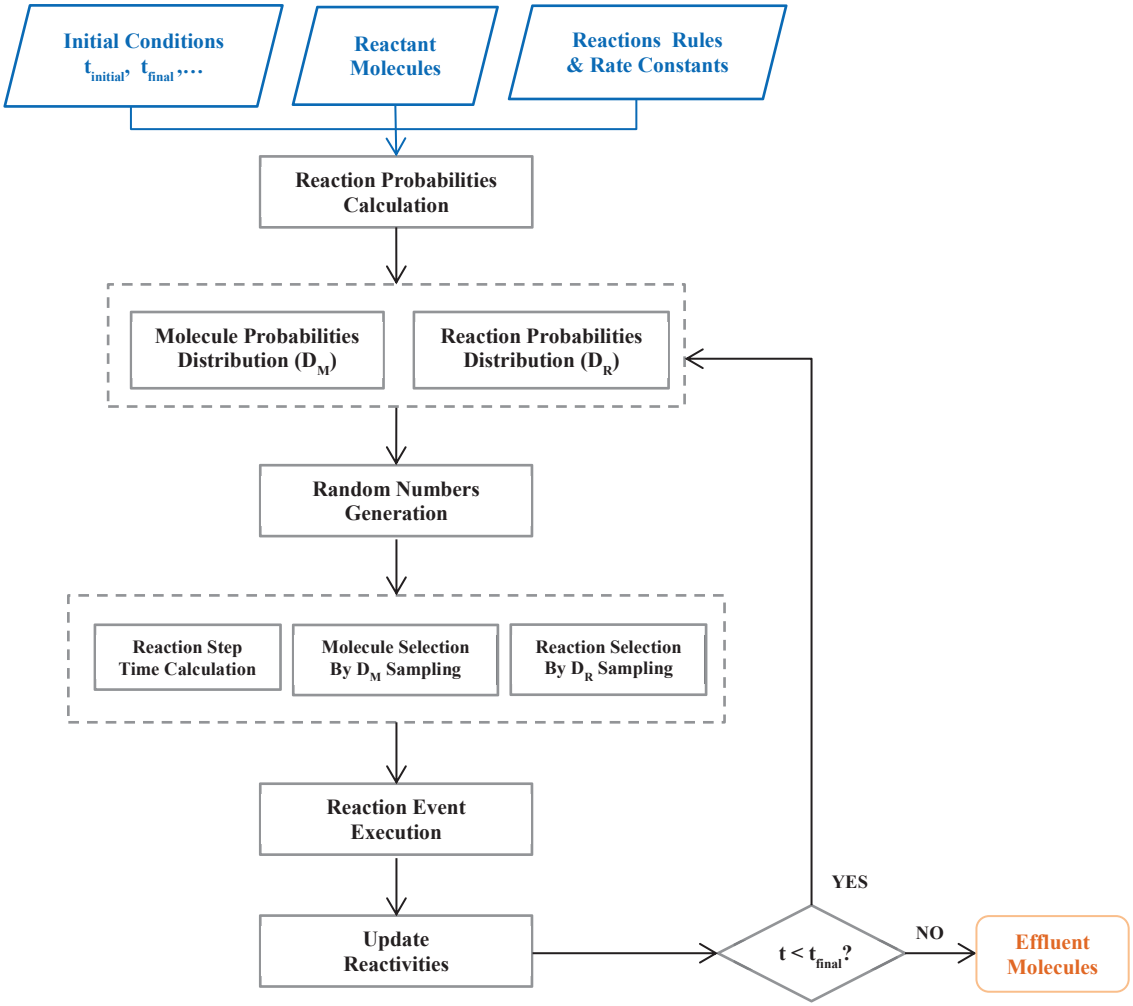


Figure VII-1 - Stochastic Simulation Algorithm (adapted from (Oliveira 2013)).

VII.1.2.1 Preliminary information

Before starting the simulation process, information regarding the simulation, the operating conditions, the reactions and their kinetic parameters are required. These are given by the user and may vary according to the mixture under study. The next sections will describe them with detail.

Process and simulation conditions

The inputs parameters related to the simulations are composed by the number of molecules, the time of the simulation, the number of simulations (iterations) and the output frequency. Obviously, the number of molecules represents the number of reactants included in the library, while the time of simulation is equivalent to the mean residence time of the hydrotreating process. The number of simulations represents the number of times that the SSA process will occur. As previously mentioned, this SSA algorithm is based on a Monte Carlo method, which is characterized by its random nature. Therefore, the results retrieved from one simulation are different from those extracted from another one. For this reason, and in order to obtain consistent results, the SSA procedure must be performed several times.

The output frequency represents the interval of time, where the algorithm will stop to calculate the mixture properties. This feature allows to observe the evolution of the process under question with time. This additional parameter depends on the number of simulations.

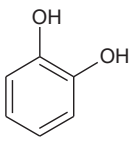
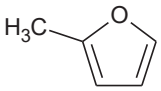
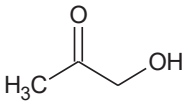
The process conditions are the operating conditions of the hydrotreating process, such as the temperature, the system pressure and the hydrogen partial pressure. These, like the simulation time, correspond to the ones used in industry or in experimental tests. All the remaining factors that may affect the process or the reactivity of the feedstock are not taken in consideration.

Reactant molecules

Regarding the reactants, a library of molecules is used to assure their representation. Obviously, this library depends on the mixture under study. In the present work, the SSA tool will not only be applied to the bio-oil complex mixture, but also to model molecules. In these cases, as there is only one molecule, a library of molecules identical to the model compound will be generated. For the case of bio-oil, the library will be directly retrieved from the molecular reconstruction method.

Considering the synthetic mixture proposed in Table VII-1, the library of reactants would be composed by catechol, 2-methylfuran and acetol. Since the composition of the model molecules is not the same, the library would be composed by 35 % of catechol, 50 % of 2-methylfuran and 15 % of acetol. As an example, for 10000 molecules, there are 3500, 5000 and 1500 molecules of catechol, 2-methylfuran and acetol, respectively.

Table VII-1 – Synthetic mixture for the illustration of SSA.

Molecule	**Catechol**	**2-Methylfuran**	**Acetol**
ID	CAT	2MF	ACE
Formula			
Fraction	0.35	0.50	0.15
Reactions	Saturation 1-Hydrodeoxygenation 2-Hydrodeoxygenation	Saturation Hydration 1-Open ring 2-Open ring	1-Hydrodeoxygenation Carbonyl Hydrogenation Hydration

Reactions rules and rate constants

In order to generate the reaction evolution of the system, the SSA method requires information concerning the reactions underwent by the molecules in the process under questions. Additionally, these reactions must come associated with their respective kinetic parameters. In the present work, the

reactions implemented in the algorithm will be linked to the hydrotreating process of organic oxygenated compounds. Although these reactions were briefly introduced in the State of the Art, a more detailed insight will be given further on the application of the model compounds and the bio-oils to the SSA.

Be that as it may, and considering the synthetic mixture proposed in Table VII-1, under hydrotreating conditions, the three compounds would potentially follow the reactions: saturation (SAT), hydrodeoxygenation (HDO), ring opening (RO), hydration (HYD) and carbonyl hydrogenation (HDT). Although only five reactions were accounted, more could happen because the molecules present more than one reactional site. For instance, catechol has two hydroxyl groups. Both represent a reactional site for HDO, meaning that catechol can undergo two different reactions depending on the site, as shown in Figure VII-2. Although in the end product is the same (phenol), both pathways are treated as independent and distinct. As for 2-methylfuran, also two reactions can be considered for the ring opening (RO), depending on which carbon-oxygen bond (site) will break. As shown in Figure VII-2, both reactions are once again treated as independent and distinct.

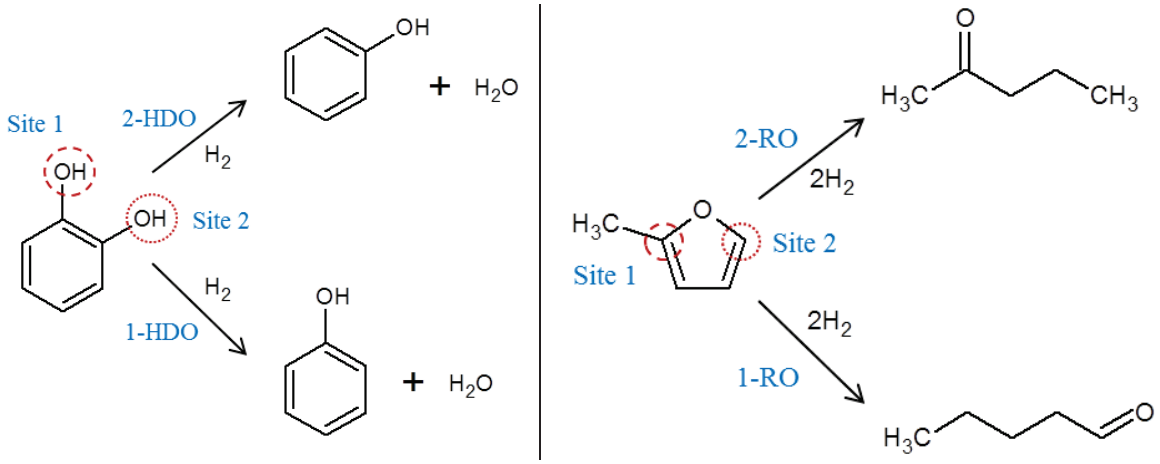


Figure VII-2 – Possible reaction sites for hydrodeoxygenation of catechol and ring opening of 2-methylfuran.

The identification of all possible reactions for a given molecule is obtained thanks to the reading of the molecule structure. As shown in Figure VII-2, each reaction is associated to a site, as functional group, a ring, a bond or even an atom. Every time a specific site is detected in a molecule, the algorithm accounts the reaction associated to the site as a possible pathway for the conversion of the molecule in question. As already seen, if a given site is repeated (e.g. the hydroxyl in the catechol), the same type of reaction is accounted twice, but as two independent and distinct pathways. In the end, the algorithm contains a list of reactant molecules with all their possible reactions, as illustrated by Table VII-1.

As previously mentioned, each reaction requires a set of kinetic parameters. To avoid the complexity of calculating the stochastic kinetic constant, Gillespie (Gillespie 1976) (Gillespie 1992) defined a relation between the stochastic and the determinist constants, as seen in Equation VII-7.

$$c_v = k_v/(N_a \cdot V)^{o_j - 1} \qquad \textit{Equation VII-7}$$

Where c_v and k_v are, respectively, the stochastic and the deterministic kinetic constant for reaction v, N_a is the Avogadro's Number, V is the volume of the system and o_j is the total order of the reaction.

In the present work, the chemical system under study is hydrotreating. In this process, the presence of hydrogen may result in monomolecular, bimolecular or even other types of reactions. However, as hydrogen is considered in excess, its impact on the chemical system is practically constant. For this reason, is possible to consider all reactions as monomolecular, meaning $o_j = 1$, thus the stochastic constant is equal to the deterministic constant. For this point on, the value given to the determinist constant will represent the stochastic constant of each reaction in the stochastic method. Such value can be obtained through the Arrhenius law, Quantitative Structure Reactivity Correlations (QSRC) or even other kinetic approaches.

VII.1.2.2 Reaction probabilities calculation

After identifying all the reactions and calculating their respective constants, the algorithm transforms this information into reactivities. As shown in Equation VII-8, the reactivity of a reaction (R_v) depends on the stochastic constant (c_v) and a combinatorial factor (h_v) that represents all the possible ways that the reactant molecules may collide and react.

$$R_v = h_v \cdot c_v \qquad \textit{Equation VII-8}$$

As previously discussed, in the present work, the stochastic constant can be defined by the deterministic constant. As for the combinatorial factor, its value can be calculated by Equation VII-9, where j is the reactive species of reaction v, N_v is the total of reactive species in the reaction v, n_j is the number of molecules of species j and p_j is the stoichiometric coefficient of species j.

$$h_v = \prod_{j}^{N_v} n_j!/(p_v! \cdot (n_j - p_j)!) \qquad \textit{Equation VII-9}$$

Considering that all the reactants have a reaction order of one and all reactions are monomolecular, there is only one possible combination. Therefore, in the end, the combinatorial factor equals one and thus the reactivities will only depend of their deterministic constant.

After knowing the values of the reactivities, the algorithm determinates the probability of each reaction occurring (P_v) via theses, as shown in Equation VII-10. These probabilities are then organized into probability distributions functions (PDFs).

$$P_v = R_v \Big/ \sum_{i}^{M} R_i \qquad \textit{Equation VII-10}$$

Where M is the total number of reactions for a given molecule, i is the reaction in question and R_i in the reactivity of reaction i.

Originally Gillespie (Gillespie 1976, 1992) proposed the application of just one PDF for all the methodology. However, in order to simplify the method and make it more user friendly, it was chosen to generate and use several PDFs for the development of the simulation. For this reason, the number of random numbers required will be three (interval of time, reaction AND reactant) instead of just two.

Reaction probabilities distribution

For a more practical illustration, let's assume the synthetic mixture in Table VII-1. Assuming a set of six molecules, catechol and acetol would represent only two and one respectively, while 2-methylfuran would be half of the set, meaning three molecules, as shown in Table VII-2. The same table shows the values considered for the reactivity of each molecule and reaction. In this example, it was assumed that the same type of reaction would have the same reactivity independently of the molecule under question. However, the reactivity of molecules can be more complex depending on several features, as for instance the molecule structure. For this reason, it is important to highlight that these values are merely used as an example and may not represent the reality.

Table VII-2 – Reactivity and probability of the reactions of the synthetic mixture.

Molecule	**Catechol**		**2-Methylfuran**			**Acetol**
ID	CAT-1	CAT-2	2MF-1	2MF-2	2-MF3	ACE
SAT	5	5	5	5	5	/
1-HDO	2	2	/	/	/	2
2-HDO	2	2	/	/	/	/
HDT	/	/	/	/	/	10
HYD	/	/	7	7	7	7
1-RO	/	/	9	9	9	/
2-RO	/	/	9	9	9	/

As can be seen, the probabilities of each reaction are then calculated via these reactivities. After cumulating the individual reactivities, these probabilities are organized into a reactions PDF, as shown in Figure I-34.

Figure VII-3 – Reactions probability distribution functions for the synthetic mixture.

Molecule probabilities distribution

Through the reactivities of each reaction it is likewise possible to define a molecule PDF. These will indicate the probability of a molecule undergoing a reaction. Equation VII-11 describes the probabilities calculation, where M is the total number of reactions for molecule j, $R_{i,j}$ in the reactivity of reaction i for molecule j and N is the total number of molecules in the mixture.

$$P_j = \sum_{i}^{M} R_{i,j} \Big/ \sum_{j}^{N} \sum_{i}^{M} R_{i,j} \qquad \textit{Equation VII-11}$$

In the end, these probabilities are once again organized into an accumulated PDF, but for the probability of molecules, as shown in Figure VII-4.

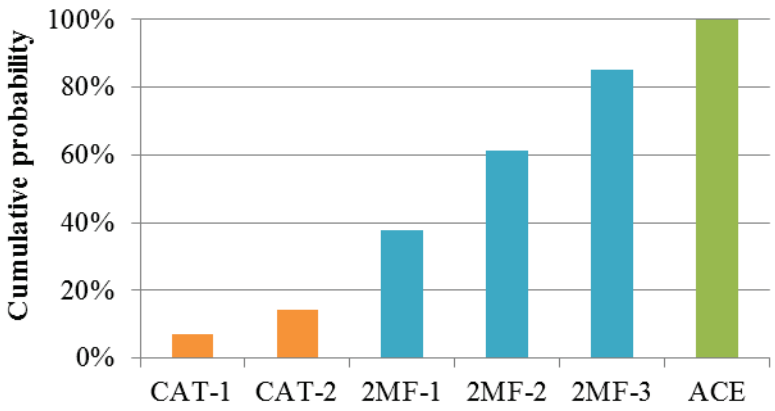


Figure VII-4 - Molecules probability distribution function for the synthetic mixture.

VII.1.2.3 Random selection of the time step, reactant and reaction

After the creation of all the accumulative PDFs, the algorithm generates three different random numbers, which vary from zero to one. The first one is to calculate the interval of time for the next reaction, the second one for the next reactant molecule and the third one for the next reaction which the chosen reactant will undergo. The next section will explain these steps in more detail with the help of the synthetic mixture.

Increment of time step

The increment of time (Δt) is a step in time from the present moment and the one where the next reaction will occur. This interval is calculated thanks to a random number (RN_1) and the sum of all the reactivities for all the molecules, as shown in Equation VII-12.

$$\Delta t = -\ln(RN_1) \Big/ \sum_{j}^{N} \sum_{i}^{M} R_{i,j} \qquad \textit{Equation VII-12}$$

The previous equation shows as well that, a part of having a stochastic nature, the interval of time varies also with the reactivity of the mixture. The higher this parameter, the smaller the interval of time, and the longer the simulation will be. However, when the number of reactions starts decreasing, due to the conversion of the possible reactants, the reactivity decreases as well and so the time interval for the next event increases, thereby shortening the simulation time.

Selection of the reactant

After the calculation of the increment of time, the algorithm resamples a second random number. This represents the accumulated probability to be read in the molecules PDF. In the end, a reactant molecule is selected as shown in Figure VII-5. In the case of the synthetic mixture, it was assumed a random

number of around about of 0.9 (90 %). When placed over the molecules PDF, the sampled reactant molecule is acetol.

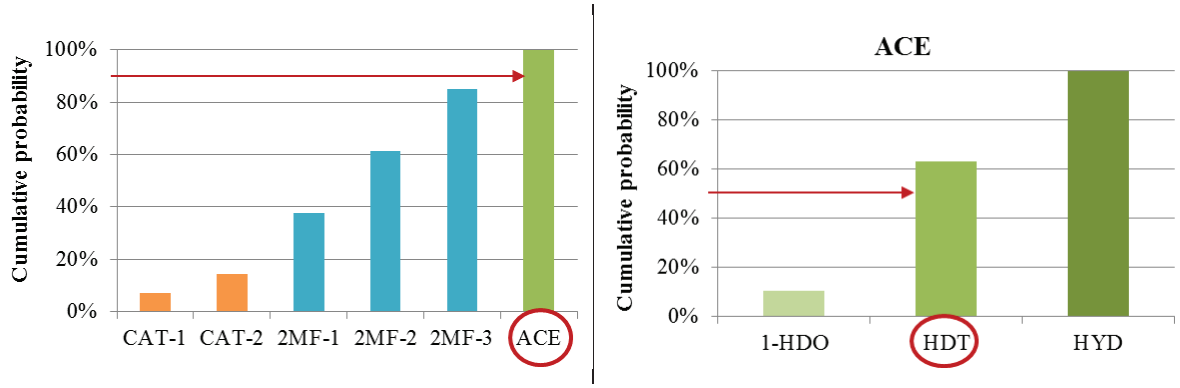
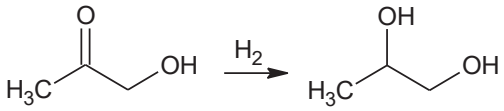


Figure VII-5 - Selection of the molecule and reaction.

Selection of the reaction

Similar to the previous PDF, the algorithm redraws a third random number. This value is the cumulative probability of the reaction PDF associated to the chosen molecule. Thanks to it, the next reaction to occur in the system is selected. In the case of the synthetic mixture, it was assumed a random number of around about of 0.5 (50 %). When read over the reaction PDF of the acetol, the selected reactant, this value results in a hydrogenation of the carbonyl group, as shown in Equation VII-13.



Equation VII-13

VII.1.2.4 Reaction event execution

Once the reactant, its reaction, and the time interval are selected, the molecular system must be updated. This process consists in the restructuration of the reactant molecule into the product molecule(s). This conversion is purely done by the modification of the molecule structure, through the elimination (addition) of atoms or substitutions or bonds, etc.

In the case of the synthetic mixture, the conversion of acetol into propylene glycol consists mainly in the elimination of the double bond and the replacement of an atom. More in detail, acetol is transformed by changing the carbon and the oxygen atoms associated to the carbonyl group into an aliphatic carbon atom and a hydroxyl. Although this step may appear simple, for more complex molecule structures and reactions (as ring opening), the reaction step is also more complicated.

VII.1.2.5 Update reactivities

Regarding the reactivities, this procedure is more complex, as it not only concerns the reactant, but also the product(s). Initially, the PDF associated to the reactant molecule must be eliminated, in addition to

the removal of this molecule from the PDF that contains the global reactivities. If the generated product can still react, a new PDF must be added according to its reactions. Furthermore, the product must be placed in the global reactivities PDF so it may be further on selected as a reactant. It is important to highlight that if more than one product is generated, the second molecule must be first added to the mixture and its reactivities must be likewise added to the PDFs library.

As shown in Figure VII-6, in the example mixture, the reactivities concerning acetol disappear and propylene glycol is added in the molecules PDF. For the same compound, a PDF for its reactions is also generated via their designated reactivities. For now, propylene glycol is also a molecule susceptible to be selected and transformed into a new product.

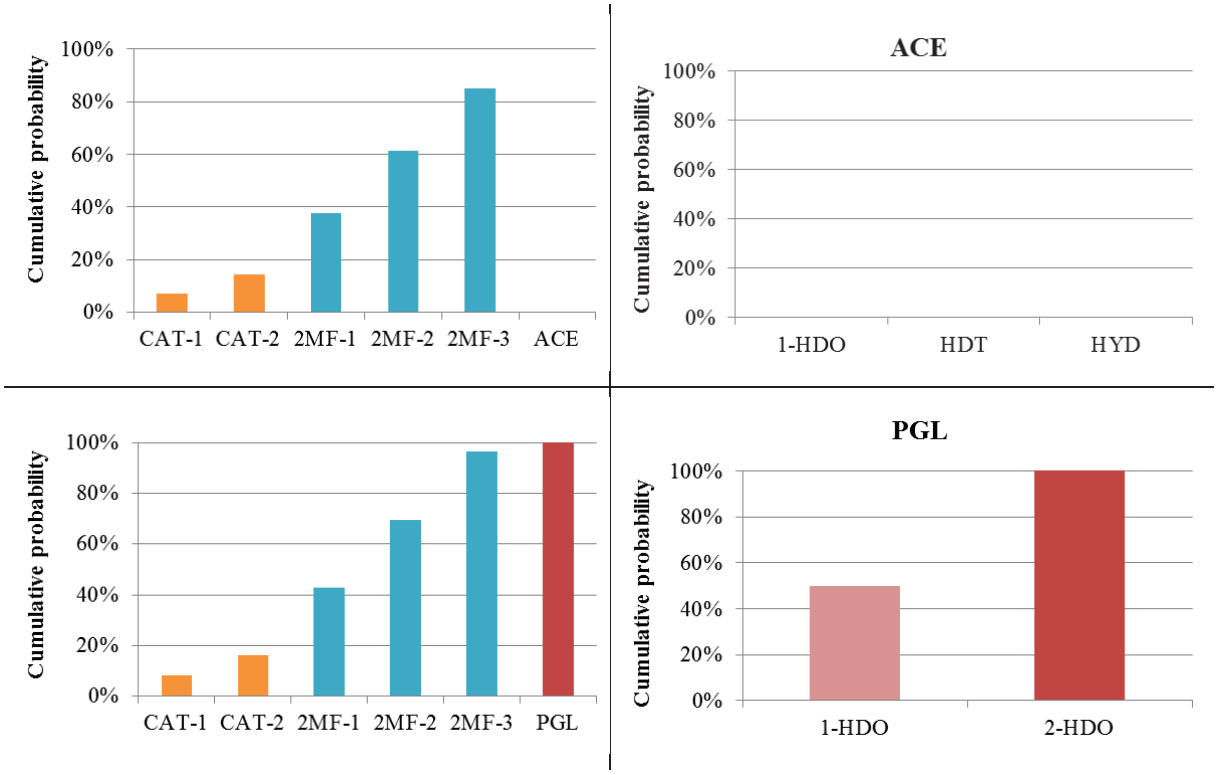


Figure VII-6 – Update the reactivities of the chemical system of the synthetic mixture.

VII.1.2.6 Evolution of the chemical system

While the time defined in the simulation does not reach the final time defined by the user, the process described previously will occur. In general, after fulfilling the reaction event and the update of the reactivities, three new random numbers will be redrawn for the calculation of the time interval and the selection of a next reactant and reaction. These steps are then followed by the reaction event execution and the reactivities update. Once again, the process will go back to the sampling procedure and continue to proceed with the following steps, until the time reaches the final time or there are no more reactants in the mixture (no more possible reactions). The flow sheet in Figure V-1 presents a summary of the continuation of this process.

To finish, in the SSA method there is another important feature which allows to follow the simulation behavior. Unlike the original SSA proposed by Gillespie (Gillespie 1976, 1992), the present algorithm can trace the evolution of the system through a series of time intervals (Δt_R) defined by a number of registration points (N_R) given by the user. These intervals are obtained thanks to Equation VII-14, where t_F is the final time of simulation given by the user.

$$\Delta t_R = t_F / N_R \qquad \textit{Equation VII-14}$$

In practice, if the simulation is defined for 3600 s and the number of registration points for 100, the interval of time will be 36 s. This means that every 36 s, the simulation stops, after the reactivities update step, and calculates and prints the mixture properties together with further information concerning the simulation of the process. The mixture properties are obtained from the molecule properties and compositions through linear rules. In their turn, the molecular properties are extracted from group contributions methods or correlations, while the composition of all molecules is equimolar.

VII.1.3 Algorithm precision

As previously seen, the SSA is a Monte Carlo method, which means that it is associated to the randomness of events. Also, as known to the world, randomness does not really imply precision, rising questions about the reliability of the algorithm and how much precise can the SSA actually be.

In methods based on the law of large numbers, the precision of the simulations (= results) is connected to the number of effectuated samplings. In the SSA, the number of samplings is directly associated to the number of simulated molecules and reactions. Furthermore, the number of simulations has impact on the number of molecules and reactions to be sampled, resulting in an indirect influence in the number of samplings and, consequently, in the algorithm precision. Therefore, the increase of the number of molecules, reactions and simulations will increase the precision of the SSA. However, since the number of reactions cannot be directly changed, only the other two features may be adapted in order to obtain a better the precision.

Be this as it may, the higher the number of molecules and simulations, the higher will be the calculation time needed to fulfill all simulations. For this reason, it is necessary to define a balance between the number of molecules and simulations and the calculation time, which allows the SSA to have enough precision to be considered a valid tool. As previously studied by Oliveira (Oliveira 2013), a set of 10000 molecules and 50 simulations, which represents 1 h and 15 min, is enough to achieve the defined objective. In the present thesis, those values were implemented.

VII.1.4 Algorithm computer limitations

Although an increase of the number of molecules may increase the precision of the SSA, this action may as well lead to the unexpected end of the simulations. As it happens, each molecule requires 85 kB to store its information, thus the increase of molecules will require more computer memory. If this is not available, the algorithm will not generate any result.

Even though the number of molecules may not be changed by the user, this value can still grow. In the SSA, there are certain reactions which generate more than one product, as cracking, dealkylation, demethoxylation, decarbonylation or decarboxylation. In just cases, the increase of the number of molecules is out of the user hand, forcing him to find other solutions.

As seen by Oliveira (Oliveira 2013), to overcome this problem the user may use a 64-bit system or simulate the initial feedstock by dividing it and individually simulate each new fraction. Oliveira (Oliveira 2013) proposed as well the modification of the molecule structure format. As mentioned, a molecule requires 85 kB, from where 79 kB represent the molecule structure. So, the replacement of this format with a more compact one, as SMILES and InCHI, could as well be a solution.

VII.2 Reactor thermal behavior

As stated previously in the State of the Art chapter, the stochastic simulation algorithm was already applied to oil fractions, as light cycle oils (Lopez Abelairas et al. 2016; Oliveira et al. 2012a), vacuum gas oils (Alvarez-Majmutov et al. 2016; Alvarez-Majmutov and Chen 2017), vacuum residues (Oliveira et al. 2013a; Oliveira et al. 2013b, 2014), and biomass fractions, as lignin pyrolysis. For all the cases, the SSA simulates an isothermal process. On the contrary of these examples, the compounds tested in the present work are extremely sensible to the temperature, as seen in The Experimental results chapter. For this reason, and in order to equate the experimental and simulation results, the SSA was extended to the thermal behavior of the pilot unit.

In an isothermal state, the algorithm received as input an initial and final time which match at the beginning and the end of the reaction. In the dynamic case, the initial time will now represent the beginning of the experimental test, while the final time, will be the end of the experimental test. This is represented by the extension of the dark blue zone in Figure VII-7 to the light blue zone.

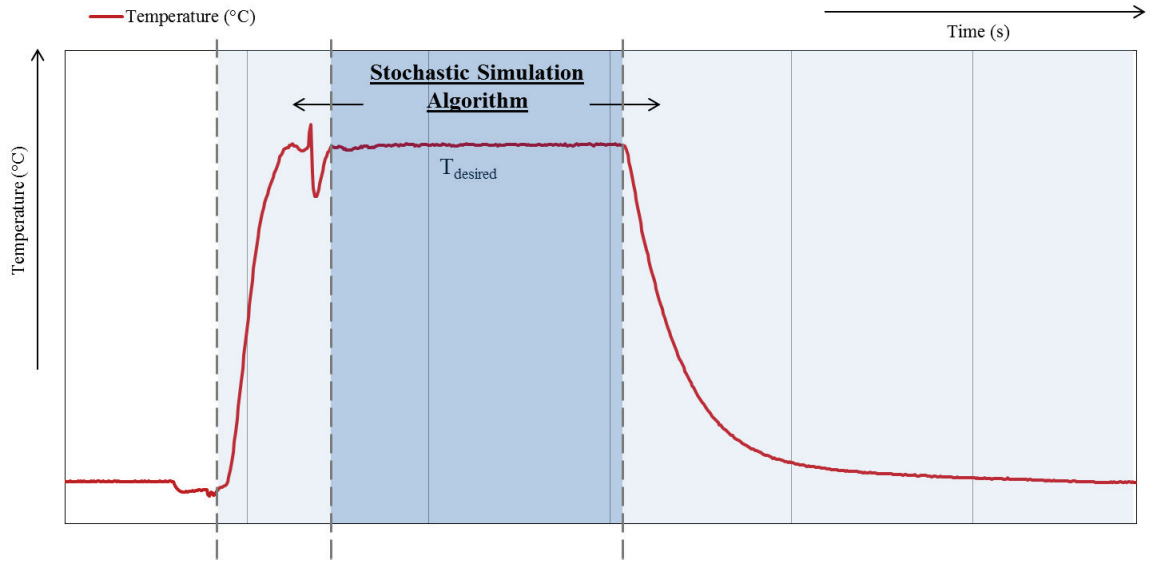


Figure VII-7 - Representation of the extension of the stochastic simulation algorithm to the temperature behavior.

Furthermore, in the dynamic state, the algorithm also receives as input the temperature profile collected from the computer associated to the pilot unit. Through the integration of this profile, shown by Equation VII-15, the algorithm can simulate the real behavior of the temperature during an experimental test.

$$T^{sto}_{t+\Delta t^{sto}} = T^{sto}_{t} + \left(\frac{dT}{dt}\right)^{exp} \cdot \Delta t^{sto} \qquad \textit{Equation VII-15}$$

Where $T^{sto}_{t+\Delta t^{sto}}$ and T^{sto}_{t} are the temperature at interval $t + \Delta t^{sto}$ and t respectively, Δt^{sto} is the interval of time defined in the SSA and $(dT/dt)^{exp}$ is the variation of the temperature with time in the profile extracted from the pilot unit. This parameter is given by Equation VII-16, where $T^{exp}_{t+\Delta t^{exp}}$ and T^{exp}_{t} are the temperature at interval $t + \Delta t^{exp}$ and t, respectively. In this expression, Δt^{exp} is the interval of time defined in the temperature profile from the pilot unit and is always fifteen seconds.

$$\left(\frac{dT}{dt}\right)^{exp} = \frac{\left(T^{exp}_{t+\Delta t^{exp}} - T^{exp}_{t}\right)}{\Delta t^{exp}} \qquad \textit{Equation VII-16}$$

As clearly shown Equation VII-15 and Equation VII-16 correspond to the stochastic and the experimental profiles respectively. Therefore, the time (t) in the first corresponds to the time defined by the SSA and in the second, to the actual experimental time.

The extension of the algorithm consists in the inclusion of the temperature behavior during an experimental test into the gear of the stochastic simulator, as shown in Figure VII-8. In the end, the temperature varies with the evolution of the time defined by the algorithm. Alongside this variation, the temperature dependent reactivities are recalculated, as are their probabilities and PDFs.

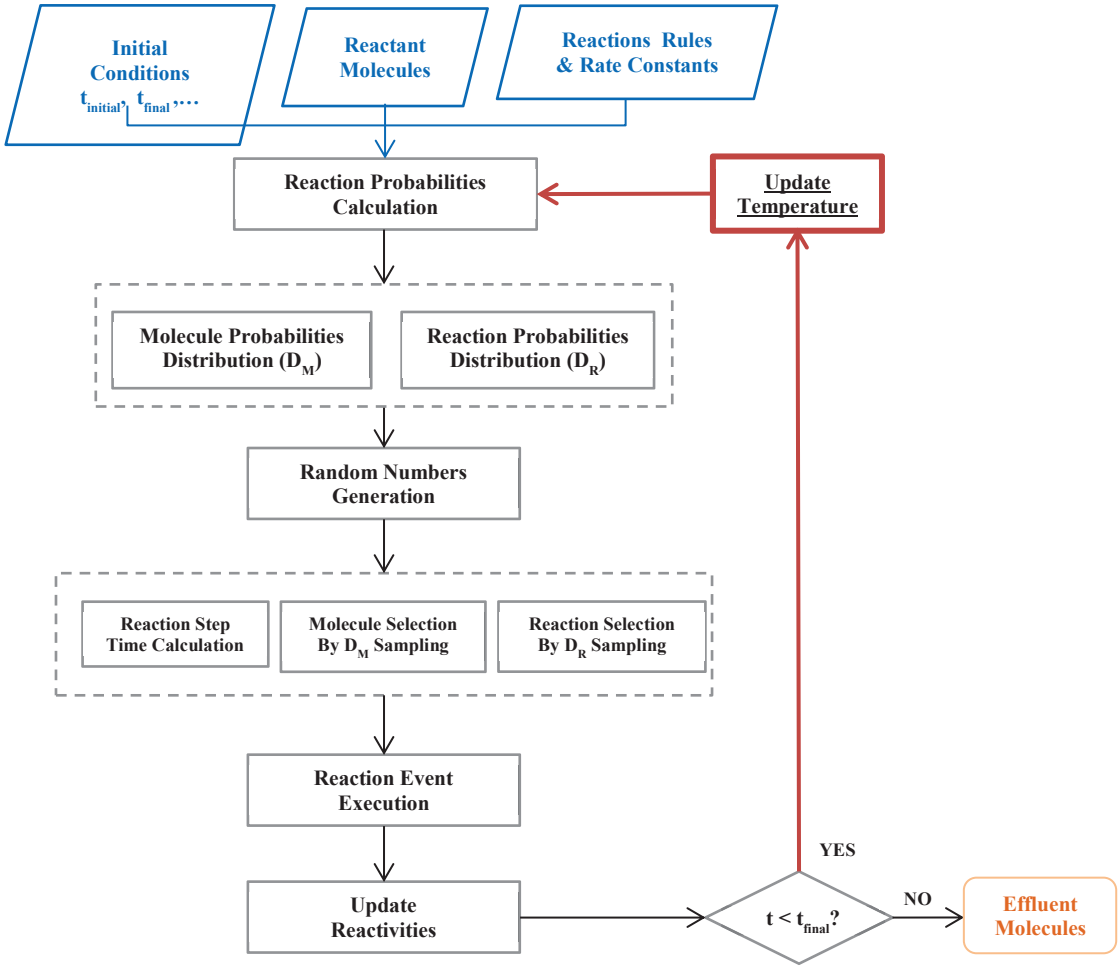


Figure VII-8 - Stochastic Simulation Algorithm with the temperature behavior.

The new feature changes the comparison process between the experimental and simulated results. Before, in the stated isothermal examples, to compare results at different reaction times, it was enough to perform only one simulation which would englobe all the desired times. For example, to compare the experimental and predicted conversion of a given reactant at 1 h, 3 h and 5 h, a simulation of 5 h would be sufficient, since it would cover all the desired times. However, in the present work such characteristic cannot be applied, as the heating and the cooling of the reactor must be always accounted for in each simulation. Therefore, when the thermal behavior is considered, one experimental test requires one simulation, which is specific for the test conditions. These procedures increase the time required to generate a kinetic model with the SSA.

Figure VII-9 shows the prediction of the temperature by the simulator. The four lines represent the four different temperatures used in the present work. As can also be seen, the integration of the temperature enables the correct prediction of the experimental profile. This method can even account with the stabilization of the temperature at the beginning of the reaction. It is important to note that the profiles do not start at the same time, because the extraction of the experimental profile from the pilot was not done for the same initial period of time.

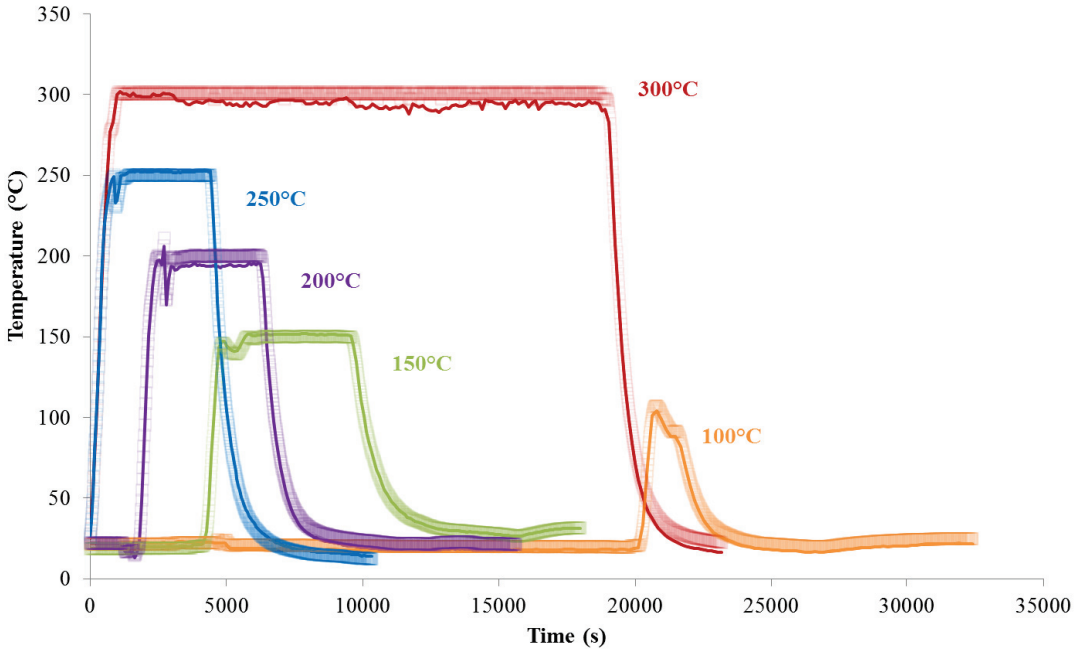


Figure VII-9 – Representative scheme of the temperature behavior at different temperatures: Squares - experimental data, Full lines – stochastic simulator.

Be that as it may, Figure VII-9 also shows certain differences between the experimental and predicted. As it happens, the integration of the temperature, besides depending on the experimental profile, depends on the stochastic time. In its turn, this parameter not only is randomly generated, but also depends on the mixture reactivity. This last evolves with the conversion of reactants and generation of new reactive products. Therefore, the prediction depends on the experimental profile, the stochastic nature and the reactivity of the molecules in the mixture.

In the end, the predicted temperature may present certain deviations, due to several reasons. If the experimental profile has precise fluctuations, the algorithm will be forced to satisfy them. As seen in in Figure VII-9, there are slight deviations in the experimental profile from the desired value of 300 °C (visible when zooming into that region). Because the effect of temperature is very sensitive and experimentally dependent, predicts these oscillations. Afterwards, the integration will continue. However, if several inconsistencies are located in the same region, as in profile at 300 °C, the simulator will always be trying to predict them and starts accumulating some deviations. Eventually this will disappear, when the profile return to a stable zone. The remaining profiles do not have precise oscillations, only coherent fluctuations as the kick of the controller. Therefore, the predicted profiles of these temperatures are coherent with the experimental data. In summary, the user should be careful when obtaining the experimental data.

The stochastic nature of the method is also an important parameter in the temperature prediction. As mentioned before, the stochastic simulation algorithm never reproduces the same results. Therefore, the

prediction of the temperature varies from iteration to iteration, as shown in Figure VII-10. Although always close to the experimental data, the maximum deviation is around about 8 %. A solution for this situation could be the creation of error limits. For example, the predicted temperature should be within a 2 % error range (measure error of the thermocouple) from the experimental temperature. However, the stochastic time usually does not fit the deterministic time from the experimental profile. The same can be seen for the temperature, meaning that the predicted temperature is usually within a range of experimental temperature, complicating the application of the proposed solution.

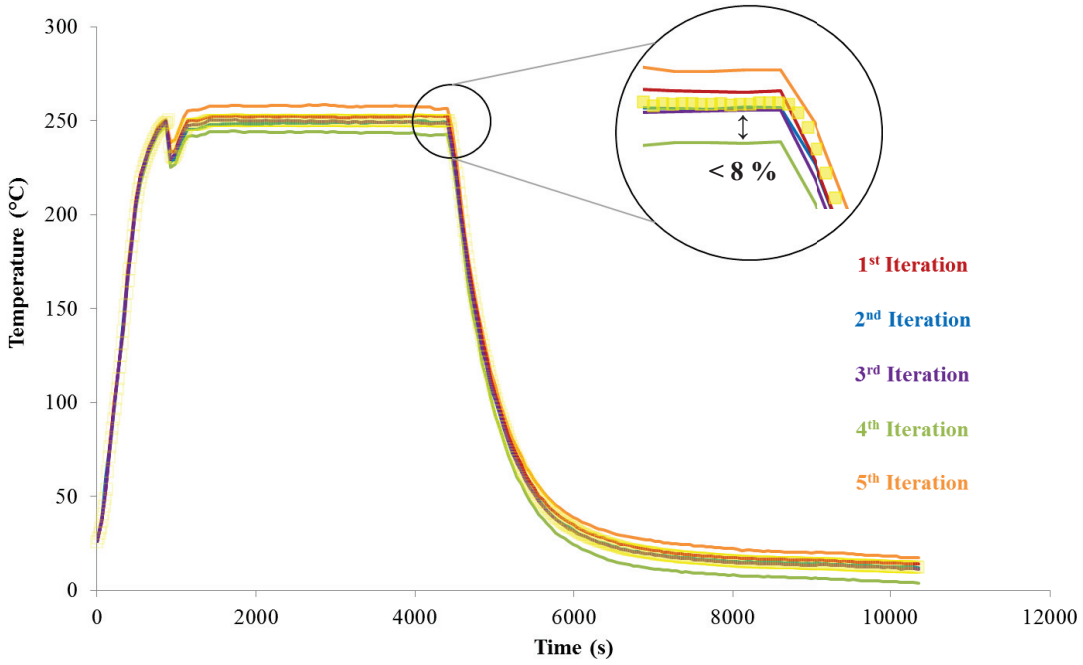


Figure VII-10 – Representative scheme of the temperature behavior after five independent iterations with the same inputs: Yellow squares- experimental data, Full lines – stochastic simulator.

Unlike the experimental profile and the stochastic nature, the mixture reactivity does not damage the predicted temperature profile. As mentioned before, the reactivity inversely affects the interval of time (see Equation VII-12), and as consequent the time. If the reactivity of the mixture is too small, the interval of time will tend to big values. These will increase too much the time, which when used for the integration of the temperature will create an inconsistent prediction of the temperature, as the experimental profile is badly read and as consequence applied. However, when the reactivity of the mixture reaches small values and the interval of time is too big, the algorithm will force to match the deterministic step of the experimental profile: 15 s. For this reason, the mixture reactivity will never damage the temperature prediction. In the most extreme perspective, the mixture reactivity will just force the algorithm to predict a profile exactly the same the experimental data.

VII.3 Adaptation to Model Compounds

As presented in Objectives and strategy chapter, the present thesis aims to understand the reactivity of bio-oils under hydrotreating conditions. However, the adaptation of the SSA tool and a reactivity study to complex matrices is not easy task. For both reasons, the present work began with single model compounds, which were typical of bio-oils and then evolved to bio-oil.

Before continuing, and as a remark, the SSA is a method for the simulation of complex matrices under a given process. Therefore, the application of this tool to single model compounds (not complex matrices) may seem useless and unnecessary. However, because the aim is to slowly evolve from a known simple feed to a more complex and unknown feed, it was decided that the present work would firstly apply the methodology to single molecules and then to bio-oil, allowing an easy adaptation of the SSA and a better understanding of the reactivities at stake.

The next sections will present the adaptation of the SSA to two model compounds of bio-oil: Guaiacol and Furfural. For both, the algorithm will be extended to account with the reaction rules associated to the hydrotreating process of both compounds. Furthermore, a brief summary of the kinetics applied to the methodology, and extracted from Deterministic results chapter, will be presented together with the results and discussion concerning them.

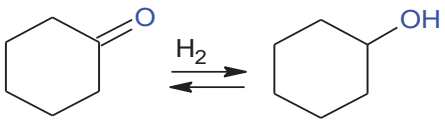
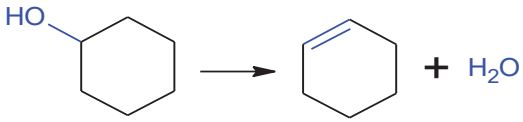
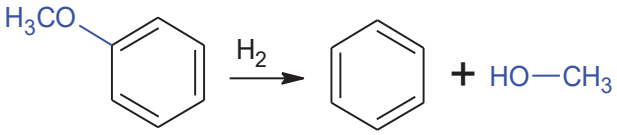
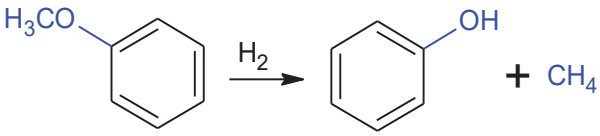
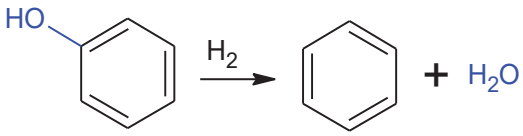
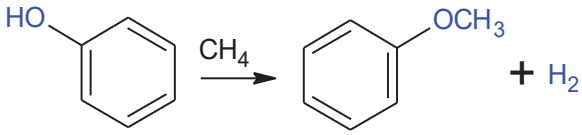
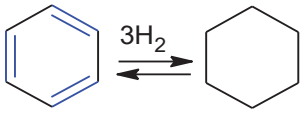
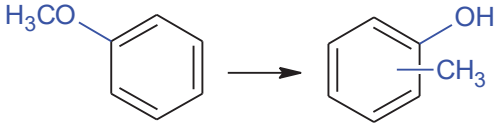
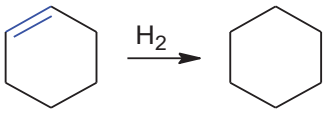
VII.3.1 Guaiacol hydrotreating

Guaiacol is a molecule typically found in bio-oil mixtures. This compound is composed by a benzene ring connected to a hydroxyl and a methoxy group in two different but neighbor carbons. Such structure clearly indicates that guaiacol is mostly generated from the degradation of lignin in the fast pyrolysis process.

VII.3.1.1 Reaction network

As seen in the State of the Art chapter, under hydrotreating conditions and over a NiMo/γ-Al_2O_3 catalyst, guaiacol undergoes mainly demethoxylation (DEMO), demethylation (DEM) and transalkylation (TRANS). The generated products may then follow hydrodeoxygenation (HDO) and saturation (SAT). Apart for these, other undesirable reactions may occur as the equilibrium hydroxyl/carbonyl (COH), dehydration (DHYD) and methylation (MET) of hydroxyl groups. All these reactions are presented in Table I-18. In total, there are ten reactions associated to the hydrotreating of guaiacol and its derivatives. At this point, it is important to highlight that, although guaiacol may form coke through macromolecules reactions, none of these pathways were considered in the SSA.

Table VII-3 – Reaction rules of guaiacol under hydrotreating.

Name	Site	Reaction
Carbonyl Hydrogenation (COH)	Cycle-C=O	
Dehydration (DHYD)	OH	
Demethoxylation (DEMO)	OCH_3	
Demethylation (DEM)	OCH_3	
Hydrodeoxygenation (HDO)	OH	
Methylation (MET)	OCH_3	
Saturation (SAT)	Cycles	
Transalkylation (TRANS)	OCH_3	
Unsaturated Hydrogenation (UHDT)	C=C, C≡C	

To simulate the hydrotreating of guaiacol in the SSA, it was necessary to add the ten reactions to algorithm. In fact, Oliveira (Oliveira 2013) already implemented the saturation reaction (equilibrium saturation/desaturation) for the case of oil fractions, leaving only nine reactions. These reactions were adapted to the SSA in the present thesis.

In the case of the saturation reaction, all aromatic carbons will be saturated with a hydrogen. Now without double bonds, the cycle is considered a saturated ring. If the inverse is performed, meaning if all carbons in the ring lose a hydrogen, the cycle becomes once again an aromatic.

For the hydroxyl/carbonyl equilibrium reaction a similar scheme is considered at the level of the carbon-oxygen bond. When the hydroxyl is oxidized, the oxygen loses the hydrogen and creates another bond with the aliphatic carbon to which the atom was already connected. This type of reaction generates a carbonyl attached to the ring, meaning that the ring initially cannot be aromatic. The inverse reaction consists in the saturation of the carbon and the oxygen with hydrogen, thus breaking the double bond between both.

Regarding the dehydration, the compound under question loses a molecule of water. In detail, the hydrogenolysis between the oxygen and the carbon atoms generates a water molecule with the help of the one hydrogen from a neighbor carbon. In the end, the molecule is stabilized by creating a second bond between the previous substituted carbon and the neighbors which donated the hydrogen. Although in theory such type of reactions may appear, according to Elliot (Elliott 2007) olefins converse extremely fast, meaning that the presence of such type of composed is hardly detected. This conversion is given by the saturation of both olefin carbons by hydrogen as shown in the unsaturated hydrogenation (UHDT) in Table I-18.

In the case of demethoxylation, demethylation and transalkylation, most of the molecule transformations occur at the level of the methoxy group, meaning that these three reactions are mostly specific to the methoxy family in the bio-oil. For this reason, in the algorithm, demethoxylation, demethylation and transalkylation can only happen to methoxy groups connected to a carbon ring. In a higher detail, demethoxylation consists in the hydrogenolysis between oxygen and the carbon in the ring. This reaction generates two products, where one is always methanol. Regarding demethylation, the hydrogenolysis occurs at the level of the aliphatic carbon in the methoxy group and the oxygen in the same group. Similar to demethoxylation, this reaction generates two products, where one is always methane. Finally, transalkylation is exactly the same process as demethylation. However instead of generating the main product and methane, this reaction implements the methyl group into the ring again, after breaking the oxygen-carbon bond in the methoxy group. Another reaction associated to demethylation is methylation. Although it is the inverse of the former, both reactions are not reversible. In the present work, it was considered that methylation would only occur at the level of the hydroxyl group. Therefore, through the addition of a methane to the hydroxyl group, this reaction can generate a methoxy group. It is important to highlight that a change in the concept of the methane molecule in the source code, transformed this previous monomolecular reaction into a bimolecular reaction. For this reason, methylation requires a brief preadaptation, alongside the SSA, before bimolecular reactions can be applied in the simulations.

To finish, guaiacol and its derivatives may undergo hydrodeoxygenation. This reaction consists in the breaking of the hydroxyl-carbon bond by hydrogen replacement. This reaction, either in reality or in the algorithm, can occur to any molecule with a hydroxyl group. In general, hydrodeoxygenation is the ideal reaction for the hydrotreating of bio-oils. Even producing large amounts of water, which is undesirable, hydrodeoxygenation just involves the removal of oxygen without breaking any further bonds or losing any carbon atoms.

In the guaiacol reaction network, the presence of gas reactions is also probable. These reactions, already identified in the State of the Art chapter are summarized in Table VII-4. As the previous reactions, all these pathways were implemented in the algorithm, making a total of sixteen reactions.

Table VII-4 – Gas reactions that may happen during hydrotreating.

Name	Site	Reaction
Methanol Decomposition (METD)	$HOCH_3$	$HOCH_3 \rightarrow CO + 2H_2$
Methanol Steam Reforming (METSR)	$HOCH_3$	$HOCH_3 + H_2O \rightarrow CO_2 + 3H_2$
Carbon dioxide Methanation (CDM)	CO_2/CH_4	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$
Carbon monoxide Methanation (CMM)	CO/CH_4	$CO + 3H_2 \leftrightarrow CH_4 + H_2O$
Water Gas Swift (WGS)	CO/ CO_2	$CO + H_2O \leftrightarrow CO_2 + H_2$

As shown in Table VII-4, methanol generated by demethoxylation can be degraded into carbon monoxide and hydrogen. Furthermore, with the increase of water in the mixture, methanol can also be transformed into carbon dioxide and hydrogen. None of these reactions are desirable, since there is the loss of carbon atoms, which means a reduction in the bio-oil energetic value.

As for carbon monoxide and dioxide, both can undergo hydrogenation into methane and water. These reactions are reversible. Besides methanation, carbon monoxide and dioxide are reversible as well thanks to the reaction water gas swift. As the methanol reaction, gas forming reaction is not desirable, since it implies the loss of carbons in the liquid effluent.

VII.3.2 Furfural hydrotreating

Furfural is also a molecule typically found in bio-oil mixtures. This compound is composed by a furanic ring connected to an aldehyde group in the carbon neighbor to the oxygen. Such structure suggests that furfural is mostly generated in the fast pyrolysis process from the degradation of hemicellulose, but it may appear as well from the degradation of cellulose derivatives.

VII.3.2.1 Reaction network

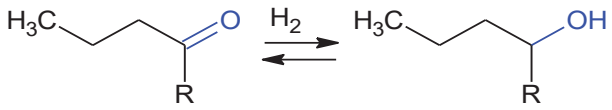
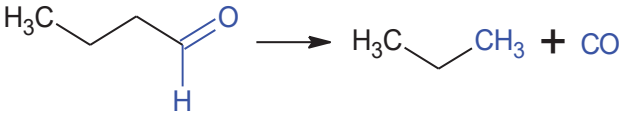
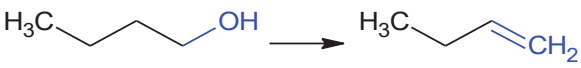
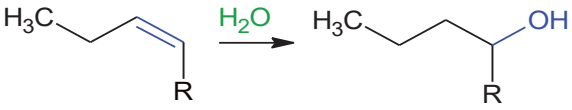
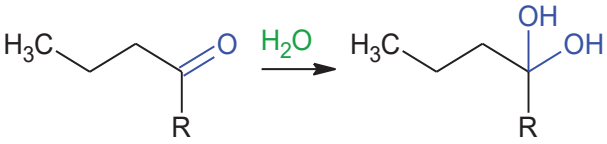
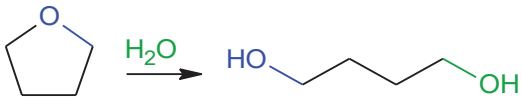
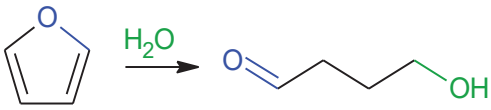
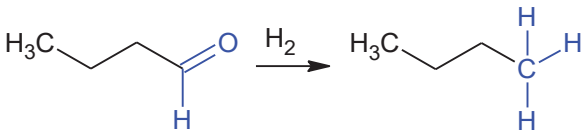
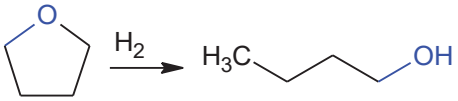
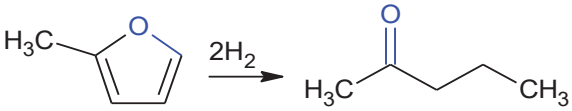
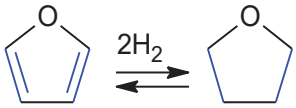
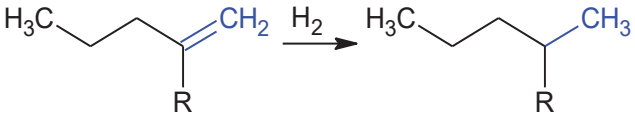
As seen in the State of the Art chapter, under hydrotreating conditions and over a $NiMo/Al_2O_3$ catalyst, furfural undergoes mainly decarbonylation (CO) and carbonyl hydrogenation (COH). The generated products may then follow hydrodeoxygenation (HDO) direct hydrogenolysis (DH), saturation (SAT) and ring opening (RO). Apart from these reactions, other undesirable reactions may occur as the dehydration (DHYD), hydration (HYD) and hydroxyl dehydrogenation, which is reversible (carbonyl hydrogenation (CO HDT)). All these reactions are presented in Table VII-5. In total, there are fourteen reactions associated to the hydrotreating of furfural and its derivatives. Similar to the reactional network of guaiacol, any reaction representative of the formation of macromolecules or polymerization was implemented in the code.

To simulate the hydrotreating of furfural in the SSA, it was necessary to add the fourteen reactions to algorithm. In fact, discounting the reactions already implemented by Oliveira (Oliveira 2013) and in the hydrotreating of guaiacol, only nine were newly added.

Concerning reactions as dehydration, hydrodeoxygenation and unsaturated hydrogenation, the same behavior of guaiacol reaction network can be expected for furfural. Regarding the equilibrium reactions, saturation is also quite similar to the reaction implemented for guaiacol, with the only difference beginning the type of ring. For the carbonyl hydrogenation, the theory behind the reaction is the same as in guaiacol: the carbonyl can be hydrogenated into a hydroxyl, while the hydroxyl can also be dehydrogenated into a carbonyl. The only difference resides in the nature of the carbon connected to the oxygen.

In the case of decarbonylation, only aldehydes can react. In fact, this reaction consists in the removal of the entire aldehyde group from the molecule. By losing the hydrogen characteristic from the aldehyde, this functional group is transformed into carbon monoxide. Similar to demethoxylation, demethylation and methanol decomposition and steam reforming, this reaction removes a carbon from the molecule. As previous discussed, such type of reactions are undesirable for the process, as the energetic value of the liquid effluent decreases.

Table VII-5 – Reaction rules of furfural under hydrotreating.

Name	Site	Reaction
Carbonyl Hydrogenation (COH)	C=O; HC=O	
Decarbonylation (CO)	HC=O	
Dehydration (DHYD)	OH	
Hydration (HYD)	C=C	
	C=O, HC=O	
	Cycle-O	
	Cycle=O	
Hydrodeoxygenation (HDO)	OH	
Hydrogenolysis (DH)	HC=O	
Ring Opening (RO)	Cycle-O	
	Cycle=O	
Saturation (SAT)	Cycles	
Unsaturated Hydrogenation (UHDT)	C=C, C≡C	

Regarding the hydration, several types of molecules (= sites) may undergo this reaction. Although hydration targets mostly double bonds between carbons or carbon-oxygen, it can also occur in the ether bridge of furans and tetrahydrofurans. In the case of chains, a water molecule connects, in the form of hydroxyl, to the only carbon of the double bond, for carbonyls, or to the carbon which is the most substituted (more stable carbocation) for carbon-carbon bonds. During the process, the extra hydrogen of the water is transfer to the unstable atom, which is either the oxygen or the other carbon of the double bond. In the case of cycles, the water will connect, in the form of hydroxyl, to the most substituted carbon neighbor of the oxygen. This will promote the opening of the ring into a dialcohol, for the tetrahydrofurans, or an aldehyde molecule with an alcohol functional group, for the furans.

Concerning the ring open reaction, both furans and tetrahydrofurans can undergo such transformation. These rings may be alone in the molecule core or connected to other rings. For all cases, the oxygen-carbon connection with undergo hydrogenolysis. For both the oxygen and the carbon, a hydrogen is added, generating either an alcohol for tetrahydrofurans, or a carbonyl for furans. Unlike hydration, ring opening reaction may occur in both carbon neighbors independently of their degree of substitution. In the algorithm, both hypotheses are considered, meaning that ring opening reaction always generates two possible reactions that are later on redrawn as potential reaction pathways. Such flexibility allows the method to generate more diverse products as seen in the reality. For this reaction, it was likewise assumed that saturated cycles with bisubstituted, olefinic and carbonyl carbons would not open. In the case of the bisubstituted carbons, the number of rings with two sidechains in the same carbon is not common in bio-oil (Staš et al. 2014).

Regarding the gas reaction, only carbon monoxide and dioxide methanation, plus water gas sift can be expected in the furfural network. For these, the reactional behavior is once again equal to the one explained in the guaiacol section.

VII.4 Conclusion of Chapter VII

The chapter presented a description of the methodology behind the stochastic simulation algorithm and its adaptation to the model compounds used in the present thesis: guaiacol and furfural. The chapter also included the transformation of the simulator from an isothermal mode to a dynamic system.

In general, the SSA is a method usually applied to mixture with complex reaction networks. The method is able to describe a reaction network molecule by molecule through a discrete evolution in time, while keeping track of the molecular detail. This feature gives the method the power of generating the reaction network while the simulation is unrolling, thereby avoiding the pre-definition of the reaction pathways. The evolution of the system is given by a Monte Carlo approach intrinsic to the algorithm.

As seen in the Experimental results chapter, the dynamic periods (heating and cooling) of the batch reactor are important in the reactivity of both model compounds. For this reason, it is crucial to take them into account in the simulation of the process. For the dynamic adaptation, the temperature profile during the non-isothermal periods was implemented in the algorithm. The *in silico* profile is obtained based on the experimental points extracted from the pilot unit and the time in the simulation. Since the latter is stochastic and reactivity dependent, the *in silico* profile will also be stochastic and reactivity dependent. Additionally, the profile will obviously depend of the experimental profile. The experimental profile is the most important parameter and if not well extracted, the *in silico* profile may have some deviations.

In the present thesis, the SSA will simulate the hydrotreating process of two model compounds (guaiacol and furfural) and also of a bio-oil fraction. Concerning the hydrotreating of guaiacol and furfural, all possible hydrotreating related reactions identified in the State of the Art were implemented in the algorithm. Amongst these, certain undesirable reactions were also added, as hydration, methylation and gas reactions. Other undesirable reactions that may occur during hydrotreating were not implemented due to their bimolecular nature. In total, for the hydrotreating of guaiacol and furfural, twenty-six reactions were implemented in the algorithm: methanol decomposition and steam reforming, methanation of monoxide and dioxide of carbon (reversible reaction), water-gas swift (reversible reaction), carbonyl hydrogenation (reversible reaction), decarbonylation, demethoxylation , dehydration, demethylation, hydrodeoxygenation, methylation, transalkylation, unsaturated hydrogenation, hydration and the opening of rings. These last two reactions have several types depending of the molecule structure. To conclude, in the source code, the implemented reactions have restrictions and assumptions in order to avoid impossible or improbable products.

Chapter VIII. Stochastic results

This chapter will present the results of the stochastic simulation of both model compounds with the parameters obtained by the deterministic model and also the results of the stochastic estimation of the kinetic parameters. As mentioned before, the deterministic parameters will serve as starting point for the estimation of the kinetic parameters for the stochastic models. The aim of this chapter is to obtain a set of kinetic parameters which will later on be applied for the simulation of bio-oil.

The following chapter includes the results of the two kinetic models. Similarly, to the deterministic results, the experimental data used, the simulation with the deterministic parameters, the stochastic estimation procedure and a comparison between both methods will be presented.

VIII.1 Hydrotreating of guaiacol

As seen in the Experimental approach section, the hydrotreating of guaiacol was performed over a reduced NiMo/γ-Al_2O_3 and at 13 MPa ten times at different reaction times and temperatures. Thanks to the results obtained from these tests, it is possible to simulate the hydrotreating of guaiacol. This section will present, for the hydrotreating of guaiacol, the experimental data, the simulation with the deterministic results and the estimation of parameters through the SSA.

VIII.1.1 Experimental and simulation data

The experimental data used in the stochastic simulation is exactly the same used in the deterministic model. As mentioned before, the experimental tests of the hydrotreating of guaiacol were obtained in a batch reactor under 13 MPa of hydrogen over a reduced NiMo/γ-Al_2O_3 at a range of temperature between 150 °C to 300 °C and a range of reaction time between 1 h to 5 h. A blank test (0 h) at 300 °C was also done. These tests all had constant water content.

As seen in Chapter VII, the SSA describes a process molecule by molecule. Therefore, the results of the SSA are normally presented in molecules. In order to later on compare with the deterministic results, the results of the SSA were transformed in concentration. The concentration of guaiacol was obtained by Equation VIII-1, while the rest of the compounds by Equation VIII-2.

$$C_{guaiacol,t} = \frac{N_{guaiacol,t} \cdot n_{guaiacol,0}}{N_{guaiacol,0} \cdot V} \qquad \textit{Equation VIII-1}$$

$$C_{i,t} = \frac{N_{i,t} \cdot n_{guaiacol,t}}{N_{guaiacol,t} \cdot V} \qquad \textit{Equation VIII-2}$$

Where $C_{guaiacol,t}$ is the molar concentration of guaiacol at a given time t, $N_{guaiacol,t}$ is the number of molecules of guaiacol at a given time t, $n_{guaiacol,0}$ is the number of moles of guaiacol at the beginning of the experiment (0), $N_{guaiacol,0}$ is the number of molecules of guaiacol at the beginning of the simulation, $n_{guaiacol,t}$ is the number of moles of guaiacol at a given time t, $N_{guaiacol,0}$ is the number of molecules of guaiacol at a given time t, $C_{i,t}$ is the molar concentration of compound i at a given time t, $N_{i,t}$ is the number of molecules of compound i at a given time t and V is the volume of the reactor (0.5 L). The experimental concentrations used for the estimation of the kinetic parameters are summarized in Table VI-1.

Concerning the simulation process, there are several inputs required by the algorithm as seen in VII.1.2.1. First of all, the simulator needs the temperature profile for each, together with the time of each test, given by the last value in the profile. This time will match the time of each iteration in the simulation. The pressure of the process is also needed by the algorithm. In the present case, the process is isobaric for all the experimental tests.

Then, the simulation also requires the number of molecules to handle, the number of iterations and the number of registration points. As seen in the State of the Art, the number of molecules should be a large value in order to achieve results, which are suitable and comparable with the deterministic model. The number of iterations should also be large, because the higher the number of iterations, the more precise the results are. However, if the number of molecules and the number of iterations is too large, the simulation time will be too long. Therefore, a balance between the number of molecules and iterations and the time is needed. According to Oliveira (Oliveira 2013), a balance between the both can be achieved using ten thousand molecules (10 000) and fifth iterations (50). Regarding the number of registration points, it was selected the maximum allow by the algorithm: 190. Thanks to these registration points, it was be possible to follow the behavior of the hydrotreating process during the dynamic and the isothermal periods.

Finally, the kinetic parameters, in addition to the reactions rules also need to be selected. These depend on the nature of the process to simulate. In the present work, the process will be the hydrotreating of guaiacol. The kinetic parameters and reactions used for the simulation of this model compound will be discussed in the following section.

VIII.1.2 Deterministic parameters

As mentioned before, the deterministic approach intended to acquire an idea of the kinetic parameters range and accelerate the parameters estimation through the SSA for a larger reaction network. However, before starting the estimation in the SSA, there is the need to analyze the impact of the deterministic parameters in the SSA. As it happens, the results of the stochastic simulation match the determinist

results if the exact parameters are used, as seen in the open literature (Oliveira et al. 2012a, 2013b) (Oliveira et al. 2014) (Oliveira et al. 2016) (Vinu and Broadbelt 2012) (Shahrouzi et al. 2008). Nevertheless, and since the SSA can calculate the equilibrium constant for each molecule, the results will not be the same. The SSA calculates the equilibrium constant through the standard Gibbs free energy of formation and Equation VIII-3. In its turn, the standard Gibbs free energy of formation is obtained through Joback and Reid group contribution method (Joback and Reid 2007).

$$K_{eq} = exp\left(-\frac{\Delta G_0}{T \cdot R}\right)$$ Equation VIII-3

Apart from the equilibrium constant, the remaining kinetic parameters and reactions are exactly the same as those used in the deterministic model: Demethoxylation, demethylation, hydrodeoxygenation, saturation (reversible reaction) and transalkylation. The kinetic parameters associated to these reactions are summarized in Table VI-3.

VIII.1.2.1 Results

Figure VIII-1 presents the results obtained for the hydrotreating of guaiacol by the SSA through the application of the kinetic parameters obtained from the deterministic model. In general, the concentration predicted by the SSA for guaiacol, methanol and benzenediol has similar results and trends to those obtained by the deterministic model. For methane, although the results are not similar, the same trend can be detected, as shown in Figure VIII-1 *d)*. Regarding, methylcyclohexanediol and cyclohexanol, neither the results nor the trends are similar to those of the deterministic model.

The deviations between the stochastic and deterministic simulations may be due to the equilibrium constant of saturation and also the presence of other molecules, which are not present in the deterministic model. Concerning the equilibrium constant, for the deterministic model it was assumed the correlation of the equilibrium constant of cyclohexanediol↔benzendiol for all aromatization reactions. However, here the equilibrium constant is calculated for each compound about to undergo aromatization. Therefore, the equilibrium constant for methylcyclohexanediol↔methylbenzenediol and cyclohexanol↔phenol are different from the one of cyclohexanediol↔benzendiol. For this reason, the results for the concentration of methylcyclohexanediol and cyclohexanediol vary, while the concentration of benzenediol is basically the same in both simulations. To conclude, due to the presence of new molecules, the reactivity of certain compounds varies. As mentioned before, the SSA creates molecules due to reaction rules, so new molecules may be generated, even if these are not present in the deterministic approach. This feature of the SSA may impact the reactivity of the molecules, and since the SSA evolves thanks to the reactivity of the molecules, the results here generated may be different from those of the determinist approach, as the case of methane.

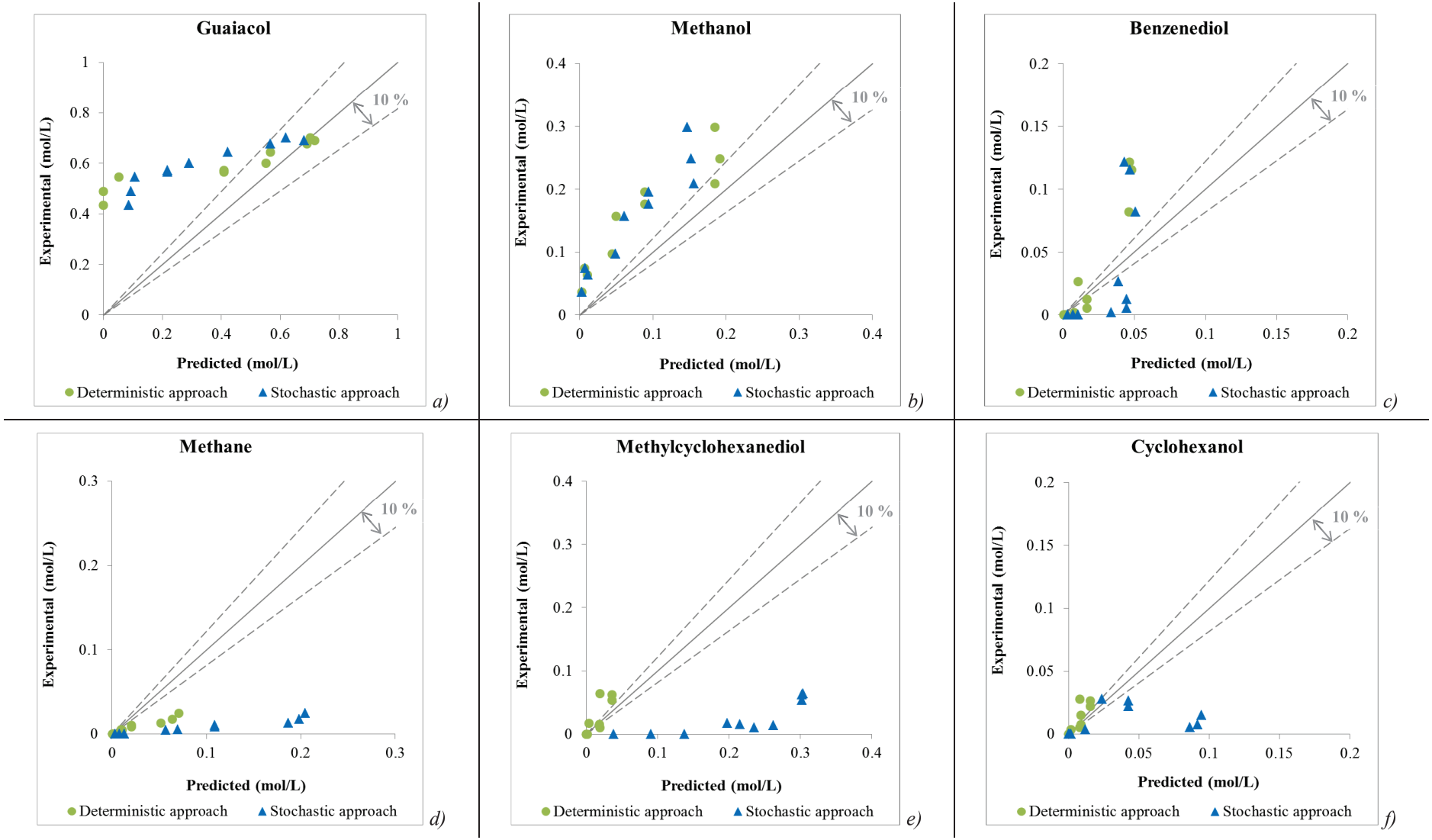


Figure VIII-1 – Comparison of the stochastic and deterministic with deterministic data: a) Guaiacol, b) Methanol, c) Benzenediol, d) Methane, e) Methycyclohexanediol and f) Cyclohexanol

VIII.1.3 Estimation of parameters

As presented in the previous section, the parameters obtained by the deterministic model are not suitable for certain compounds. However, these parameters represent an appropriate range of kinetic parameters and thus, can be used as starting point for the estimation of parameters via the SSA. As previously explained, the stochastic approach does not have an optimizer implemented. For this reason, and even with a base of kinetic parameters, the estimation procedure will be manual. Here it is important to note that a simulation in the SSA, with 10000 molecules and 50 iterations, takes around about 1 h and 15 min (Oliveira 2013). However, the addition of the dynamic periods of the reactor and the running of several simulations simultaneously, may increase the simulation time to around about 2 h or more. Therefore, the estimation procedure in the SSA is way more exhaustive and time consuming.

Regarding the reactions, the same used in the previous section and in the deterministic model will be applied in the estimation of the parameters via the SSA: Demethoxylation, demethylation, hydrodeoxygenation, saturation (reversible reaction) and transalkylation. As seen for the deterministic approach, kinetic constants for all the reactions follow the Arrhenius law. Therefore, ten kinetic parameters (pre-exponential factor and activation energy) may be accounted for the five reactions proposed in the guaiacol reaction network. These will then be estimated by the SSA thanks to experimental data available for the hydrotreating of guaiacol and summarized in Table VI-1.

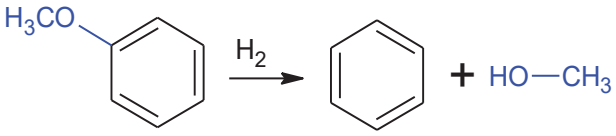
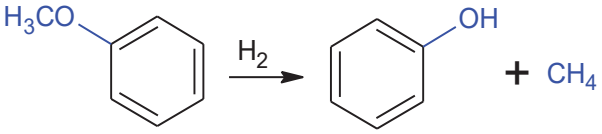
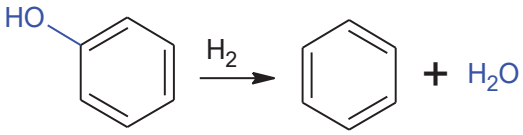
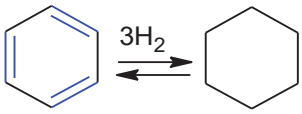
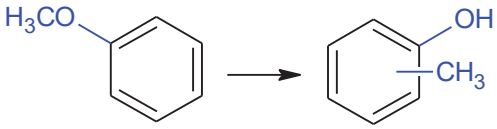
VIII.1.3.1 Results

The kinetic parameters estimated by the SSA for the experimental data of the hydrotreating of guaiacol are summarized in Table VIII-1. Similar to the deterministic model, only the molecules with higher selectivity and the reactant were accounted in the estimation process. Actually, and in order to later compare, only the compounds use in the deterministic model were used here: guaiacol, methanol, benzenediol, methane, methylcyclohexanediol and cyclohexanol. However, other molecules as phenol, cyclohexanediol and methylphenol could also be fitted by the SSA, unlike the deterministic approach. Here, only the pre-exponential constants were re-estimated. The activation energies are similar to those of the deterministic model, being therefore valid their coherency with the open literature.

The parameters of Table VIII-1 can be illustrated by Figure VIII-2 for each of the compounds used in the estimation procedure. In general, the stochastic model can predict within 10 % error margin all the experimental tests of guaiacol below 300 °C. For the remaining compounds, the model can only predict within 10 % error margin punctual experimental tests, as the concentrations of benzenediol at the blank test and 1 h at 300 °C, the concentration of methanol at 5 h and 300 °C, and finally the concentrations of methylclyclohexanediol at 250 °C and at the blank test at 300 °C. Although the remaining concentrations are under or overestimated, the predicted concentration of certain compounds as

methanol, methylclyclohexanediol and cyclohexanol, as the blank test at 300 °C and the experimental tests at 200 °c and 250 °C, are quite close the 10 % error area.

Table VIII-1 – Estimated kinetic parameters for the HDT of guaiacol in the SSA at 13 MPa over a reduced NiMo/γ-Al_2O_3 at a range of temperature between 150 °C to 300 °C and a range of reaction time between 0 h to 5 h.

Name	Reaction	k_0 (/s)	Ea (kJ/mol)
Demethoxylation (DEMO)		1 500	75
Demethylation (DEM)		500	82
Hydrodeoxygenation (HDO)		50	82
Saturation (SAT)		100	70
Transalkylation (TRANS)		1 000	75

For the experimental points at 150 °C, compounds as methane, benzenediol, methylcyclohexanediol and cyclohexanol were not detected as mentioned before. For the case of methane, benzenediol and cyclohexanol, the model predicts concentration values small enough to be zero (report to appendix C), thus predicting well what is expected from the experimental tests. However, for methycyclohexanediol, the model predicts higher concentrations, which are no longer negligible.

The deviations between the model and the experimental data may come from several origins, such as the absence of some reactions (methylation, equilibrium OH↔CO, etc.) (Ozagac 2016) (Wang et al 2013), neglecting of the catalyst effect, the imprecision of the kinetic parameters, and, the most probable cause, catalyst deactivation. The first two hypotheses require the redefinition of the source code, while the third cause requires more experimental data and more simulations. As for catalyst deactivation, dedicated experiments should be conducted to validate this hytothesis. Be this as it may, to achieve a solution for at least one of these hypotheses, more work and time is needed.

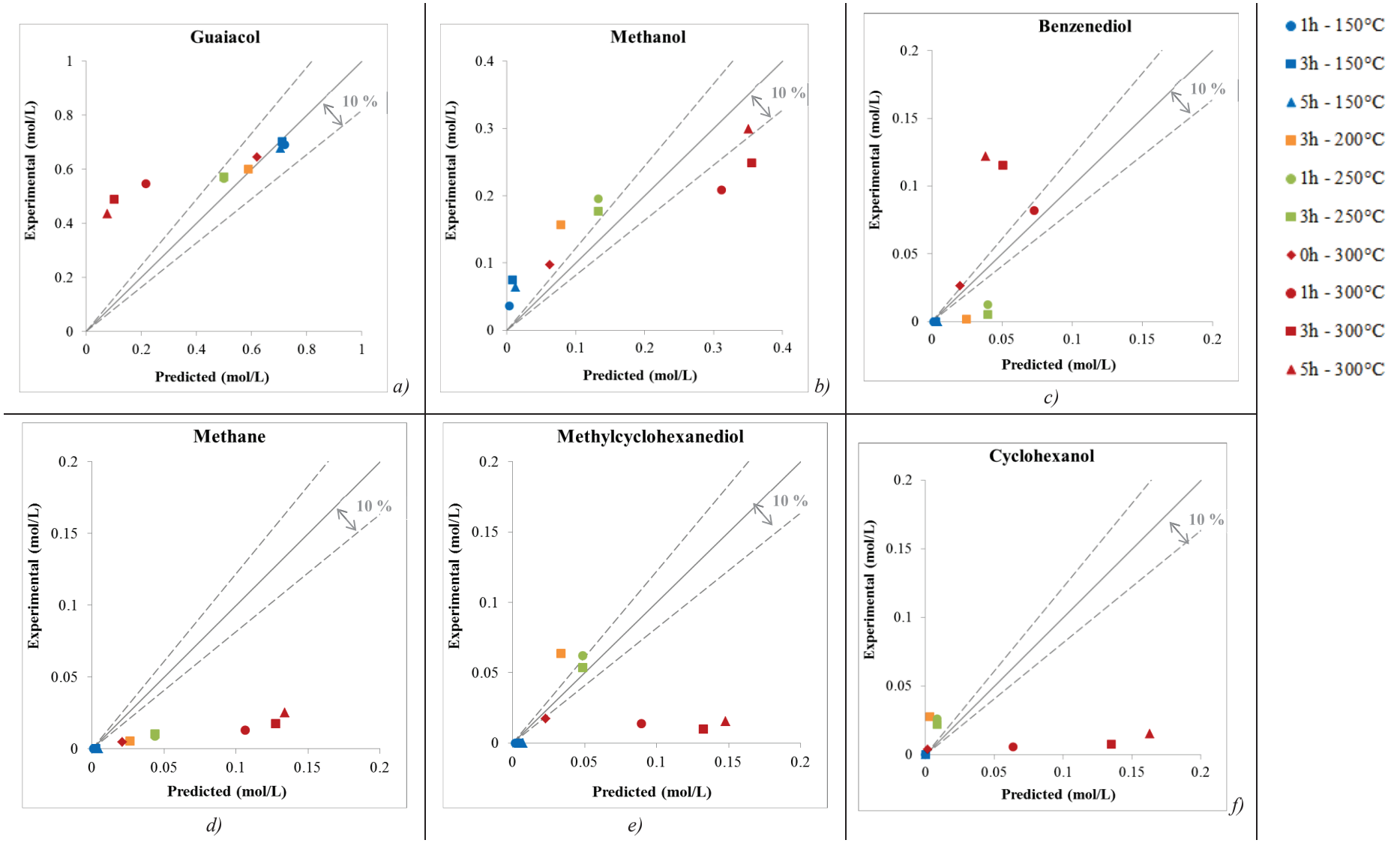


Figure VIII-2 – Parity plots of the stochastic model of guaiacol HDT: a) Guaiacol, b) Methanol, c) Benzenediol, d) Methane, e) Methycyclohexanediol and f) Cyclohexanol.

VIII.1.4 Comparison of kinetic models

Figure VIII-3 compares the deterministic and stochastic model for the concentrations of all the compounds used in the estimation of parameters. In general, the stochastic model can predict better the concentration than the deterministic model of the major compounds, guaiacol and methanol. However, in the case of the methanol, the stochastic model overestimates the concentration at high temperatures (300 °C). Be that as it may, both models predict the similar trends for this compound, as well as for guaiacol and benzenediol. For this last compound, the stochastic model overestimates more than the deterministic approach the experimental tests at low temperatures (150 °C – 250 °C). On the contrary, for the tests at higher temperature (300 °C), the stochastic model overestimates the concentration of methycyclohexanediol and cyclohexanol. Regarding methane, both models share the same trend. However, the stochastic model overestimates its concentration.

As mentioned before, the deterministic models accounts for less compounds and for a correlation for the equilibrium constant for saturation, which is identical for all saturations. On the contrary, the stochastic model can calculate the equilibrium constant of saturation for each specific compound undergoing aromatization and accounts with further more products. Even with a better approach for the equilibrium constant, the diversity of compounds, and thus reactivity, increases the complexity of the fitting by the stochastic approach.

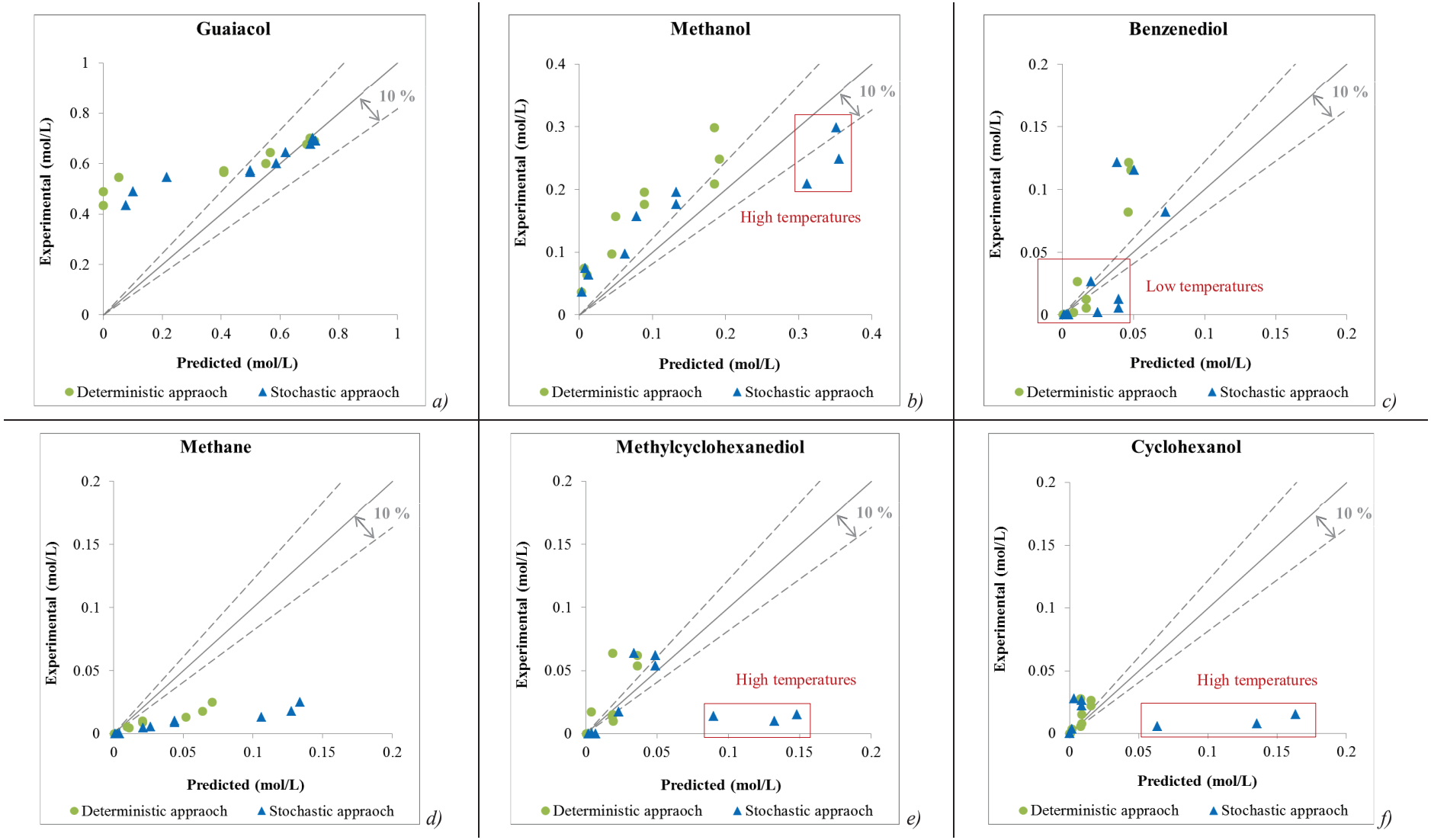


Figure VIII-3 – Comparison of the stochastic and deterministic models of guaiacol HDT: a) Guaiacol, b) Methanol, c) Benzenediol, d)Methane, e) Methycyclohexanediol and f) Cyclohexanol..

VIII.2 Hydrotreating of furfural

As seen in the chapters of the experimental part, the hydrotreating of furfural was performed over a reduced NiMo/γ-Al_2O_3 and at 13 MPa ten times at different reaction times and temperatures. Thanks to the results exploited from these tests, it is possible to simulate the hydrotreating of furfural. This section will present for the hydrotreating of guaiacol, the experimental data, the simulation with the deterministic results and the estimation of parameters through the SSA.

VIII.2.1 Experimental and simulation data

The experimental data used in the stochastic simulation is exactly the same used in the deterministic model. As mentioned before, the experimental tests of the hydrotreating of guaiacol were obtained in a batch reactor under 13 MPa of hydrogen over a reduced NiMo/γ-Al_2O_3 at a range of temperature between 100 °C to 150 °C for 1 h. Two blank tests (0 h) at 100 °C and 250 °C was also done. These tests all had constant water content.

Similar to guaiacol, the results of the SSA were transformed in concentration. The concentration of furfural was also obtained by Equation VIII-1 by replacing guaiacol by furfural. The concentration of the remaining compounds was once again obtained by Equation VIII-2. For furfural, the experimental concentrations used for the estimation of the kinetic parameters are summarized in Table VI-4.

As explained before, the simulation process needs the temperature profile and the time of each test. The latter is given by the last value in the temperature profile. This time will match the time of each iteration in the simulation. Similar to guaiacol, the process is isobaric for all the experimental tests. Concerning the number of molecules to simulate, the number of iterations and the number of registration points, the same value used in the hydrotreating of guaiacol were assumed: Ten thousand molecules (10 000), fifth iterations (50) and one hundred and ninety registration points (190). For the reactions and its kinetics, these depend on the hydrotreating of furfural and will be discussed in the following section.

VIII.2.2 Deterministic parameters

As already explained for the hydrotreating of guaiacol, the deterministic approach intended to acquire an idea of the kinetic parameters range and accelerate the parameters estimation through the SSA for a larger reaction network. Nevertheless, and since the SSA can calculate the equilibrium constant for each molecule, there is the need to analyzed the impact of the deterministic parameters in the SSA. AS mentioned previously, the SSA calculates the equilibrium constant through the standard Gibbs free energy of formation and Equation VIII-3. In its turn, the standard Gibbs free energy of formation is obtained through Joback and Reid group contribution method (Joback and Reid 2007).

Similar to guaiacol, besides the equilibrium constant, the kinetic parameters and reactions are exactly the same as those used in the deterministic model: Carbonyl hydrogenation (reversible reaction), decarbonylation, hydrodeoxygenation, ring opening and saturation (reversible reaction). The parameters associated to these reactions are in Table VI-6.

VIII.2.2.1 Results

Figure VIII-4 presents the results obtained for the hydrotreating of furfural by the SSA through the application of the kinetic parameters obtained from the deterministic model. The concentration predicted by the SSA for furfural has similar results to those obtained by the deterministic approach. For furfuryl alcohol, the second major compound in the effluent, the SSA can predict the experimental points at 100 °C, but not those at higher temperatures. Even so, the SSA can predict better the experimental point at 250 °C than the deterministic model, as shown in Figure VIII-4 *b)*. For the remaining compounds, the SSA cannot predict the same results or trends, with exception of furan, which the deterministic model cannot predict suitably either.

As seen for the hydrotreating of guaiacol, the deviations between the stochastic and deterministic simulations may be due to the equilibrium constants of aromatization and dehydrogenation of the hydroxyl and also due to the presence of other molecules, which are not present in the deterministic model. Concerning the equilibrium constant, for the deterministic model it was assumed the correlation of the equilibrium constant of benzendiol↔cyclohexanediol for all aromatization reactions and the equilibrium constant of furfural↔furfuryl alcohol for all the hydroxyl dehydrogenations. However, and as mentioned before, the SSA calculates the equilibrium constant for each compound about to undergo aromatization or dehydrogenation of the hydroxyl. Therefore, the constants used in the SSA do not match to those used in the deterministic (except for the furfural↔furfuryl alcohol reaction). For this reason, the concentrations may vary, in particular for other compounds besides the furfural and furfuryl alcohol. To conclude, due to the presence of new molecules, the reactivity of certain compounds varies. As remark for guaiacol, the SSA creates molecules due to reaction rules, so new molecules may be generated, even if these are not present in the deterministic approach. This feature of the SSA may impact the reactivity of the molecules, and since the SSA evolves thanks to the reactivity of the molecules, the results here generated may be different from those of the determinist approach. For example, in the SSA THF alcohol and methylfuran can undergo OH↔CO equilibrium (inverse of carbonyl hydrogenation) and saturation, respectively. Since both reactions were not accounted before in the deterministic model, and are also impacted by the approximations of the equilibrium constants, the concentrations of both products diverge not only from the results of deterministic model, but also from the experimental data.

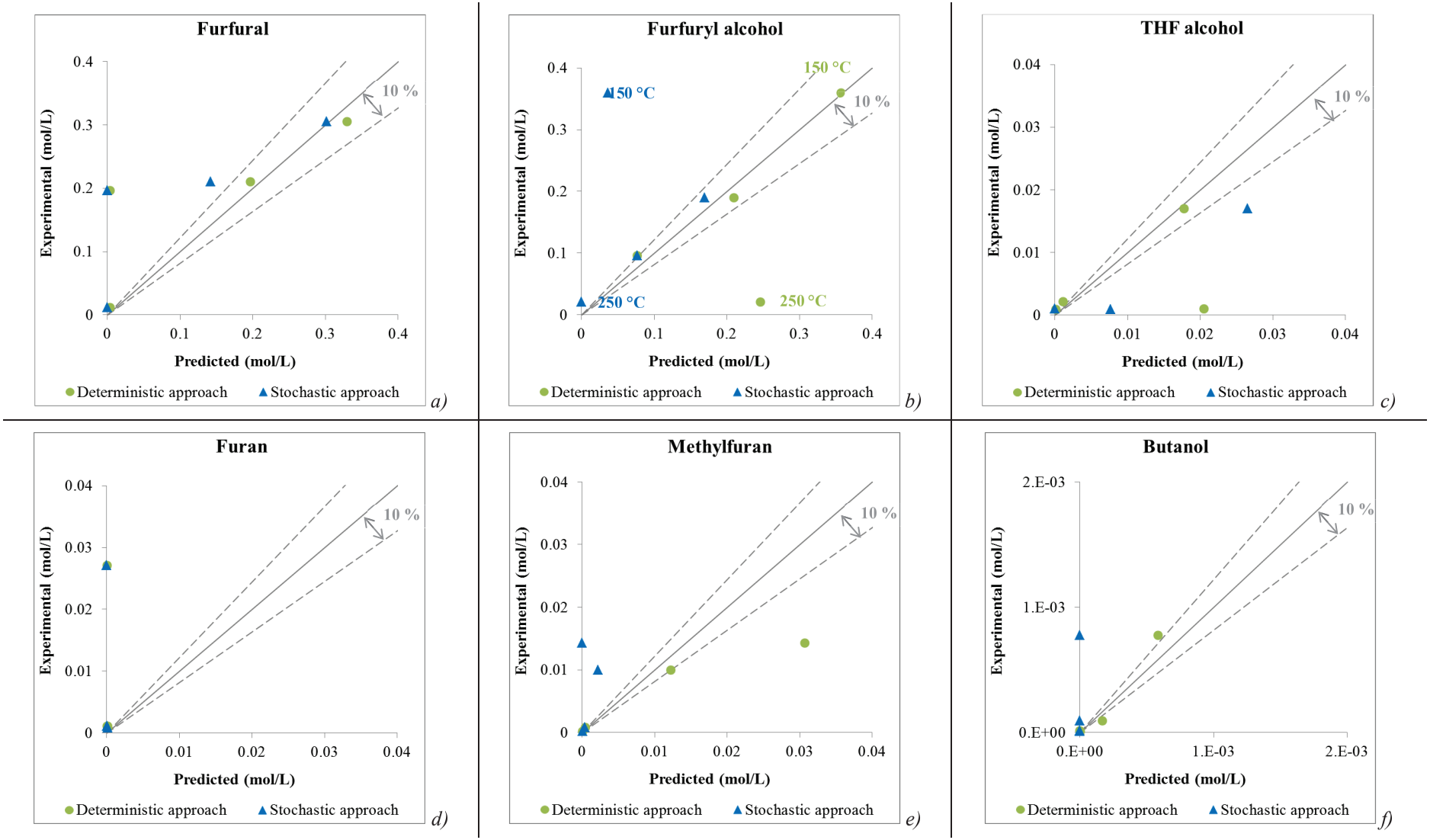


Figure VIII-4 – Comparison of the stochastic and deterministic with deterministic data: a) Furfural, b) Furfuryl alcohol, c) THF alcohol, d) Furan, e) Methylfuran and f) Butanol.

VIII.2.3 Estimation of parameters

As presented in the previous section, the parameters obtained by the deterministic model are not suitable for certain compounds. However, and similar to the hydrotreating of guaiacol, these parameters represent an appropriate range of kinetic parameters and thus, can be used as starting point for the estimation of parameters via the SSA. In the present case, as for guaiacol, the addition of the dynamic periods of the reactor and the running of several simulations simultaneously, may increase the simulation time to around about 2 h or more. Therefore, the estimation procedure here applied in the SSA is way more exhaustive and time consuming.

Regarding the reactions, the same used in the previous section and in the deterministic model will be applied in the estimation of the parameters via the SSA: Carbonyl hydrogenation (reversible reaction), decarbonylation, hydrodeoxygenation, ring opening and saturation (reversible reaction). As seen for the deterministic approach, kinetic constants for all the reactions follow the Arrhenius law. Therefore, ten kinetic parameters (pre-exponential factor and activation energy) may be accounted for the five reactions proposed in the furfural reaction network. These will then be estimated by the SSA thanks to experimental data available for the hydrotreating of guaiacol and summarized in Table VI-4.

VIII.2.3.1 Results

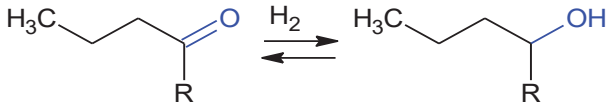
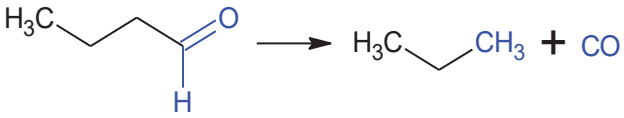
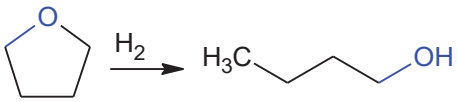
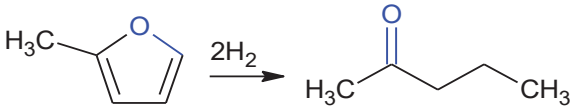
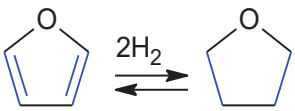
The kinetic parameters estimated by the SSA for the experimental data of the hydrotreating of furfural are summarized in Table VIII-2. Similar to the deterministic model, only furfuryl alcohol, furfural, THF alcohol, methylfuran, furan and butanol were accounted for the estimation of process. Here, only the pre-exponential constants were re-estimated. The activation energies are the same as those of the deterministic model, being therefore valid their coherency with the open literature.

The parameters of Table VIII-2 can be illustrated by Figure VIII-5 for each of the compounds used in the estimation. The stochastic model can predict reasonably well the concentrations at 100 ºC and 150 ºC for most of compounds. At 250 ºC, the model cannot predict any concentration. As mentioned before, at high temperatures other reactions, as hydration, start to emerge. For this reason, the reactivity of the system changes, together with the effluent composition. Furthermore, hydration reactions compete with the hydrogenation reactions for the catalytic sites (Wang et al. 2013), thus reducing the generation of desirable molecules, as hydrocarbons.

The stochastic model also has some difficulties in predicting the concentration of butanol and the concentration of furan at lower temperatures. In the case of furan, this may be related with the absence of bimolecular molecules, which consume furanic species for the generation of macromolecules (Ozagac 2016). This can also affect the prediction of the concentration of butanol. If furan is consumed through other reactions besides those defined in the SSA, the model will never be able to correctly predict furan

and hence butanol. In addition, the presence of more pathways which lead to butanol may increase the difficulty of the adjustment.

Table VIII-2 – Estimated kinetic parameters for the HDT of furfural by the SSA at 13 MPa over a reduced NiMo/γ-Al_2O_3 at a range of temperature between 100 °C to 250 °C and a range of reaction time between 0 h to 1 h.

Name	**Reaction**	**k_0 (/s)**	**Ea (kJ/mol)**
Carbonyl Hydrogenation (COH)		10 000	56
Decarbonylation (CO)		3 000	70
Hydrodeoxygenation (HDO)		12 000	75
Ring Opening (RO)	 	600	60
Saturation (SAT)		60	60

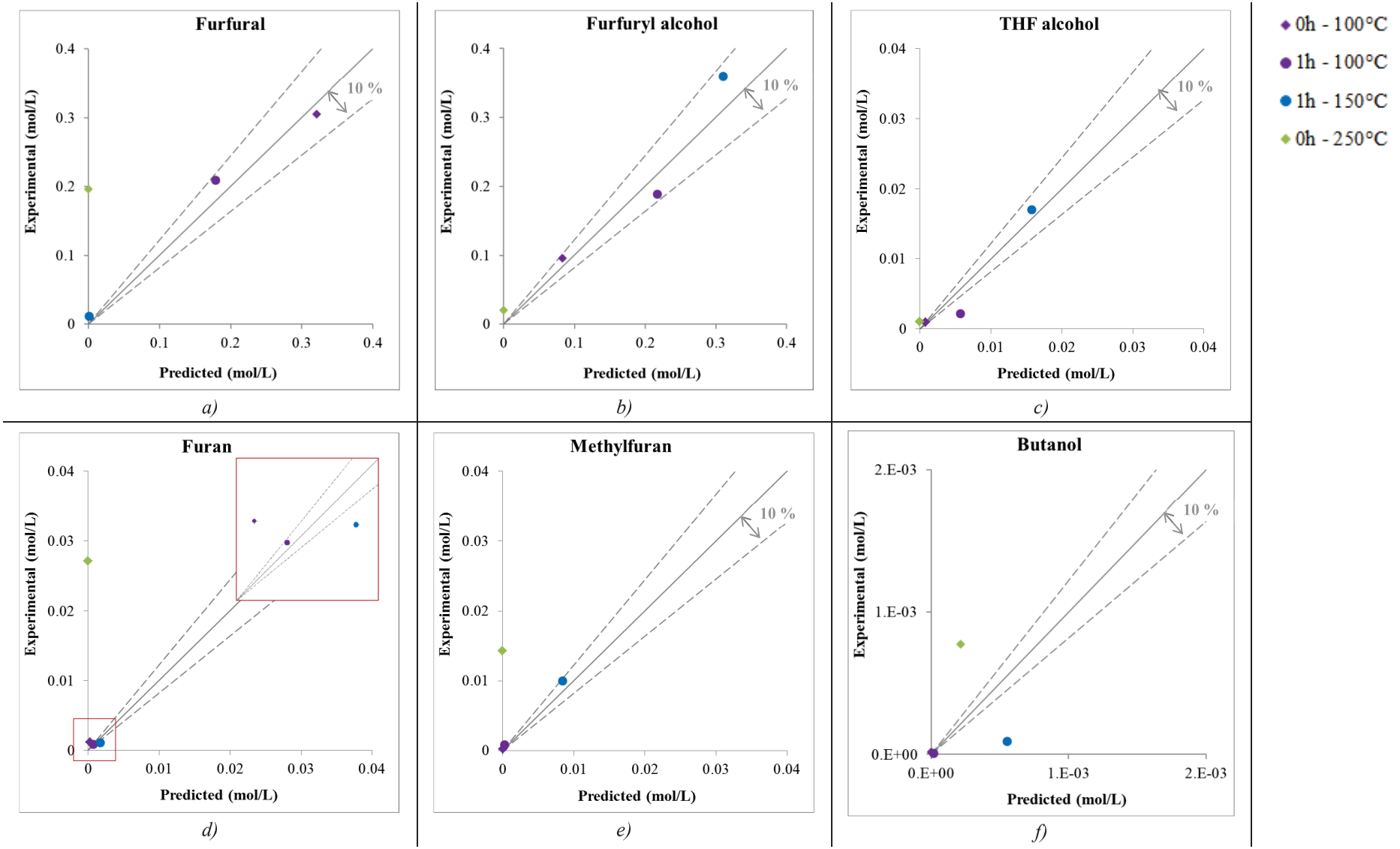


Figure VIII-5 – Parity plots of the stochastic model of furfural HDT: a) Furfural, b) Furfuryl alcohol, c) THF alcohol, d) Furan, e) Methylfuran and f) Butanol.

VIII.2.4 Comparison of kinetic models

Figure VIII-6 compares the deterministic and stochastic model for the concentrations of all the compounds used in the estimation of parameters. In the case of furfural, both models have similar results and the same trend. Also, both can predict reasonably well the concentration of furfural. A similar behavior is seen for furfuryl alcohol. Both models present similar results and trend, except for the experimental point at 250 ºC, which is better predicted by the stochastic model.

In the case of THF alcohol, stochastic model can predict better the concentrations than the deterministic model, except the experimental point at 1 h and 100 ºC. However, the model can predict the concentration with a degree of deviation inferior of that of the experimental point at 250 ºC of deterministic model. Regarding furan, methylfuran and butanol, both models have different results and trends. For the two former compounds, the stochastic model can predict better the concentrations than the deterministic model, even if associated to certain deviations in the case of furan. However, the deterministic model can predict better the concentration of butanol than the stochastic model.

To conclude, the parameters estimated by the SSA can predict better the majority of the concentrations than the parameters estimated by the deterministic approach. The main reason behind this higher performance is the definition of the equilibrium constants. While the deterministic model is based on approximate equilibrium constants (report to section VIII.2.2), the SSA calculates each equilibrium constant for each reaction. Through accounting with the nature of the reactions, the simulator (model) can achieve better performance in the predictability of the concentrations.

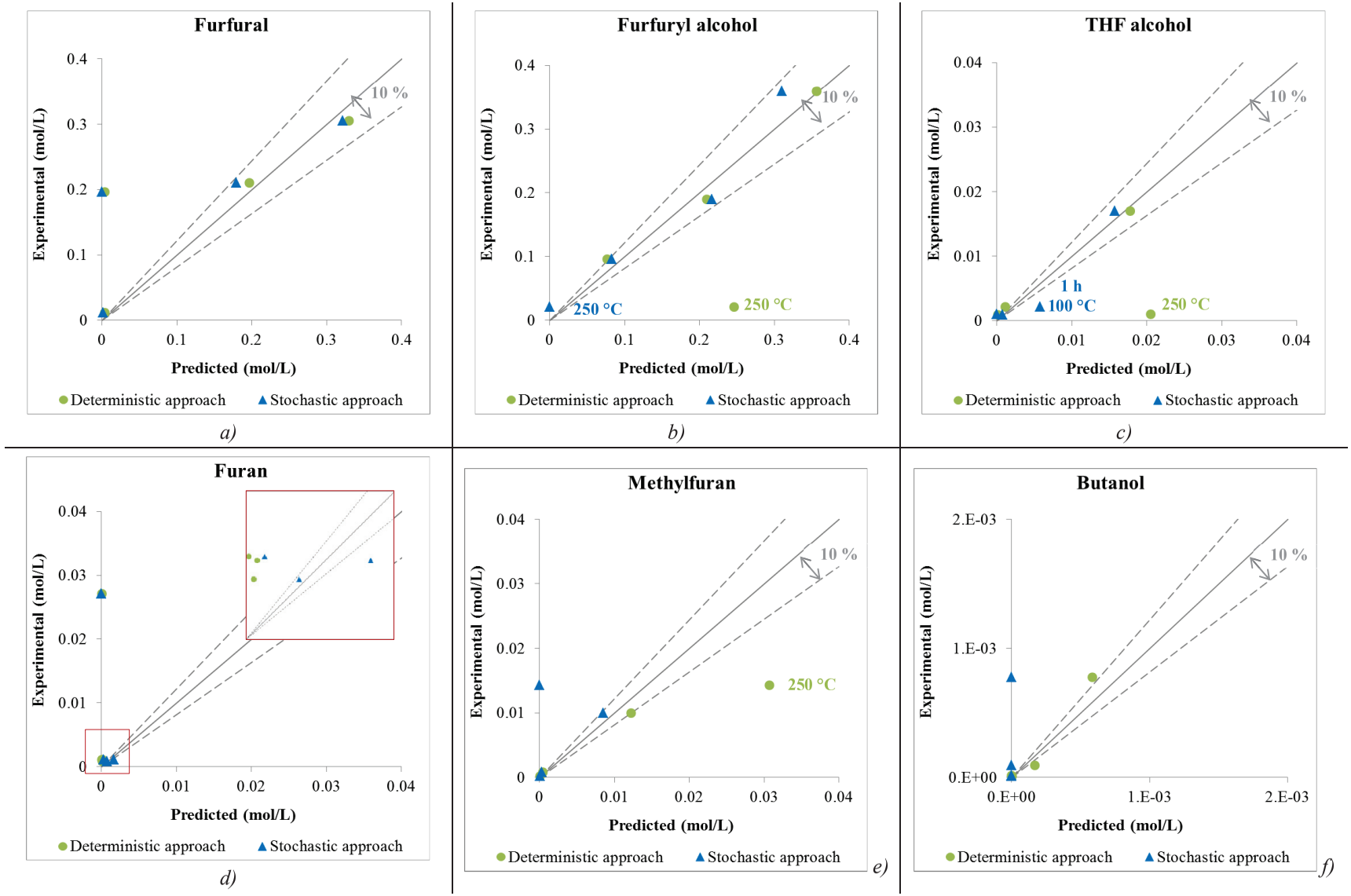


Figure VIII-6 – Comparison of the stochastic and deterministic models of furfural HDT: a) Furfural, b) Furfuryl alcohol, c) THF alcohol, d) Furan, e) Methylfuran and f) Butanol.

VIII.3 Conclusion of Chapter VIII

In the present chapter, the results of the stochastic simulation with the parameters of the deterministic model and the stochastic model were presented and discussed for both model compounds: guaiacol and furfural. These models were done using the experimental data available, the deterministic model, the stochastic simulator and the procedure of estimation of parameters.

For guaiacol, the stochastic simulations with the deterministic parameters predict similar trends for guaiacol, methanol and benzenediol to those obtained by the deterministic model. However, the remaining compounds present a higher concentration than expected. The deviations may be due to the equilibrium constant of aromatization and also to the presence of other molecules, which are not present in the deterministic model.

In the case of the stochastic model, the SSA can predict the concentration of the major compounds, guaiacol and methanol, while the remaining molecules are overestimated or underestimated. The deviations found in the model may be justified by catalyst deactivation, unsuitable parameters, the absence of reactions and the disregard of the impact of the molecule structure and the catalytic sites in the reactivity. Concerning the kinetic parameters, it was assumed the same activation energy values. However, the pre-exponential constants were re-fitted via the SSA. The pre-exponential factors found for the proposed reaction network follow the order describe in Equation VIII-4.

$$k_{0,DEMO} > k_{0,TRANS} > k_{0,DEM} > k_{0,SAT} > k_{0,HDO} \qquad \textit{Equation VIII-4}$$

For furfural, the stochastic simulations with the deterministic parameters predict similar trends for furfural to those obtained by the deterministic model. However, the remaining compounds did not present the same results or trends. Similar to guaiacol, the deviations may be due to may be due to the equilibrium constant of aromatization and also the presence of other molecules, which are not present in the deterministic model. Furthermore, the correlation considered for the equilibrium constant of OH↔CO may also impact the results.

In the case of the stochastic model, the SSA can predict reasonably well the concentrations of most of compounds at low temperatures. At higher temperatures, the model cannot predict any concentration, probably due to the presence of reactions which were not accounted for neither in the deterministic nor in the stochastic model, due to their incompatibility. Concerning the kinetic parameters, it was assumed the same activation energy values. However, the pre-exponential constants were re-fitted via the SSA. The pre-exponential factors found for the proposed reaction network follow the order describe in Equation VIII-5.

$$k_{0,COH} > k_{0,HDO} > k_{0,RO} > k_{0,CO} > k_{0,SAT} \qquad \textit{Equation VIII-5}$$

To conclude, the SSA allows the creation of a kinetic model for each model compound. The model for guaiacol presents some shortcomings, which with a correct kinetic approach may be reduced. However, the lack of time prevents such solution. The model for furfural can represent very well the concentration of the major effluents at lower temperatures. With further experimental data and the inclusion of bimolecular reactions, the kinetic model here presented could be extended to higher temperatures. Nevertheless, it is important to remark once more the difficulty associated to the fitting of both models, especially by the SSA. Besides the lack of experimental data and intrinsic information about kinetics, the time consumed by the SSA hinders an effective estimation of so many parameters (10).

PART IV: Hydrotreating of bio-oils

"Here, bio-oil has > 200 oxygen-containing compounds, representing nearly all kinds of oxygenated organics and oxygen function groups, which can undergo intertwining interactions, making it extremely difficult to understand HDO catalysis reaction pathways, mechanisms, kinetics, and property-reactivity correlation of hydrotreating catalysis."
(Wang et al. 2013)

Table of Contents

Chapter IX. Molecular reconstruction methodology

In collaboration with Rita Ribeiro

As seen in State of the Art chapter, the most suitable molecular reconstruction strategy for the present work consists in a coupling method between the Stochastic Reconstruction (SR) and the Reconstruction by Entropy Maximization (REM) techniques. Although both methods can be applied individually, the global performance of the reconstruction increases when they are coupled.

The following chapter will briefly show both procedures, focusing on their principle and functionality. Furthermore, this chapter will include the adaptation of the Stochastic Reconstruction and the Entropy Maximization techniques to bio-oils and their application. It is important to highlight that the adaptation of both techniques to the bio-based products was partly done by Ribeiro (Ribeiro 2018).

IX.1 Stochastic reconstruction

The stochastic reconstruction method generates an equimolar set of molecules. For such, this procedure considers that any molecule of a feedstock can be viewed as a combination of functional and structural attributes (as type of molecule, number of rings and etc.). In its turn, each one of these attributes is represented by a probability distributions function (PDF) (as histograms, gamma distributions and etc.). By random sampling, these allow the molecular characterization of the mixture. These steps will be presented with further detail while unrolling this section.

IX.1.1 Theoretical aspects

As previously mentioned, the stochastic reconstruction method assumes that a molecule can be characterized by a group of structure attributes. The idea is as simple as creating a character in a video game or creating a personal Lego man. However, while human characteristics are arms, legs, head, hair, color hair, type of hair, etc, molecules features are cores, number of cores, rings, type of rings, chains, heteroatoms, type of heteroatoms, etc.

As seen in State of the Art chapter, Boduszynski (Boduszynski 1987) (Boduszynski 1988) saw that petroleum cuts could be described by statistic distributions which vary according to their boiling point, molecular weight and number of carbons. Thanks to this principle, a team led by Klein (REF), was able to apply this concept to the reconstruction of oil fractions. This approach, as the stochastic simulation algorithm, is done thanks to a Monte Carlo method.

Although the stochastic reconstruction is a tool with potential, it also presents several disadvantages. The stochastic reconstruction method requires at least two distributions. These must describe the size of the molecules and their functional groups (FG). Both distributions report the basic information of a

mixture and without them, the method is not able to construct a coherent set of molecules. The same idea is applied to human kind. When trying to identify an individual, size (height and width) and sex (male/female) are the main features and without them, the individual identification is more complicated.

In summary, the stochastic reconstruction is based on two principles: all molecules are an ensemble of molecular attributes which are all described by a probability distribution function. The generation of a correct mixture depends mostly on both features and their coherency with the reality.

IX.1.2 Method description

As illustrated in Figure IX-1, stochastic reconstruction method not only needs to know the molecule attributes and its distributions, but also how these are organized. Thanks to this last step, named of building diagram, it is possible to establish a consistent and coherent guideline for the reconstruction of a molecule. After assemble all the PDFs, the SR technique generates several molecules (a mixture) through a random sampling process of these. The pure component properties are calculated afterwards following to the calculation of the mixture properties through linear mixture rules.

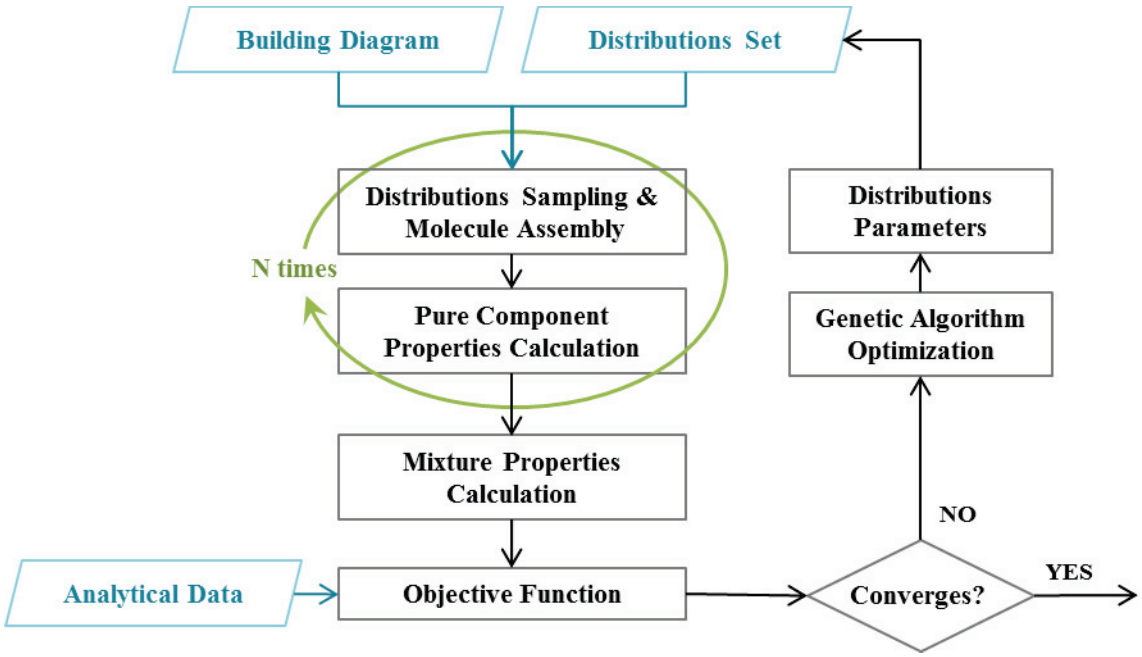


Figure IX-1 - Stochastic Reconstruction diagram.

The calculated characteristics will afterwards be compared with the given analytical data through an objective function. If different, the process undergoes an optimization loop, named of genetic algorithm, which will modify the parameters of the distributions. These will once again be tested by the process described above until the properties match and the number of iterations reaches the end. For a better understanding of the SR method, this section was divided into several steps corresponding to the order seen in Figure IX-1: Preliminary information (building diagram and distributions set), Generation of

molecules (distributions sampling and molecule assembly), Calculation of properties (pure and mixture properties), Objective function, Optimization loop and finally the new set of parameters.

IX.1.2.1 Preliminary information

As mentioned before, the generation of a correct mixture, depends mostly on the molecular attributes and their PDFs and their coherency with the reality. All grouped in the building diagram, meaning that this is the key feature for a correct reconstruction. This section will discuss in further detail the molecular attributes, the PDFs and the building diagrams.

Molecular attributes

The molecular attributes must be selected according to the mixture to be reconstructed. These must represent the great majority of the molecules in the mixture. The selection of structural blocks is based not only on empirical knowledge, but also on experimental data. For this reason, it is crucial to possess sufficient information to achieve a good characterization of the feedstock under study.

For the case of oil fractions, attributes as the type of molecule, number of cores, type of ring, number of heterocyclic rings, type of heteroatom, length of sidechains, etc. are usually applied. These are related with the oil fractions classical features, as can be seen in the work of (Hudebine 2003; Lopez Abelairas et al. 2016; Oliveira 2013). For bio-oil, certain of these attributes can be implemented, as the type of molecule or ring. However, due to the diversity of bio-oil and the high amount of oxygen, other more related attributes will be created.

Considering a given mixture of alcohols, mainly composed by *n*-paraffins, isoparaffins and saturated and aromatic rings. Typical molecules of this mixture, designated ***A***, are shown in Figure IX-2. This mixture could have as attributes: the type of molecule, length of chains, type of alkyl chain, number and type of rings and finally the probability of an atom accepting a hydroxyl group.

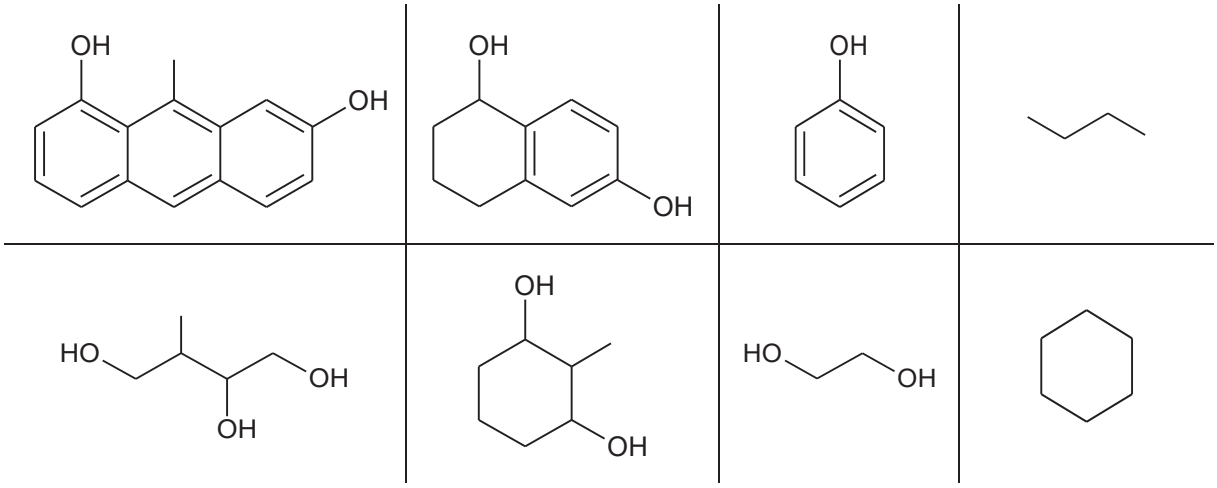


Figure IX-2 - Typical molecules of the example mixture A.

Probability Function Distributions

As seen before, each molecular attribute is associated to a probability function distribution, which returns a value/information for the given attribute when sampled. These distributions must have a great flexibility, in order to reach the several molecules of the mixture. Moreover, these PDFs also must have a small number of parameters to not damage the stochastic nature of the process. It is important to denote that not all distributions can describe in a coherent way all the molecular attributes. Therefore, the process of association must consider the range of values that an attribute can have. This means that by adjusting the range of parameters (maximum and minimum allowed), it possible to shape the PDF into a distribution close to the molecular attribute behavior.

In the source code used in the present work, there are 24 PDFs. At the moment, the PDFs mostly used in the reconstruction are the histogram, the gamma and the exponential distributions. The remaining 21 PDFs can be found in the reference (Ribeiro 2018). In further detail, the histogram is a discrete distribution, which is only applied when the attributes have a narrow range of possible values. The parameters of this PDF match the number of possible attributes minus one. On the contrary, the gamma and exponential distributions are continuous functions. These are suitable for attributes with a wider range of possible values. For attributes with large values, a gamma distribution is the most suitable. This function presents two parameters: shape parameter (α) and the scale parameter (β), however, and in order to reduce the number of parameters, it is considered that β is two times α. When large values are to be avoided, it is more appropriate to use an exponential distribution. This one possesses only one parameter (α).

To be implemented and used in the SR method, the continuous distributions need to be discretized. Only then is it possible to create a match between them and the values of the attributes. Figure IX-3 *a)* shows a continuous distribution, while Figure IX-3 *b)* represent the discretization (integration) of the same function. Besides this step, the distributions are transformed into cumulative distribution functions (CDFs), as presented in Figure IX-3 *c)*. Histograms are also transformed into CFDs.

The CDFs, as the PDFS, always have a minimum and maximum value, where usually the minimum is equal to zero, except for attributes where zero would mean the non-existing of a molecule, as the number of cores, or the contradiction of another PDF, as the length of chains. For the maximum value, histograms assume the biggest value of the attribute, while the continuous distributions may tend to infinite, where the results are negligible (improbably to happen). For this reason, these are truncated in the 0.99 probability and afterwards renormalized to assure a probability sum of 1, as shown in Figure IX-3 *d)*. In some cases, if the 0.99 probability matches a large value, which is not suitable for the attribute under question, the truncation process is done in the number acceptable for the given attribute (e.g. number of rings per core).

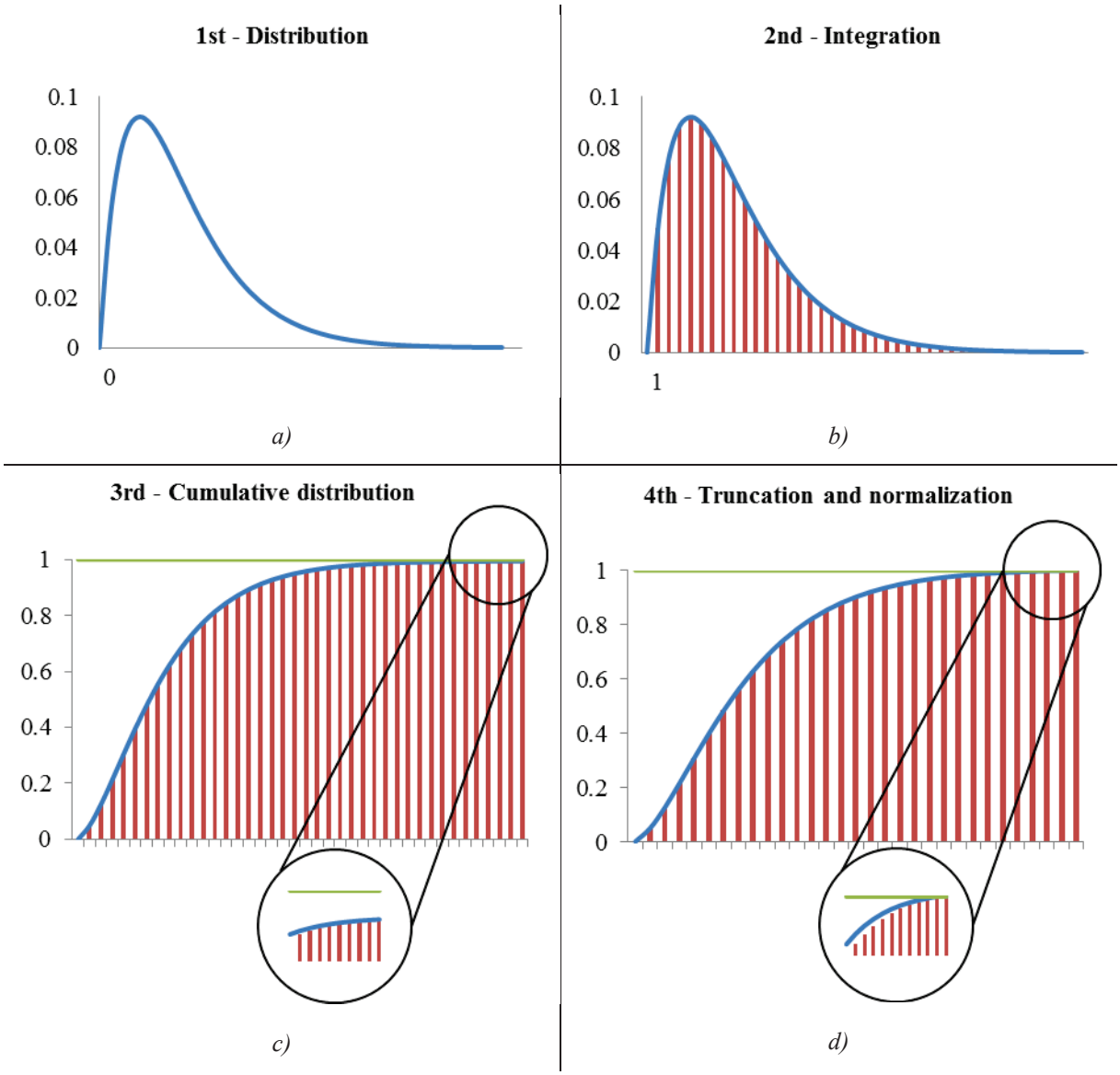


Figure IX-3 - Adjustments of the distributions for the sample procedure: a) Probability distribution function, b) integration of the PDF, c) accumulation of the PDF, now a cumulative distribution functions (CDFs) and d) truncation and normalization of the CFDs.

The choice of the PDF is related to the mixture features. As explained before, the PDF are related to the range of parameters. A small range requests a histogram, while a larger range requires a continuous function. Here, the attributes cannot have large values, an exponential distribution is most suitable than a gamma distribution. Obviously, this selection requires a certain knowledge of the mixture under question. When that information is unknown or not clear, the selection of the PDF for each attribute becomes more difficult. In these situations, the user must blindly choose and then, by trial and error, try to find the best PDF for the attribute.

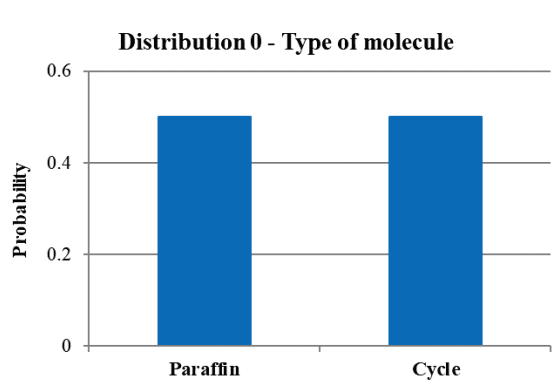
In the case of mixture A, the molecular attributes may be defined by the PDF defined in Table IX-1 and represented in Figure IX-4. For the type of molecule, it is either a paraffin or a ring, while for the type of ring, it is either a cyclohexane or a benzene and for the type of alkyl chain, it is either a methyl, or an

ethyl, or it does not exist (this feature will be explained later). As all attributes have a narrow range, all can be represented by a histogram. The same type of PDF is used for yes/no questions. For instance, the probability of an atom accepting a hydroxyl group is a yes/no question, which can be represented by histogram. Regarding the number of rings, a wider range can be considered. However, molecules with many rings may be suspicious, so from the wider range, the larger values should be avoided. Therefore, this attribute can be represented by an exponential distribution. For the length of the chains, it depends on the size of chains that can be found in the mixture. Small chains can be described by an exponential, while medium and large chains can be described by a gamma function. A histogram could also be used for the small chains. However, this function would have more parameters to fit than the exponential distribution. Here, as in any modeling tool, overparametrization is not desirable. Assuming that the mixture contains large chains, the attribute can be described by a gamma function.

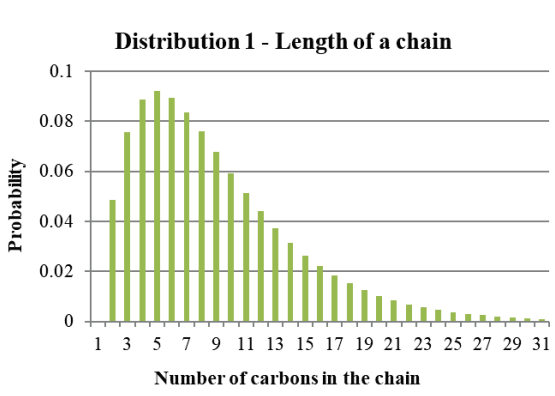
Table IX-1 - Molecular attributes and the respective PDF for mixture A.

PDF Number	Molecular attribute	PDF Type	Parameters Number
0	Type of molecule	Histogram	1
1	Chains length	Gamma	1
2	Type of alkyl chain	Histogram	2
3	Number of rings	Exponential	1
4	Type of ring	Histogram	1
5	Hydroxyl group	Histogram	1

To conclude, this section, imagine that actually mixture A has few and small chains, as in Figure IX-2, and molecules with several cores. In this case, the predicted molecules (and mixture) would be incorrect. In such situation, not only the PDF of the length of chains would be badly selected, but also an attribute would be missing. Only by reconstructing the mixture with an exponential distribution and an attribute for the cores, it would be possible to obtain a correct prediction. So, in the end the choice of the attributes and the PDFs will always be a trial and error method versus the unknown.



a)



b)

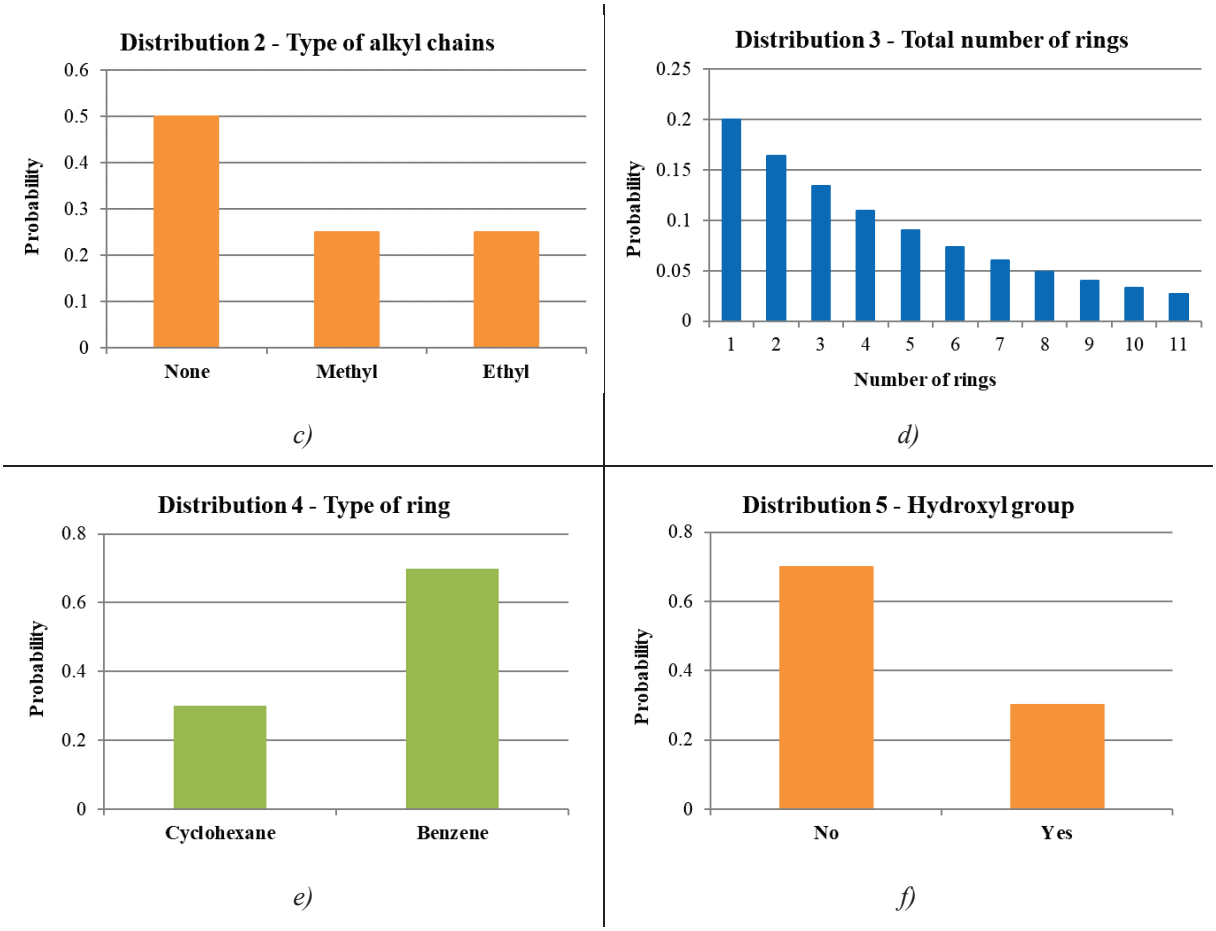


Figure IX-4 – PDFs for the example mixture A: a) distribution 0 – type of molecule; b) distribution 1 – length of chains; c) distribution 2 – type of alkyl chain; d) distribution 3 –total number of rings; e) distribution 4 - type of ring and f) distribution 5 – hydroxyl group.

Building diagram

As mentioned before, a building diagram is a decision tree of molecular attributes (structural blocks). For the edification of a building diagram, firstly there is the need to identity all the molecular attributes and their PDFs. Afterwards, the creation of the building diagram should be done since the most basic information, as type of molecule, until the more detailed information, as type of heteroatom. The idea is to create a guideline for a molecule construction, which starts from the base until the top, similar to the construction of a house, which starts from the ground to the walls to the rooftop.

Besides allowing a hierarchic and coherent organization, the main advantage of the building diagram concept is the preclusion of impossible and/or improbable molecules. Thanks to blocks organization, the building diagram is able to identify and discard all molecules that infringe the thermodynamic and/or probabilistic grounds. In the end, only the correct molecules are maintained. These may later be discarded if there is no convergence between the synthetic and real mixture properties.

Considering the example mixture, when organizing all attributes, a potential building diagram could be Figure IX-5. Other schemes can be created from the same structural attributes, by reorganizing them in

different positions in the tree. According to Figure IX-5 organization, when constructing a molecule from mixture A, the first question is the type of molecule. This can be a paraffin or a cycle molecule.

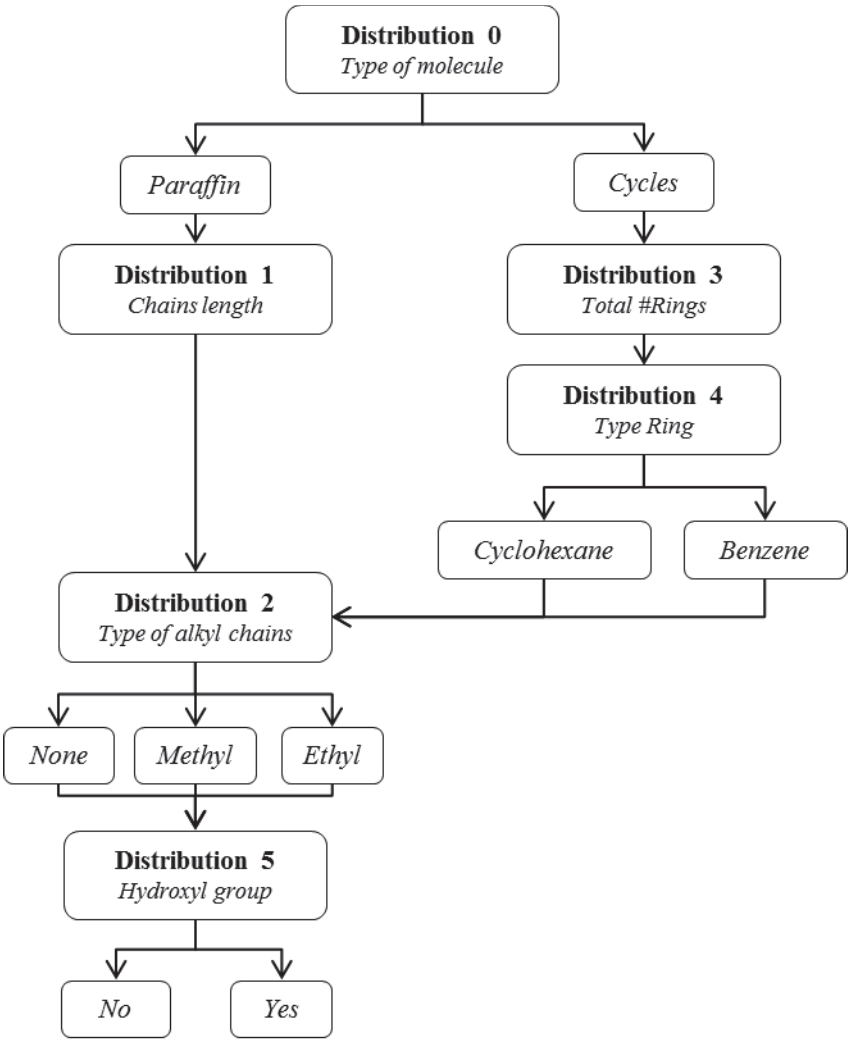


Figure IX-5 - Building diagram for mixture A.

In the case of being a cyclic molecule, how many rings will this molecule have? After knowing the number of rings, the building diagram will check for each ring, what is its type. The rings may be cyclohexanes (saturated rings) or benzenes (aromatic rings). Afterwards, the presence and size of possible side chains is chosen. This distribution checks for all the carbons that can be substituted in the ring (>CH_2 and >CH-). For those, the algorithm will sample a type of alkyl chain. If the answer is "none", the carbon atom will stay as it is. If the answer is a "methyl" or a "ethyl", a sidechain with one carbon or two carbons, respectively, will be added to the carbon in the ring. To conclude, if the cyclic molecule has sidechains, the algorithm will check if the terminal atom can be replaced by a hydroxyl.

In the case of being a paraffin, distribution 1 will give the size of the paraffin (length). Next, the presence and size of possible side chains are selected. As for the rings, this distribution will check for all the carbons that can be substituted (>CH_2 and >CH-). As it will be explained later, the alkyl chains

(branches) can only be added to core atoms of the paraffin (>CH_2 and >CH-), never terminal atoms (CH_4, -CH_3). As before, if the answer is “none”, the carbon atom will stay as it is. If the answer is a “methyl” or a “ethyl”, a branch with one carbon or two carbons, respectively, will be added to the carbon in the paraffin. As in the rings, the algorithm will check if the terminal atom, in this case of the paraffin and its branches (if existent), can be replaced by a hydroxyl.

IX.1.2.2 Generation of molecules

After defining all the inputs (preliminary information), the algorithm starts to run. The first step is the generation of molecules. This is done by sampling the CDF of the attributes according to the building diagram. The concepts used in the reconstruction of the molecules (name of the atoms, chains, substituents, rings, etc.) can be found in detail in the work Hudebine (Hudebine 2003) and Oliveira (Oliveira 2013). After the sampling of the attributes, the algorithm assembles all the attributes in the form of a molecule. This section will discuss in further detail these two features.

Distributions sampling and Molecules assembly

In general, the sampling process is based on a Monte Carlo method. As exampled in the Stochastic modeling methodology chapter, these methods are based on random events. The Monte Carlo method chooses a structural block value/information through randomization of a value between zero and one (both excluded). The result will then be equivalent to a given probability in the cumulative distribution and, as a consequence, to a specific value of the molecular attribute.

Considering mixture A and its features (see Figure IX-4 and Figure IX-5), several types of molecules can be sampled through the process presented before. As shown in Figure IX-6 *a)*, after random sampling (Figure IX-6 *a)* - 0.7), **molecule A** is defined as cyclic compound. According to the building diagram (Figure IX-5), the total number of rings are sampled (Figure IX-6 *c)* - 0.4). **Molecule A** is now a cyclic compound with three rings. Afterwards, the type of each ring is sampled (1^{st} – 0.35)(2^{nd} – 0.85)(3^{rd} – 0.6), all three rings are benzenes, as shown in Figure IX-6 *e)*. At the moment, **Molecule A** is a molecule with three benzenes.

In the case of **molecule B**, paraffin was the type of molecule sampled (Figure IX-6 *b)* - 0.3). According to the building diagram (Figure IX-5), the length of the paraffin is sampled (Figure IX-6 *d)* - 0.5), giving eight carbons. At the moment, **Molecule B** is a linear hydrocarbon with eight carbons (octane).

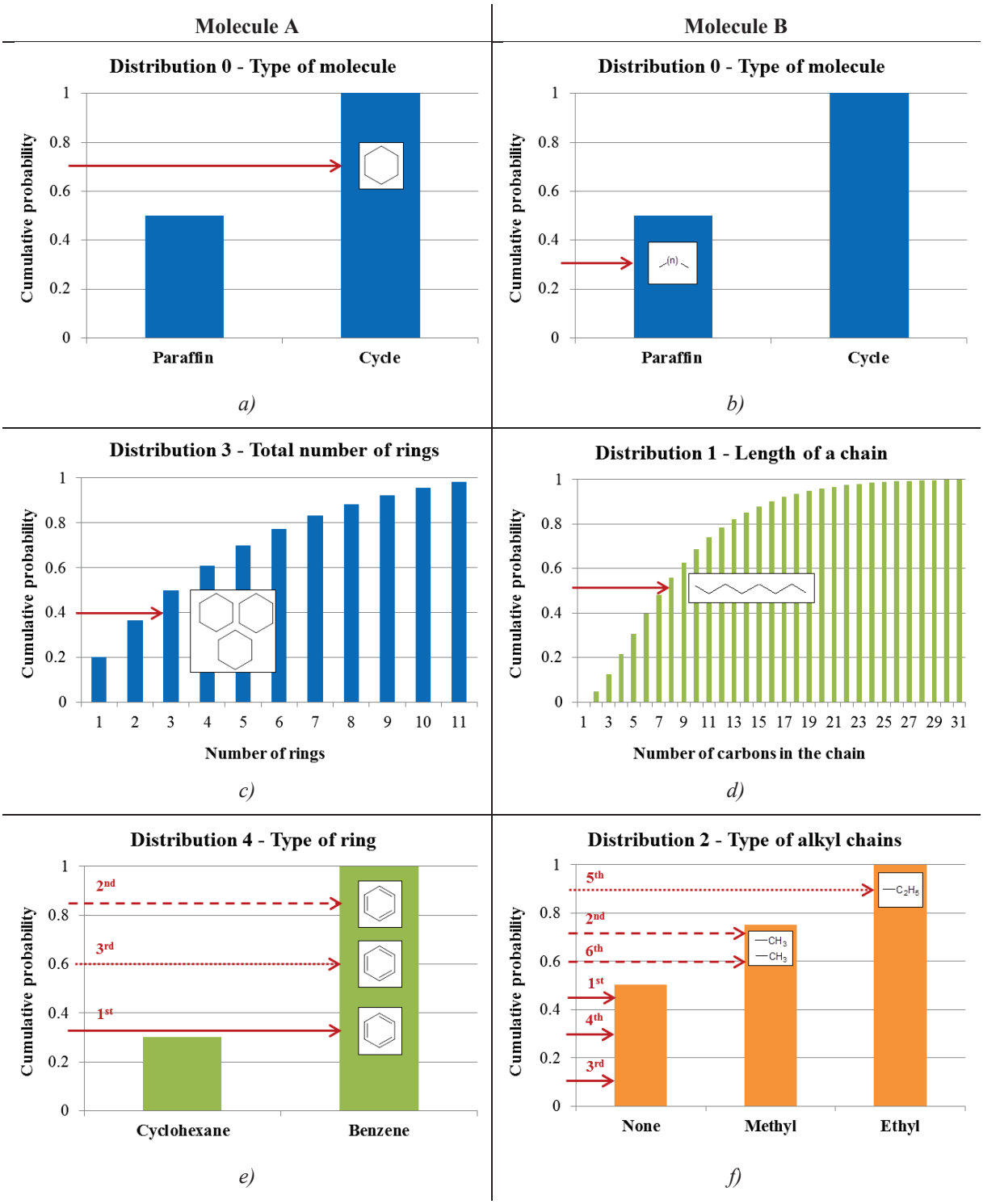


Figure IX-6 – Sampling of the CDFs for the example: a) sampling of the type of molecule for molecule A; b) sampling of the type of molecule for molecule B; c) sampling of the total number of rings for molecule A; d) sampling of the length of chains for molecule B; e) sampling of the type of rings for molecule A; f) sampling of the type of chains for molecule B.

Before continuing the sampled process, the previous structures are assembled. This procedure is done thanks to a technique designated hexagonal map. Briefly, this technique consists in grouping the structural blocks one by one in the map. Only after, the chains and the remaining attributes are sampled

and added to the core of the molecule. The assembly of molecules follows certain rules, as described in references (Hudebine 2003; Oliveira 2013) and here summarized:

Rings

- The construction of a core starts by connecting all aromatic and heterocyclic rings. Naphthenic rings are the last to be added in to the core.
- If heterocycles are sampled, these are the first to be placed in the map. The following aromatic rings and remaining heterocycles (if existent) are randomly added.
- The position next to the heteroatom of a heterocycle are forbidden (impossible liaison).
- Five atoms rings are placed in hexagonal units of the map, leaving an atom unfiled (peak). The positions next to this are forbidden (impossible liaison).
- The aromatic cycles cannot be placed in positions that have an impair number of empty peaks.
- The probability of a ring to assume a free position in the map is proportional to the number of cycles around that position (condensed molecules).

Chains

- Rings can only have branches in peripheral carbons which can be substituted.
- Chains can only have branches in non-terminal carbons which can be substituted.
- According to experimental data of heavy oil fractions, >CH- atoms in saturated rings have a lower probability of accepting a branch. The same was assumed for bio-oils.
- In a molecule with several cores, the first chains are inter-core chains.
- The cores are connected according to their creation order (the 1st with the 2nd and so on)
- The heteroatoms are added to the chains by replacing an existent atom.
- Chains cannot have two heteroatoms function groups in series.

In the case of **molecule A**, the first ring of the three benzene rings sampled is placed in the center of the map, as seen in Figure IX-7 A. Since there are no chemical constraints, any side of the ring can receive a ring neighbor with a probability of 1/6 (1position over 6 possible). The second ring is chosen through random sampling. As seen in Figure IX-7 B, the second ring is located at west of the first ring. Unlike before, the dimer is associated to chemical constrains. In the positions mark with an ✗the molecule cannot have ring neighbors. The dimer can accept any neighbor in the remaining positions. All have a 1/6 probability (1position over 6 possible). As seen in Figure IX-7 C, the third ring is located at southwest of the second ring. Once again, the trimer is associated to chemical constrains. These are represented by ✗, leaving the remaining as possible positions. Since there is no more rings, the algorithm will continue the sampling process.

In the case of **molecule B**, there is no rings to be positioned. Instead, there is a *n*-paraffin of eight carbons. This structure is as well placed in the middle of the map. If any further features need to the added, as the present example, the algorithm will continue the sampling process

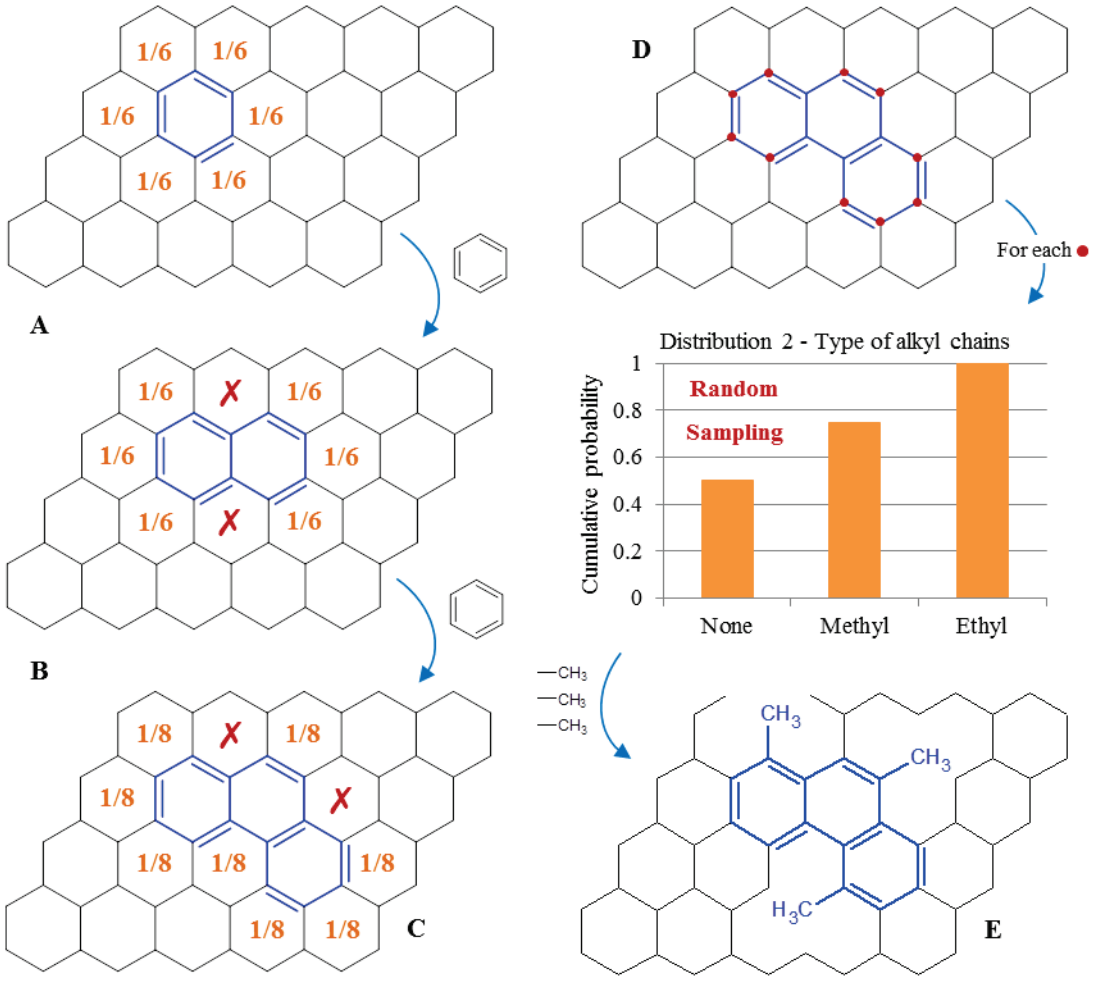


Figure IX-7 - Assembly of structural blocks into a molecule.

For **molecule A**, the algorithm will sample for each carbon that can be substituted (>CH_2 and >CH-, mark as red dot), its potential sidechain, as seen in Figure IX-7 D and E. For all the ten possible carbons, only three were substituted and all for methyl sidechains. The other seven fell in the probability of being "none". Afterwards, for each terminal carbon of the sidechains, it will be sampled the probability of replacing the carbon by a hydroxyl. For all the three possible carbons, two were replaced by a hydroxyl, as seen in Figure IX-8 *a)*.

For **molecule B**, the algorithm will also sample the type of branches for each carbon that can be substituted (>CH_2 and >CH-) (similar to Figure IX-7 D and E). For all the six carbons possible, three were substituted by branches: two methyl groups and one ethyl. The other three fell in the probability of being "none", as shown in Figure IX-6 *f)*. Afterwards, the algorithm will check the terminal carbons of the isoparaffin. For each, the probability of replacing the carbon by a hydroxyl will be sampled. For

all the five possible carbons, two were replaced by an hydroxyl, as seen in Figure IX-8 *b)*. In the end, the **molecule A** and **molecule B** appear as seen in Figure IX-8 *c)* and *d)*.

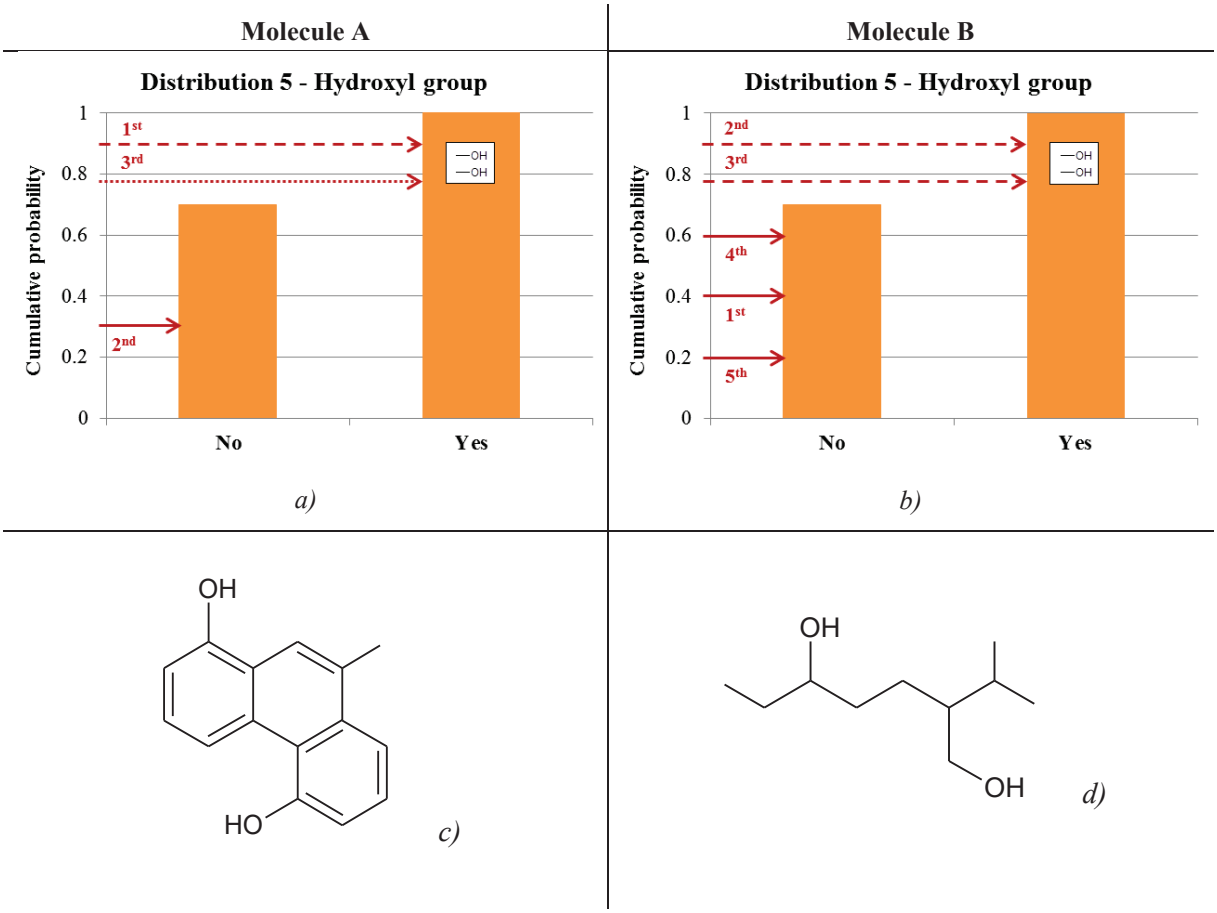


Figure IX-8 - Example molecules resultant of mixture A: a) sampling of the acceptance of a hydroxyl for molecule A, b) sampling of the acceptance of a hydroxyl for molecule B, c) Molecule A and d) Molecule B.

The process here presented for the generation of a molecule is repeated *N* times. By reproducing N molecules, the algorithm can create synthetic mixture with *N* molecules. The convergence between this mixture and the real feed is evaluated by the mixture properties. For this reason, the pure properties of each generated molecule are calculated. By linear mixing rules of these, the algorithm can finally obtain the mixture properties and compare both mixtures. The process of properties calculation is presented in the next section.

IX.1.2.3 Calculation of properties

As presented in Figure IX-1, after the generation of a molecule, the algorithm calculates its pure component properties. When the set of N molecules is finished and all their pure component properties are defined, the algorithm calculates the mixture properties through these.

Pure components properties

Through pure component properties, it is possible to discover crucial information regarding not only the atoms abundance and connectivity, but as well about the molecule structure and its chemical and physical characteristics. The calculation of these is done either by inspection of the structure, or by group contributions methods or by correlations. All the pure component properties used in the stochastic reconstruction and the entropy maximization can be found in Table IX-2. More information concerning these methods can be found in (Bondi 1964), (Joback and Reid 2007), (Oliveira 2013), (Ourique and Silva Telles 1997), (Sastri and Rao 1992) (Ribeiro 2018) and (Bour 2017).

Table IX-2 - Properties of the SR and REM methods.

Molecular properties calculation methods		
Direct inspection of the structure	**Group contribution / Correlations**	
	Normal boiling Point	Cohesion energy
Chemical formula	Melting point	Standard enthalpy of formation
Molecular weight	Critical temperature	Standard Gibbs free energy of formation
^{13}C NMR	Critical pressure	Standard entropy of formation
^{1}H NMR	Critical volume	Octanol/water partition coefficient
Mass spectrometry	SARA families	Van der Waals volume
PIONA families	Specific gravity	Surface Area
Basic nitrogen	Conradson carbon	Double Bond Equivalent
Oxygenated functional groups	Dynamic viscosity	Polystyrene equivalent molecular weight
FT-ICR/MS Families	Kinematic viscosity	Molecular refractivity
		Refractive index

Mixture properties

The properties of the synthetic mixture are calculated by the pure component properties of the molecules present in the mixture. Considering the mixture as ideal, its properties can be calculated by linear mixing rules, as shown in Equation VII-1, where MP_j is a j mixture property, x_i is the molar fraction, of compound i, $PP_{j,i}$ is a j pure component i property and N is the number of molecules.

$$MP_j = \sum_{i}^{N} x_i \cdot PP_{j,i} \qquad \textit{Equation IX-1}$$

The mixture properties used in the stochastic reconstruction and the entropy maximization are presented in Table IX-3. In the end, these are either an average value for the mixture, as is the case of the average molecular weight, or a value for fractions, as is the case of the SARA families, where each family has a

fraction and the sum equals one hundred. The mixture properties are used in the objective function, where the generated synthetic mixture is evaluated.

Table IX-3 – Mixture properties of the SR and REM methods.

Mixture properties calculation methods		
Elemental analysis	True boiling point distillation	Average molecular weight
^{13}C NMR	Simulated distillation	Average specific gravity
^{1}H NMR	SARA families	Average dynamic viscosity
Mass spectrometry	PIONA families	Average kinematic viscosity
Oxygenated functional groups	Conradson carbon	Average Double Bond Equivalent
FT-ICR/MS Families	Z-average molar mass	Number average molar mass
Size Exclusion Chromatography	Dispersity index	Mass average molar mass
Kendrick diagram		

IX.1.2.4 Objective function

As mentioned before, the synthetic mixture is compared to the real feed thanks to the mixture properties and through an objective function (OF). In the present work, this difference is compared by the least absolute deviations technique, as seen in Equation IX-2 and Equation IX-3. These deviations are not squared to avoid the amplification of the background noise, already associated to the stochastic nature of the process.

$$OF = \frac{\sum_{j=1}^{N_P}(w_j \delta_j)}{N_P} \qquad \textit{Equation IX-2}$$

Where N_P is the number of properties in the objective function OF, δ_j is the relative deviation of property j and w_j is the weight of property j, which is given by Equation IX-3, where $N_{M,j}$ is the number of measures of property j, $x_{k,j}^{exp}$ and $x_{k,j}^{calc}$ are the experimental and calculated values, respectively, of property j for a given measurement k.

$$\delta_j = \frac{1}{N_{M,j}} \cdot \sum_{k=1}^{N_{M,j}} \left(\frac{\left| x_{k,j}^{exp} - x_{k,j}^{calc} \right|}{x_{k,j}^{exp}} \right) \qquad \textit{Equation IX-3}$$

IX.1.2.5 Optimization of the parameters

In the case of no convergence by the objective function, the SR algorithm resources to an optimization loop. Several algorithms can be applied, however, and due to the noise associated to the SR method, it is required an optimization technique that does not use derived functions. Initially, Hudebine (Hudebine 2003) proposed a simulated annealing procedure. Afterwards, Oliveira (Oliveira 2013) adjusted to the

SR method to a genetic algorithm. The performance of both was analyzed by Schnongs (Schnongs 2005), who showed that the first technique added significant oscillations to the objective function. The author also denoted that the second technique also created oscillations, but with a minor impact. For these reasons, the genetic algorithm is the optimization procedure in the SR method.

Genetic Algorithm

Created by John Holland (Holland 1975), the genetic algorithm (GA) is a method based on the natural selection theory. The algorithm, as the Darwin's theory of evolution, rounds about genetics, parenthood, reproduction and the survival of the best individuals. The genetic algorithm is integrated in the stochastic reconstruction as the red features in Figure IX-9. In relation to Figure IX-1, it may be considered that the *genetic algorithm optimization* corresponds to the *selection*, *reproduction* and *mutation*, while the *distribution parameters* are the *initialization* and *population*.

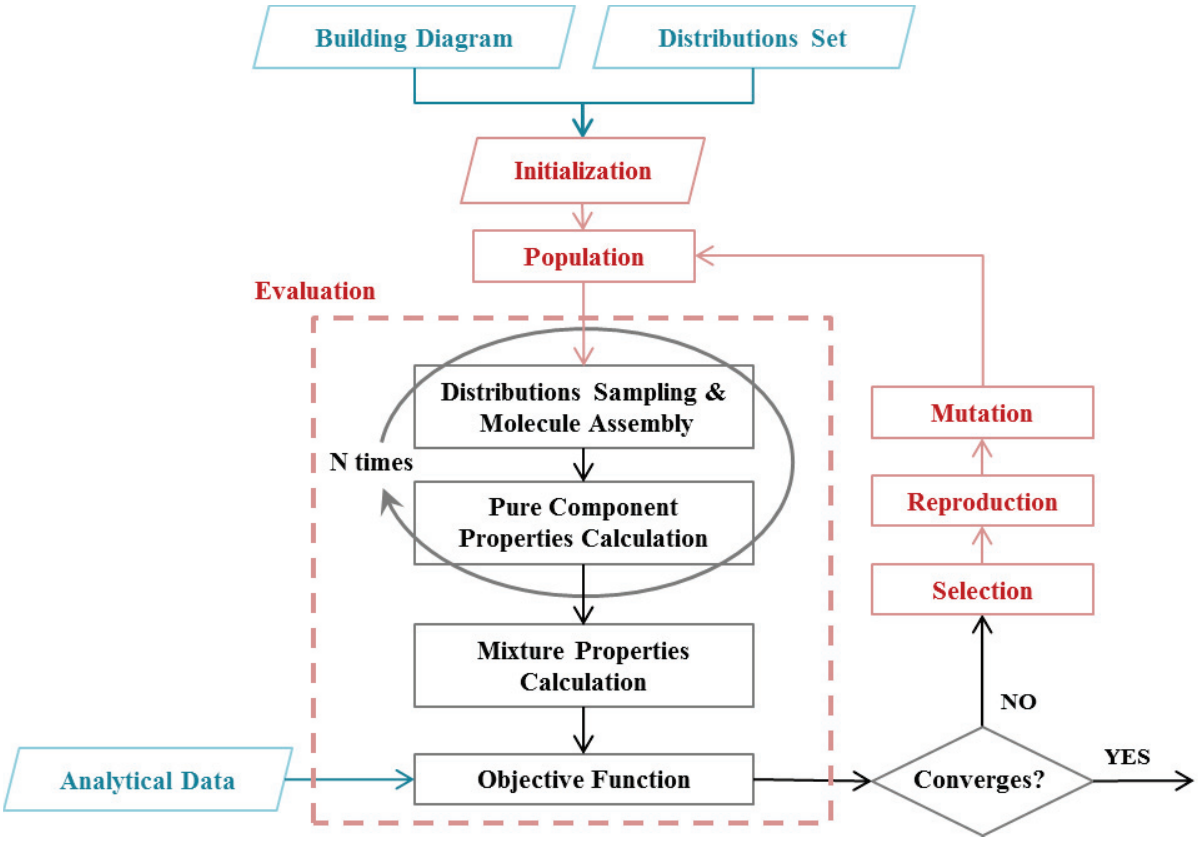


Figure IX-9 – Integration of the genetic algorithm in the stochastic reconstruction diagram.

The genetic algorithm can be divided in six different steps: the initialization of the algorithm (creation of the first individuals), the population (definition of the individuals), evaluation of the individuals by the stochastic method, selection of the best individuals, reproduction of these into new individuals and finally potential mutation of the children. Through these steps, the optimizer is able to estimate the correct set of parameters, which correctly describe the mixture.

The initialization consists in the assembly of all the parameters of all the distributions used in the building diagram. In the population step, the algorithm manages the parameters to form a population of individuals. In detail, each individual is a group of parameters, whose genes are the values of these. Since the number of distributions (and parameters) is limited, the initial population must be expanded in order to form a diversified society, increasing the area of possible solutions. This expansion is done through the random selection of a value restrained between the maximum and the minimum of a distribution parameter.

The generated population is then evaluated through the molecular reconstruction itself, as shown in Figure IX-9. This step is performed by a pre-defined number of molecules and returns the objective function value associated to each individual. In all the range, the individuals with the smaller objective function values are the more suitable to represent the mixture under study. However, obtaining a small objective function may consume to much CPU (Central Process Unit) time. Therefore, the choice of the pre-defined number of molecules is a compromise between the computing time and the representativeness of the equimolar mixture. In order to establish a balance between quality and efficiency, it was considered as optimal value the number of molecules (5000) (Oliveira et al. 2013b).

Similar to the natural selection, the more adapted individuals survive. In this step, individuals with smaller objective function are kept as parents, while the others are discarded, as seen in Figure IX-10. Afterwards, the parents will pass by reproduction, where they are grouped two by two. This procedure is as well effectuated through random selection. The reproduction consists in the exchange of genetic material between genes of the parents, as shown in Figure IX-10. The new individuals, children, are obtained thanks to the crossover points and exchange of information between the parents. To ensure the correct continuation of the genetic algorithm and stochastic reconstruction, the crossover points can never be between two parameters of the same distribution, only between different PDFs. Furthermore, the number of children must always be equal to the number of parents, meaning that the reproduction step produces always other two individuals. It is important to highlight that some parents can be reproduced more than once, while others never.

The last step of the genetic algorithm is the mutation of a new individual. This step does not always occur, but when it does, a given child undergoes an alteration in random selected genes (see Figure IX-10). This modification is done by the Parker technique (Parker 2002), where the parameters undergo a bit inversion. The mutation step is extremely advantageous, as it promotes the expansion of the optimal research area and allows to avoid local minima. In the end, a new population is set to be once again evaluated and optimized if correct set of parameters is not achieved.

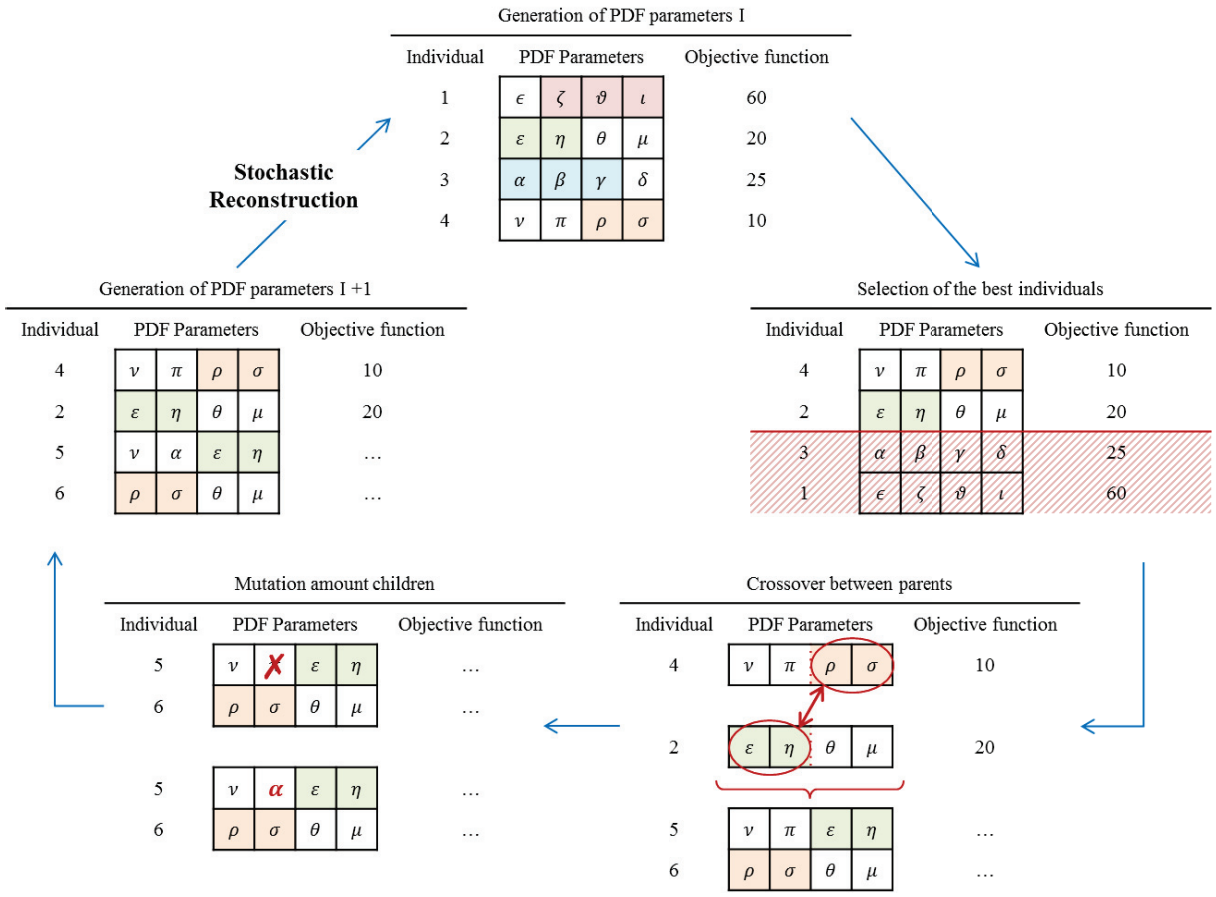


Figure IX-10 - Steps of the genetic algorithm.

To finish, the genetic algorithm will be repeated until the stochastic reconstruction method achieves the minimal number of individuals. Obliviously, this parameter is chosen knowing that the discovering of an optimal solution is a time-consuming process. Similar to this parameter, others as the number of individuals, the number of molecules used for the evaluation of an individual, the number of individuals that are chosen as parents, the number of crossover points, the probability of mutation and the reduction speed of the population must be provided by the user.

The genetic optimization loop, similar to the stochastic reconstruction, is a time-consuming technique. To surpass this drawback, it was added to the algorithm a minimal number of individuals and a reduction speed of the population. Basically, for times to times the number of individuals will be reduced until it achieves the minimal value. In addition to these two parameters, the evaluation of the population was as well modified. To reduce the calculation time of the objective function values, only the new individuals would be evaluated, leaving the old ones (parents) with their former value.

IX.2 Reconstruction by entropy maximization

The reconstruction by maximization of entropy calculates the correct composition for the original set of molecules of the stochastic reconstruction. This method fits the equimolar composition of the mixture

into the correct values by matching the properties of the synthetic and real mixtures. The optimization is performed by LaGrange parameters, who helps to reduce the CPU time required by the method. The present section will explain in further detail the method.

IX.2.1 Theoretical aspects

The reconstruction by maximization of entropy is a method based on the information theory of Shannon (Shannon 1948). This criterion (see Equation IX-4 and Equation IX-5) measures the unknown information in a system. For example, in a mixture where nothing is known, as for example the composition, certain molecules cannot be preferred to others. However, when there is information (e.g. analytical data), certain molecules can be preferred over others, if in line with the given information.

$$H(p) = -\sum_{i=1}^{N}(p_i \cdot \ln p_i) \qquad \textit{Equation IX-4}$$

$$\text{Where, } \sum_{i=1}^{N} p_i = 1 \qquad \textit{Equation IX-5}$$

In summary, and as described by Hudebine and Verstraete (Hudebine and Verstraete 2011), the "(…) Shannon entropy ($H(p)$) is a measure of homogeneity of the probability distribution p_i. The higher the entropy value, the more uniform the distribution is. When the criterion is associated with constrains, maximizing Shannon entropy is equivalent to determining the most uniform distribution which verify these constraints.".

When applied to a mixture of N molecules, the probability distribution p_i is transformed into the molar fractions x_i of molecule i, resulting in Equation I-1. In the absence of analytical data, the method cannot favor one compound over other. In this case the mixture is equimolar and the distribution of the molar fractions is uniform. When analytical information is available, these are added to the method as constraints. Then, the method changes the molar fractions so as to verify these constraints. Although, in this case the distribution of the molar fractions is no longer uniform, this stays as uniform as possible until further information deforms it again.

IX.2.2 Method description

As illustrated in Figure IX-11, the reconstruction be entropy maximization requires analytical data, a library of molecules and mixing rules (see Equation VII-1). The analytical data and the mixing rules are the same as the ones applied in the stochastic reconstruction. In its turn, the library of molecules is the equimolar mixture resultant for the stochastic reconstruction.

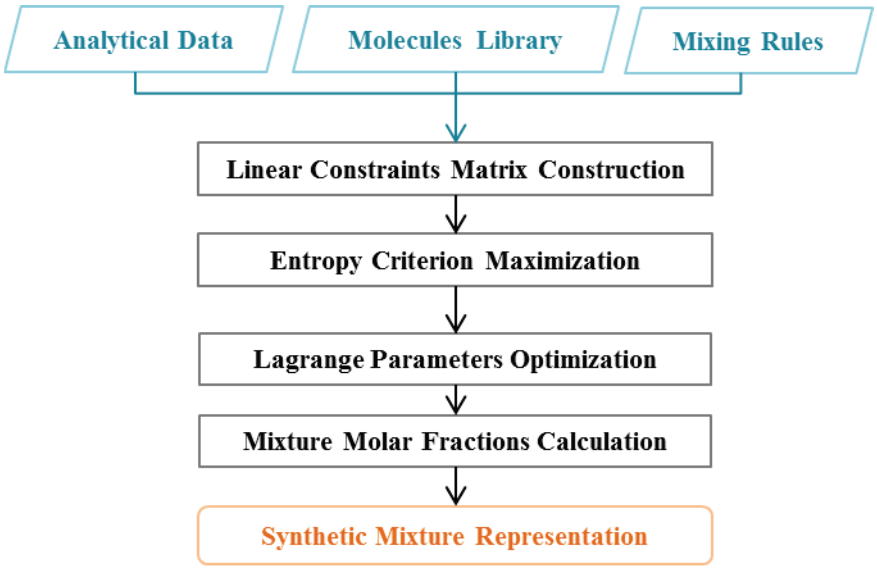


Figure IX-11 – Scheme of the reconstruction by entropy maximization.

The reconstruction by entropy maximization method starts by construction of a matrix with the constraints, obtained thanks to the analytical data and the mixing rules. These constraints can be linear or non-linear. The last restriction depends on innumerous unknowns, which complicates the task for the entropy maximization due to the high dimension of the solution space. In the case of linear constraints, a semi-algebraic solution is possible, together with the decrease of the solution space (unknowns) and the CPU time.

When the linear constrains are introduced in the entropy criterion with Lagrange multipliers, the final expression results in Equation IX-6, where μ and λ_i are Lagrange multipliers, J is the number of constrains, f_j is the equality term of constraint j and $f_{i,j}$ is the parameter of compound i of constraint j. By deriving the previous expression in respect to the molar fraction, Equation IX-7 is obtained.

$$\xi(x) = -\sum_{i=1}^{N}(x_i \cdot \ln x_i) + \mu \cdot \left(1 - \sum_{i=1}^{N} x_i\right) + \sum_{j=1}^{J} \lambda_i \cdot \left(f_j - \sum_{i=1}^{N}(x_i \cdot f_{i,j})\right) \qquad \textit{Equation IX-6}$$

$$x_i = exp\left(-\sum_{j=1}^{J}(\lambda_i \cdot f_{i,j})\right) \Big/ \sum_{i=1}^{N} exp\left(-\sum_{j=1}^{J}(\lambda_i \cdot f_{i,j})\right) \qquad \textit{Equation IX-7}$$

By introducing Equation IX-7 in Equation IX-6, the entropy criterion is now written for J constraints, as shown in Equation IX-8. Maximizing the previous expression is now easier, since the solution space was reduced from $N + J$ unknows (report to Equation IX-6) to J unknows.

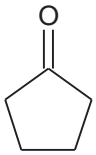
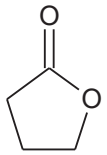
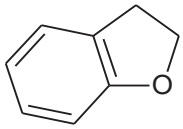
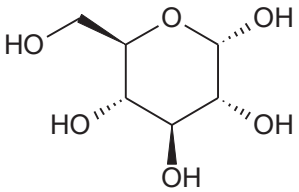
$$\xi(\lambda) = \ln\left(\sum_{i=1}^{N} exp\left(-\sum_{j=1}^{J}(\lambda_i \cdot f_{i,j})\right)\right) + \sum_{j=1}^{J}(\lambda_i \cdot f_j) \qquad \textit{Equation IX-8}$$

The maximization of Equation IX-8 can be done by classic optimization methods, as the conjugate gradient method. After obtaining the values for the Lagrange multipliers, the molar fraction of each compound can be calculated using Equation IX-7. In the end, the initial synthetic mixture will be associated compositions.

IX.3 Stochastic reconstruction adaptation

The adaptation of the stochastic reconstruction method consisted in the addition of atoms and cycles, together with properties features typical from bio-based products, as bio-oil. As mentioned previously, these modifications where done in an internship, thus further information can be found in the respective report (Ribeiro 2018).

Table IX-4 - New molecular structure added to the stochastic reconstruction algorithm.

Type atom/ring	Description	Example	
CO Carbonyl in ring	Carbonyl connected to a naphthenic carbon	Cyclopentanone	
CO Ester in ring	Carbonyl connected to a naphthenic carbon and an oxygen in a ring	Dihydrofuran-2(3H)-one	
O Ester in ring	Oxygen connected to a naphthenic carbonyl and a naphthenic carbon.		
Tetrahydrofuran	Saturated cycle with four carbon atoms and one oxygen	Dihydrobenzofuran	
Tetrahydropyran	Saturated cycle with five carbon atoms and one oxygen	D-Glucose	

Regarding the atoms and cycles addition, the algorithm was extended to reconstruct carbonyls placed in the ring, as cyclopentanone, and ester groups placed in the ring, as dihydrofuran-2(3H)-one, as shown in Table IX-4. Although this last type of functional group is not usual in bio-oils (as lignocellulose degradation products), the combination of between a furanic ring and a carbonyl can be found in the mixture under the form of furanones (Staš et al. 2014). Furthermore, the new cycles were added:

tetrahydrofuran and tetrahydropyran. Both rings are vastly present in the bio-oils, as these structures represent the main skeleton of the sugars from cellulose and hemicellulose.

Regarding the properties, the algorithm was extended to new carbon ranges for the ^{13}C NMR and to a new SEC correlation. Previously, as the main feed used as input were oil fractions, the ^{13}C NMR analysis did not account with the intervals of the spectrum correspond to the oxygenated functional groups. A new range was redefined for the carbon features according to Table I-7 in the State of the Art chapter. The adaptation of the SEC analysis to the stochastic reconstruction was proposed by Bour (Bour 2017). The author created several correlations, which were later on adapted by Ribeiro (Ribeiro 2018) to bio-oil feeds. For the present work, four new expressions were proposed. For all, Equation IX-9 was the most suitable to predict the polystyrene equivalent molecular weight ($Mw_{eq}PS$) from the molecular weight (Mw) of the reconstructed molecules.

$$Mw_{eq}PS = \left[950 \cdot arsin\left(0.64 \cdot \frac{Mw}{1000}\right) + 450\right] \cdot \frac{0.64 \cdot Mw}{1000} + 88 \qquad \textit{Equation IX-9}$$

Besides the ^{13}C NMR analysis, it was added to the algorithm a feature of direct inspection of the structure designated of oxygenated functional groups. Although this feature is not a real property and cannot be measure, it helps to have an insight in the chemical detailed of the reconstructed molecules.

IX.4 Reconstruction by entropy maximization adaptation

The adaptation of the reconstruction by entropy maximization method consisted in the update of the properties used for the optimization of the molecular fractions. From the three features mentioned previously, only the ^{13}C NMR analysis is used in this algorithm. However, as the properties sources are shared between the stochastic reconstruction and this method, there is no need to redefine once more the intervals of the ^{13}C NMR. Regarding the SEC analysis, it was required to extend the matrix of constrains. After included as a restriction, of the reconstruction by entropy maximization method could be correctly to bio-oil feeds.

IX.5 Conclusion of Chapter IX

In this chapter, the molecular reconstruction method, selected in Objectives and strategy chapter, was presented. This method consists in the coupling of a stochastic reconstruction method and a reconstruction by entropy maximization. The first generates an equimolar library of molecules which is in accordance with the experimental data used for the reconstruction. The second method calculates the molar fraction of the equimolar mixture also thanks to the experimental data. In the end, it is possible to obtain a synthetic mixture representative of the real feedstock.

In the present work, this methodology was applied to bio-oil mixtures. Therefore, certain methods required some modifications when compared to the methods dedicated to the oil fractions. As will be seen in the next chapter, to apply the SR and REM techniques, new molecular structures must be added, together with the redefinition of the ^{13}C NMR families, the adjustment of the SEC correlation to bio-oil and finally the creation of a feature designed for the account of oxygen families.

Chapter X. Molecular reconstruction results

In collaboration with Rita Ribeiro

As previously mentioned, both molecular reconstruction techniques were already applied and validated for heavy oil fractions. When applied to bio-oil, both methods may need certain adjustments, as seen before. This section will include a brief remark on typical molecules of bio-oil, the experimental information used in both algorithms and the results of the molecular reconstruction, passing through two different reconstruction building diagrams, created by Ribeiro (Ribeiro 2018).

X.1 Chemical structure characteristics

As seen in the State of the Art chapter, bio-oil is a complex oxygenated mixture. Its molecular diversity and high content of oxygen weakens its characterization. However, it is still possible to retrieve some information regarding the physical properties and chemical structure of bio-oils. The former was already detailed in the State of the Art chapter and further information will be given in the next section.

X.1.1 Functional groups

Concerning the molecules structures, it is known that bio-oil contains a high diversity of oxygenated molecules, due to the variety of oxygenated functional groups (FG) present in them (Milne et al. 1997; Staš et al. 2014). Between hydroxyls, carbonyls (ketones and aldehydes), carboxyls, there is also ethers in the form of methoxyls and esters in the form of furanones (Staš et al. 2014). Any other type of ethers or esters are unusual and represent a small fraction of the compounds present in bio-oil (Milne et al. 1997; Staš et al. 2014). Besides the functional groups, bio-oils are also characterized by heterocycles, where the main heteroatom is oxygen. These rings, as furan and tetrahydropyran (THP), are a main feature of the degradation of hemicellulose and cellulose. Apart from these, bio-oils account as well with benzene rings derivative from lignin degradation. Similar to ethers and esters, saturated rings (apart from tetrahydropyran) are not usual as well as peroxides (Milne et al. 1997; Staš et al. 2014). Even so these can be generated by undesirable reactions of bio-oil during storage and transportation, meaning that the hydrotreating process should be prepared for such. As an initial reconstruction of bio-oil, the present work will not account with these structures (excluding methoxyls and furanones). Figure X-1 summarizes some examples of bio-oil molecules.

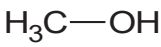
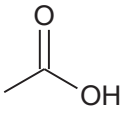
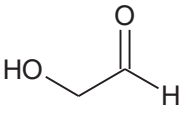
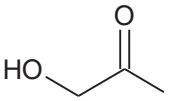
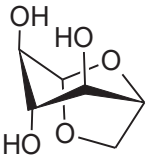
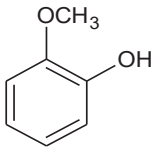
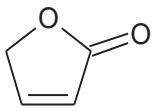
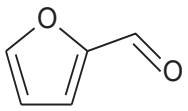
			
Methanol	Acetic acid	Glycolaldehyde	Acetol
			
Levoglucosan	Guaiacol	γ-crotonolacton	Furfural

Figure X-1 - Examples of bio-oil molecules with different functional groups and rings (Staš et al. 2014).

X.1.2 Dimensions

Besides the functional groups, bio-oil molecules may present different sizes, hence different molecular weights. There are light molecular weight species, mostly originated from cellulose and hemicellulose, and high molecular weight compounds, derivate from lignin. On one hand, the formers are mainly composed by linear molecules characterized by carbonyl, carboxyl and hydroxyl functional groups. Monomers compounds, as furans, sugars and phenolics, can be also included in the light species. Although these molecules can also be characterized by carbonyl, carboxyl, hydroxyl and methoxyls functional groups (in this case), the presence of a ring is mandatory. May these be aromatic, saturated or heterocycles, the rings usually have sidechains with oxygenated functions as terminal groups. On the other one hand, and according to Wang et al. (Wang et al. 2013), "high molecular weight compounds are primarily oligomers of phenolic compounds with molecular weights ranging from several hundred too much as 5000 or more.". According to Bayerbach and Meier (Bayerbach and Meier 2009), this heavy fraction may contain dimers, trimers, tetramers and so on. These species, like the monomers, also have sidechains with oxygenated functions as terminal groups (Bayerbach and Meier 2009). Either in the oligomers and monomers, or in the linear molecules (paraffins), the length of the chains is never large. Usually, these comprise around about five or less atoms (Bayerbach and Meier 2009; Milne et al. 1997; Staš et al. 2014). Figure X-1 and Figure X-2 present some examples of potential bio-oil molecules.

As mentioned before, furan and tetrahydropyran rings are usual in bio-oil. However, tetrahydrofuran (THF) is not. Either way, it is quite feasible to find tetrahydrofurans in pyrolytic lignin of bio-oils. This type of ring is normally in a group, as for 1H,3H-Furo[3,4-c]furan, tetrahydro- (two THFs) and benzofuran (one THF and one benzene), but can also be found alone. Both possibilities were seen by Bayerbach and Meier (Bayerbach and Meier 2009), as seen in Figure X-2, and also by Xu et al. (Xu et al. 2014) in Figure I-34 in the State of the Art chapter.

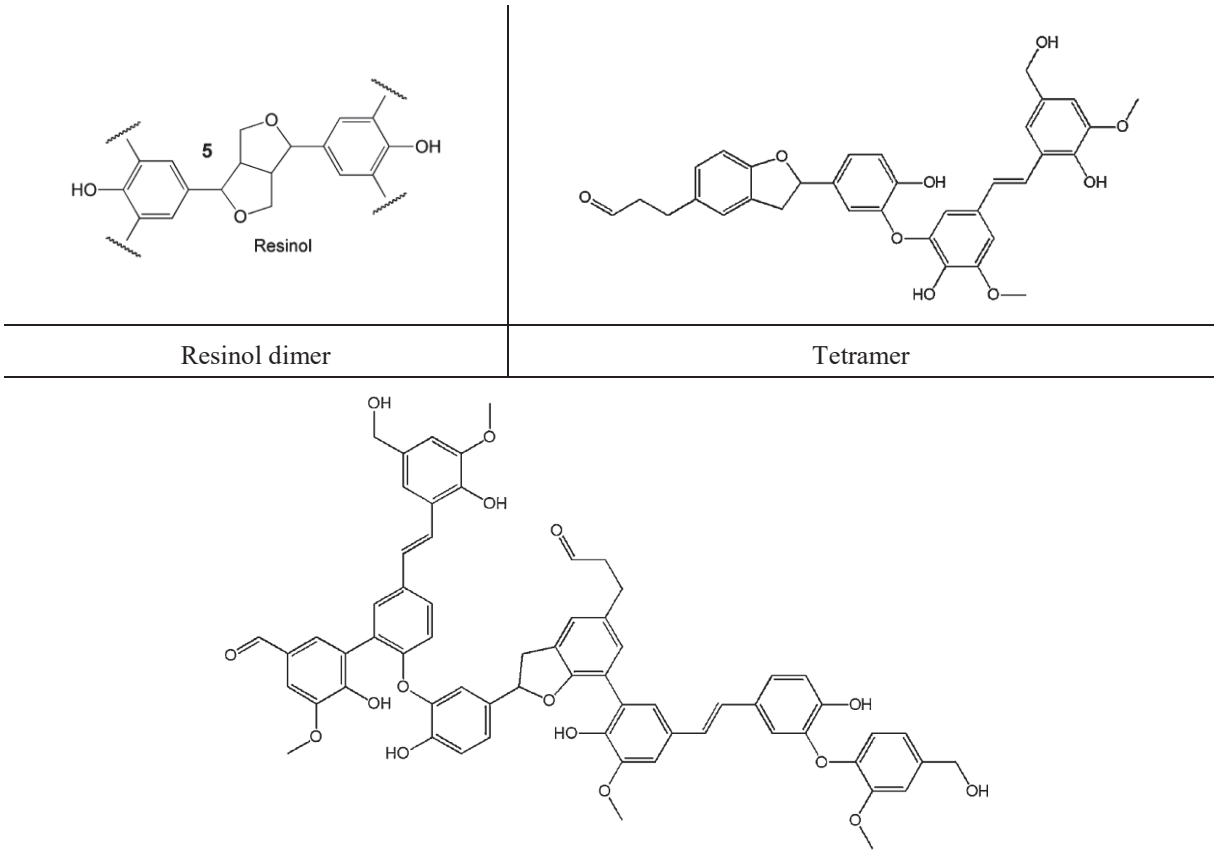


Figure X-2 - Examples of bio-oil molecules with different sizes (Bayerbach and Meier 2009).

X.1.3 Elements ratio

According to Wang et al. (Wang et al. 2013), "(...) oxygen [is present] in almost every bio-oil component (…)". On a dry basis, the ration oxygen/carbon (O/C) and hydrogen/carbon (H/C) vary as shown in Figure X-3. From the picture, it is possible. To conclude that bio-oil molecules have an O/C and H/C ratio between $0.1 \leq O/C \leq 1$ and $0.2 \leq H/C \leq 2$, respectively.

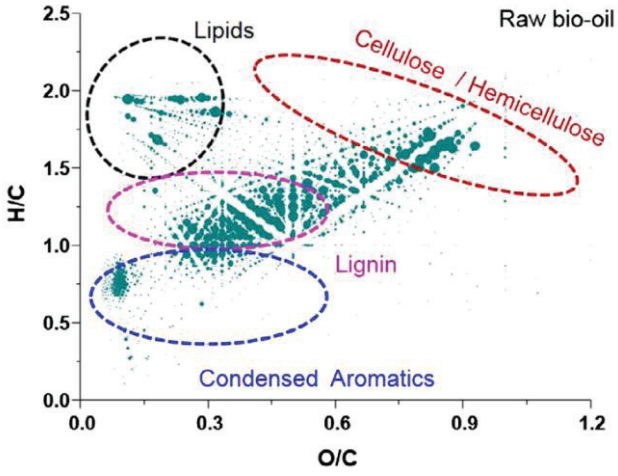


Figure X-3 - Van Krevelen diagrams of the bio-oil from red pine fast pyrolysis (Cheng et al. 2016).

X.2 Experimental data

For the reconstruction of bio-oil mixture, two samples were used. Both come from VTT Technical Research Centre of Finland Ltd. As mentioned in Experimental approach section, the samples were obtained in 2013 through fast pyrolysis of lignocellulosic biomass (forest residues). The original feed underwent fast pyrolysis at 480 °C with a residence time inferior to 1 s.

Both bio-oil samples, here designated as sample A and sample B, were analyzed either by the VTT Technical Research Centre of Finland Ltd, or by the Analytical Department of IFPEn. As presents in Table X-1, the density at 20 °C, the viscosity at 40 °C, the moisture (water) content, the elemental analysis (CHONS) in a wet basis and the nuclear magnetic resonance (NMR) of carbon 13 are the analyses available for both bio-oil samples.

Table X-1 - Available experimental data for bio-oil sample A and B.

Property		**Sample A**	**Sample B**
Density (20 °C) (g/cm^3)		1.24	1.19
Viscosity (40 °C) (mm^2/s)		31.44	14.96
Water (w%)		21.70	22.80
Elemental analysis (wet basis) (%)	Carbon	41.50	41.40
	Hydrogen	7.41	7.64
	Oxygen	49.94	50.29
	Nitrogen	0.04	0.30
	Sulfur	0.00	0.00
^{13}C NMR (%)	Aldehydes & Ketones (215-180ppm)	6.20	5.90
	Carboxylic acids & Esters (180-165 ppm)	4.10	6.70
	Aromatic C-O (165-142 ppm)	8.80	8.50
	Aromatic C-C (142-125 ppm)	3.70	4.20
	Aromatic C-H (125-95 ppm)	17.70	18.50
	Hydroxyls, Ethers, Phenols, Methoxys, Sugars (95-55 ppm)	41.60	30.70
	Aliphatic carbons: branched and long chain (55-28 ppm)	7.10	10.60
	Aliphatic carbons: short chain (28-8ppm)	10.80	14.90

Besides these analyses, the samples were also analyzed by size exclusion chromatography (SEC) and Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR/MS). This last analysis used as ionization negative electrospray ionization (ESI-), negative and positive atmospheric pressure chemical

ionization (APCI). The first mode in the most applied to bio-oil, and as consequence, will be the one used in the present thesis (Abdelnur et al. 2013; Miettinen et al. 2017; Smith et al. 2012). The experimental data regarding the SEC analysis can be found in Figure X-4 for both samples, while the *OxNySz* families obtained through FT-ICR/MS are shown in Figure X-5 likewise for both bio-oil samples.

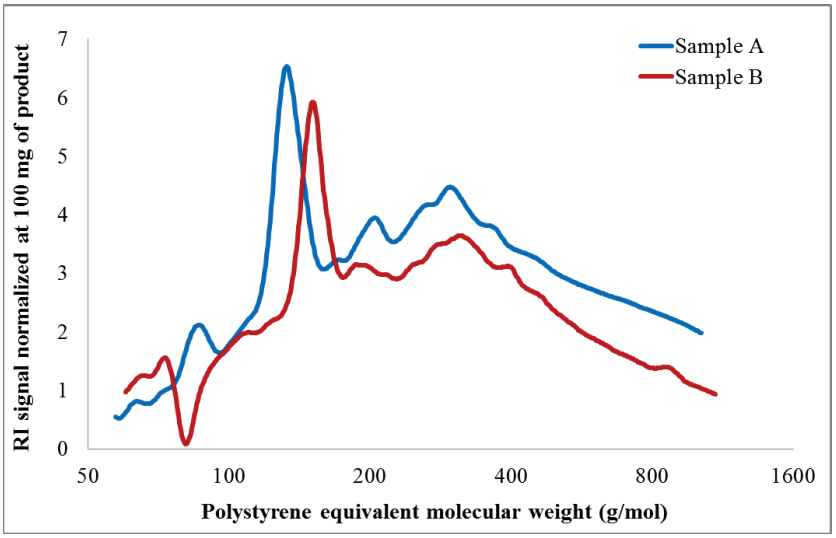


Figure X-4 - Distribution of molecular weights (polystyrene equivalent) of bio-oil samples A and B by size exclusion chromatography hyphenated with an index refractive detector.

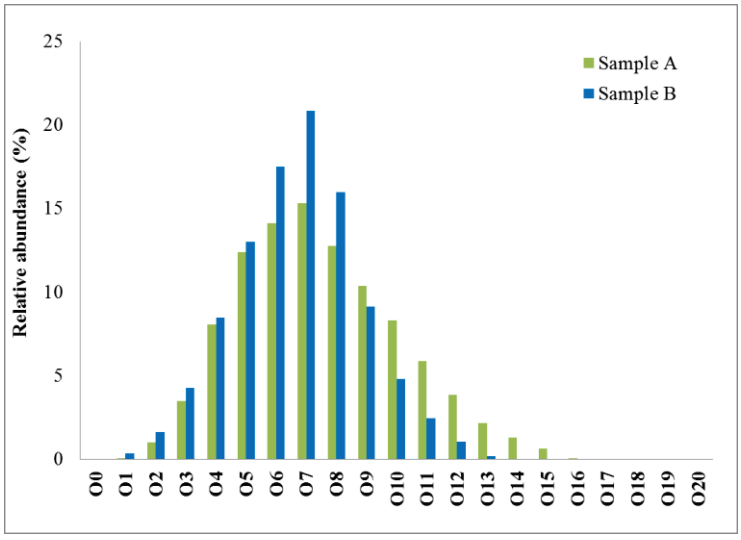


Figure X-5 - FT-ICR/MS analysis(%mol/mol) with ESI- (OxNySz families) for bio-oil samples A and B.

To reconstruct a mixture, the stochastic reconstruction method requires at least two properties. On the one hand, the algorithm needs information regarding the functional groups of the mixture. On the other one hand, the method must know the size of the molecules. As consequence, ^{13}C NMR analysis and the SEC analysis were used for the creation of the compositional model. Besides these analyses, the algorithm also used the elemental analysis to know the atoms content in the mixture.

The data of the NMR was grouped into saturated, aromatic, carbonyls, carboxyls and hydroxyl (methoxyl and phenols) bonds, as shown in Table X-2. Finally, the SEC analysis was applied by subdividing into several fractions. These are later on accumulated, producing the distributions seen in Figure X-6. It is important to note that the first two points and the last (0.0, 0.5 and 1.0) are not in the distributions, since they do not have a meaning in the analysis. The data from the elemental analysis was first transformed into a dry basis, because water molecules will not be reconstructed by the algorithm. This conversion was based on the moisture content, also given as experimental data (see Table X-1). Only then, these values were used as inputs.

Table X-2 - Available experimental data used as inputs for bio-oil sample A and B.

Property		**Sample A**	**Sample B**
Elemental analysis (dry basis) (%)	Carbon	53.79	54.12
	Hydrogen	6.48	6.67
	Oxygen	39.73	39.21
	Nitrogen	0.00	0.00
	Sulfur	0.00	0.00
^{13}C NMR (%)	Saturated	17.90	25.50
	Aromatics	30.20	31.20
	Aldehydes & Ketones	6.20	5.90
	Carboxylic acids & Esters	4.10	6.70
	Hydroxyls, Ethers, Methoxyls	41.60	30.70

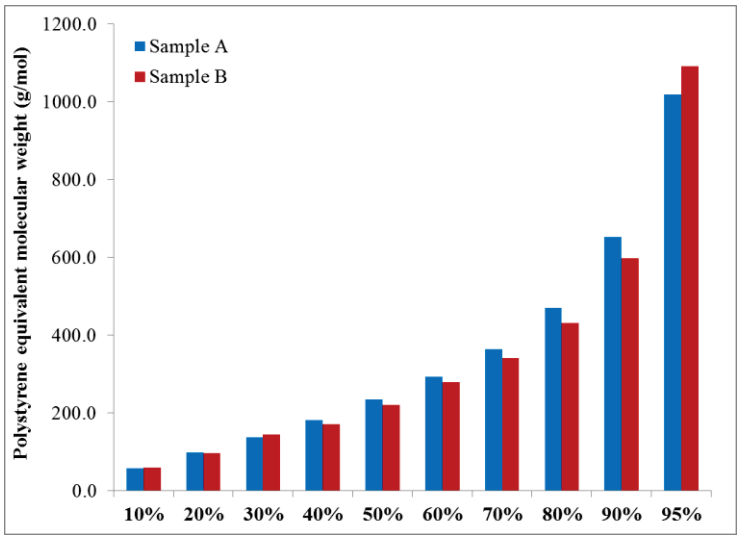


Figure X-6 - Cumulative distribution of molecular weights (polystyrene equivalent) of bio-oil samples A and B by size exclusion chromatography hyphenated with an index refractive detector

The validation of the reconstructed mixture was done by FT-ICR/MS analysis and density. The last two analyses were not adapted for bio-oil, because it was not required by the algorithm.

X.3 Molecular library results

The following section will present the results obtained for the molecular reconstruction of bio-oil sample A and B. As mentioned before, the present work was mainly done by Ribeiro (Ribeiro 2018). The author proposed several building diagrams (10) for the reconstruction of bio-oil sample A. From all, only the two bests will be presented and discussed.

Before advancing further, it is important to highlight that any of the building diagrams proposed by Ribeiro (Ribeiro 2018) included the possibility to generate olefins or isomers. The difficult associated to the characterization of these is the main reason for their absence in the building diagram (even if the algorithm is prepared). Furthermore, the diagrams that will be presented do not include as well the possibility to create esters (apart from furanones), ethers (apart of methoxyls) and peroxides. As mentioned before, these components are not usual in bio-oil mixtures. Therefore, there is no reason to reconstruct them.

In conclusion, the same inputs were used for all the reconstructed mixtures. These parameters were already optimized for complex mixtures by Oliveira ((Oliveira 2013)) in the case of light cycle oils and vacuum residues. Due to the diversity and dimensions of the bio-oil samples, the same parameters used for vacuum residues were applied ((Oliveira 2013) – page 251). Moreover, a library of ten thousand molecules was generated for each sample and each building diagram here presented. The number of molecules was determinate according to the stochastic simulation algorithm. As discussed in the State of the Art chapter, the higher the number of molecules the better the method performance. The number of molecules required was also investigated by Oliveira (Oliveira 2013).

X.3.1 First hypothesis of reconstruction

Here, it will be presented the first hypothesis of reconstruction (BD#1). Furthermore, the present section includes the results for both bio-oil samples.

X.3.1.1 Building diagram

As shown in Figure X-7, the first building diagram here presented starts with the sampling of the ***number of cores***. As previously seen, bio-oil molecules may be linear or cyclic with one or more cores. The number of cores is never too large, except for high molecular compounds as hexamers, heptamers and etc. In this building diagram, it was assumed that those molecules would exist in low quantity in the mixture. Therefore, the probability of reconstruction of these should be low as well. In summary, it was concluded that an exponential distribution would be most suitable distribution.

The previous distribution gives the number of cores in the molecules. When this value hits zero, the molecule is linear (paraffin). In this case, the algorithm will follow for the next distribution: l***ength of***

the chains. This function indicates how many atoms exist in a chain. Because several types of paraffins can be found in bio-oil, it was assumed that the length of paraffins would be described by a gamma distribution.

According to the number of cores sampled, the algorithm will select the ***number of rings per core***. The distribution regarding the number of rings is a histogram with three options: one ring, two rings and three rings. As seen before, it is unusual for bio-oil molecules, and even macromolecules, to have more than three rings per core. For this reason, it was assumed that the number of rings could be described by a histogram distribution.

As shown in Figure X-7, after having selected the number of rings per core, the algorithm samples the ***type of ring*** for each sampled cycle. As presented before, bio-oil molecules are mostly composed by benzene, furan and tetrahydropyran (THP) rings. However, in the macromolecules, it is possible to see tetrahydrofuran (THF) rings assembled together or with benzene rings. For this reason, it was assumed that the type of rings could be described by a histogram distribution with four options: benzene, furan, THF and THP.

Usually cyclic bio-oil molecules have one or more side chains. Therefore, building diagram contains a distribution for the sampling of the ***additions of sidechains*** (addition of alkyl chains). This distribution is a "yes or no question", where yes implies the addition of a sidechain. This sampling process occurs for all the atoms in the rings that can be substituted. Technically, the algorithm goes every atom to atom in the rings that can be substituted, samples a yes or a no, and if it is a yes, adds a side chain. As probably expected, the nature of this distribution can only be described by a histogram.

For each atom of the ring with a sidechain, the algorithm will select the ***type of chain***. From the examples presented for the bio-oil molecules (see Figure X-1 and Figure X-2), it is possible to conclude that sidechains are either methoxyls, or methyls or alkyl chains. Hence, it was assumed that the type of chains could be described by a histogram distribution with three options: methoxyls, methyls and alkyl chains. For the last case, alkyl chains, the algorithm samples once again the ***length of the chains***. The distribution for this attribute is the same as for the paraffins. Although some differences may be found between the size of sidechains and paraffins (Staš et al. 2014), this approximation was suitable for the description of heavy oil fractions (Oliveira 2013).

As shown in Figure X-7, after defining the core of the molecules and their ramification, the algorithm selected the heteroatoms associated with the molecules. For the reconstruction of bio-oil molecules, it was assumed that the compounds would not have nitrogen or sulfur, only oxygen. Since there are several oxygenated functional groups in bio-oils (see X.1), this section was divided in three distributions. For all paraffin and cyclic molecules with methyl or alkyl sidechains, the algorithm will sample the probability of ***replacing a carbon atom by a carbonyl group*** (addition of carbonyl to CH_2). Similar to

the addition of chains, the algorithm will go through all carbons (CH_2) which can be substituted by a carbonyl (>C=O), sampling a yes or no answer and if it is a yes, will replace the atom by a carbonyl. Once again, the nature of this distribution can only be described by a histogram.

After the addition of carbonyls, the algorithm will sample the potential ***substitution of terminal groups*** (substitution of CH_3). Also similar to the addition of chains, this distribution is a "yes or no question", where yes implies the replacement of the terminal group. This sampling process occurs for all the terminal methyl atoms in the molecules. When a yes is sampled the algorithm will sample what type of oxygen terminal atom. Once again, the nature of this distribution can only be described by a histogram.

The sampling of the ***type of substituent*** is done through a histogram function with three options: hydroxyl, aldehyde and carboxyl. As seen before, these are the three main type of termination groups in bio-oil molecules. Technically, of each methyl group that is to be replaced, the algorithm will select among a hydroxyl (OH), an aldehyde (COH) and a carboxylic (COOH), which oxygenated group will be added. After replacing all methyls, the reconstruction of the molecule is given as concluded and the algorithm will pass for the reconstruction of another.

The molecular attributes highlighted before, together with their probability function distributions and their respective number of parameters are summarized in Table X-3. In summary, the building diagram here proposed, and present in Figure X-7, has nine distributions and fourteen parameters to be estimated.

Table X-3 - Molecular attributes and the respective PDF for building diagram 1.

PDF Number	Molecular attribute	PDF Type	Parameters Number
0	Number of cores	Exponential	1
1	Length of chains	Gamma	1
2	Number of rings	Histogram	2
3	Type of ring	Histogram	3
4	Addition of alkyl chains	Histogram	1
5	Type of chains	Histogram	2
6	Addition of carbonyl to CH_2	Histogram	1
7	Substitute a CH_3	Histogram	1
8	Type of substituent	Histogram	2

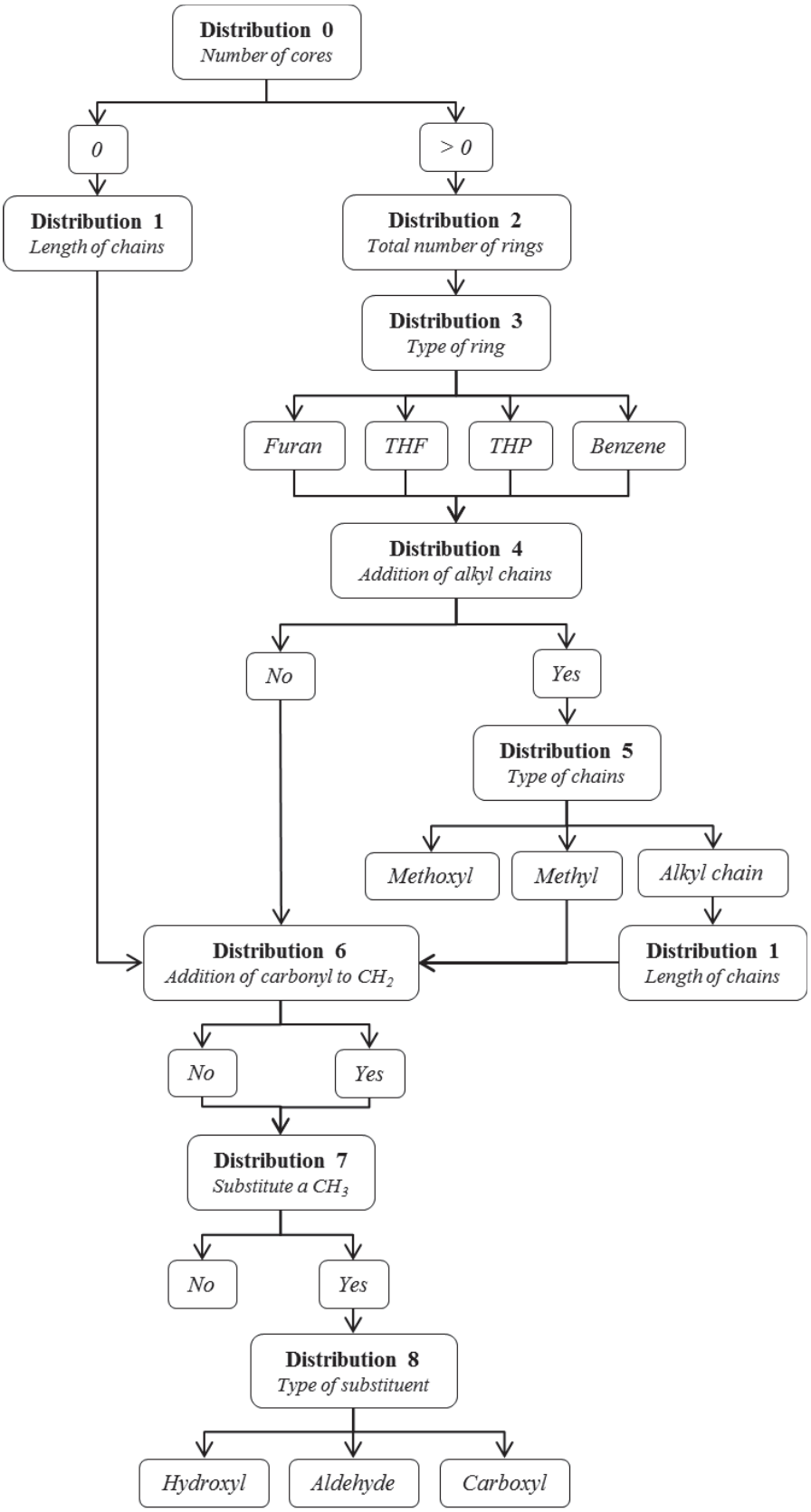


Figure X-7 - Diagram number 1 for the reconstruction of bio-oils.

X.3.1.2 Stochastic reconstruction

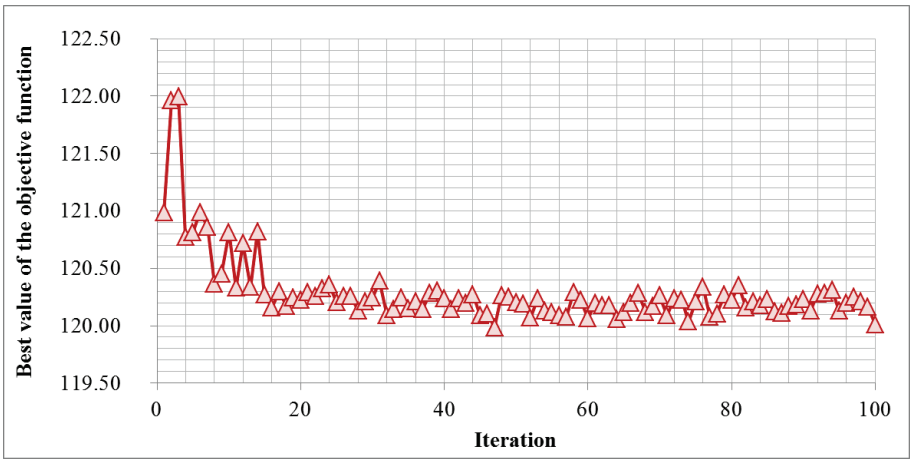
For sample A and sample B, a library of molecules was generated through building diagram 1. Here, the results will be presented and discussed. These were divided into five sections: Objective function, probability distribution diagrams, estimated properties, predicted properties (those not used in the fitting of the parameters) and finally the molecular structures.

Objective function

The objective function (OF) used by the stochastic reconstruction algorithm is present in Equation IX-2. For building diagram #1, the minimum objective function value was of 119.98 and 108.75 for sample A and sample B, respectively. The values are far from the desired result (zero). However, these were obtained after a series of two individual optimizations processes. As it happens, the SR has applied to both samples, originating a set of optimized parameters (1st optimization step). Afterwards, these parameters were implemented as input values and the SR was once again applied to both samples (2nd optimization step). This procedure intended to force the reduction of the objective function. However, and as shown in Figure X-8 (2nd optimization step results), after 100 iterations the OF barely decreased in both cases. Furthermore, after iteration 20, for both samples, the OF value is practically constant, meaning that this can hardly decrease furthermore. Thus, while using BD#1 and the experimental sets of sample A and sample B, there is no need to push further the OF.

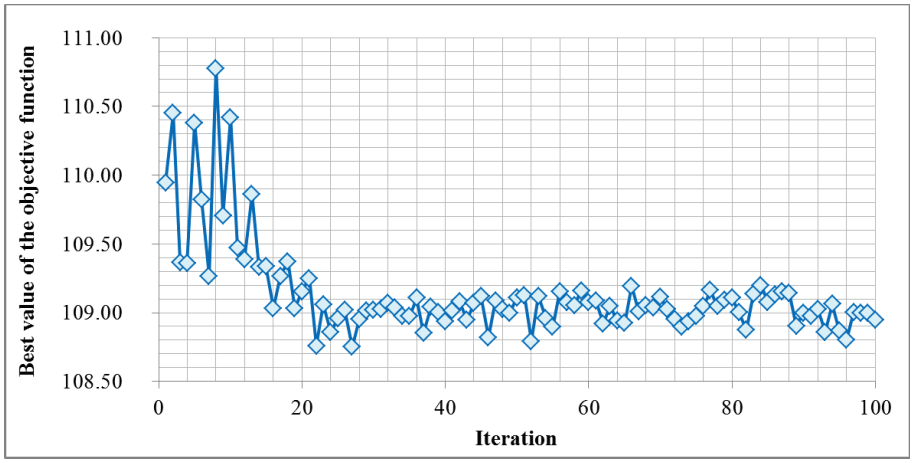
Objective function

Sample A



a)

Sample B



b)

Figure X-8 - Evolution of the nest value of the objective function with the iterations: a) Sample A; b) Sample B.

Probability distribution functions

For the previous reason, it was considered that the parameters obtained after the 2nd optimization run were the optimum values. These are present in Table X-4, together with their respective molecular attribute and probability distribution function (PDF). For a better analysis of the optimized parameters, the results of their discretized PDF are illustrated in Figure X-9.

Table X-4 - Molecular attributes and the respective PDF and optimized parameters for building diagram 1.

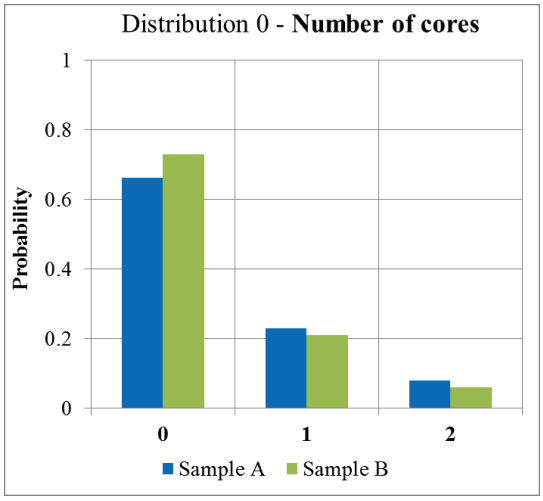
PDF Number	Molecular attribute	PDF Type	Parameters Number	Optimized parameter values	
				Sample A	Sample B
0	Number of cores	Exponential	1	0.938	0.794
1	Length of chains	Gamma	1	1.001	1.002
2	Number of rings	Histogram[1]	2	0.871	0.579
				0.103	0.349
				0.026	0.071
3	Type of ring	Histogram	3	0.874	0.933
				0.106	0.044
				0.005	0.014
				0.015	0.010
4	Addition of alkyl chains	Histogram	1	0.314	0.452
				0.686	0.548
5	Type of chains	Histogram	2	0.061	0.116
				0.937	0.883
				0.002	0.001

[1] In histograms, the extra parameter (*) is always obtained through difference (1 – the other parameters).

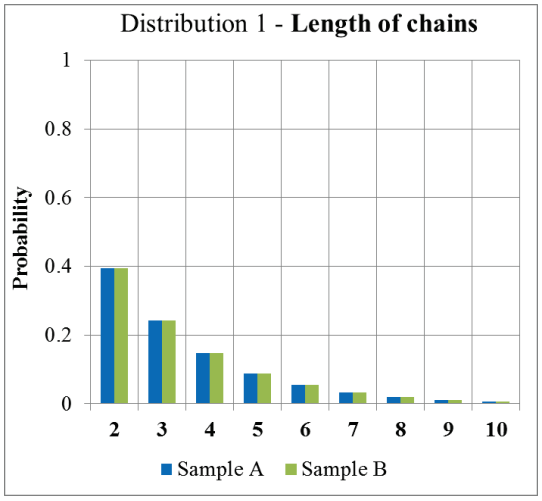
6	Addition of carbonyl to CH_2	Histogram	1	0.868	0.940
				0.132	0.060
7	Substitute a CH_3	Histogram	1	0.000	0.003
				1.000	0.997
8	Type of substituent	Histogram	2	0.746	0.687
				0.153	0.148
				0.101	0.165

As seen in Figure X-9, the PDFs follow the same trends for both samples. Even so, there is certain differences between both (also seen in Table X-4). These are originated from the distinct experimental data of sample A and sample B used by the SR for the estimation of the parameters.

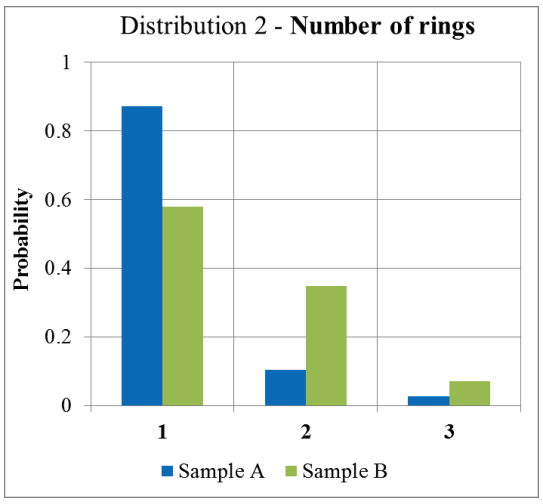
Building diagram # 1 - Sample A and Sample B



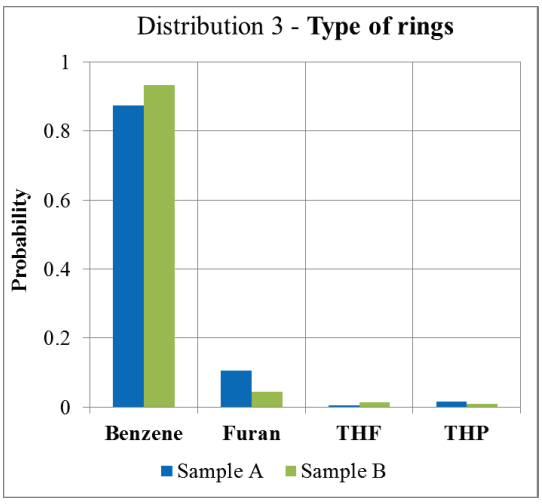
a)



b)



c)



d)

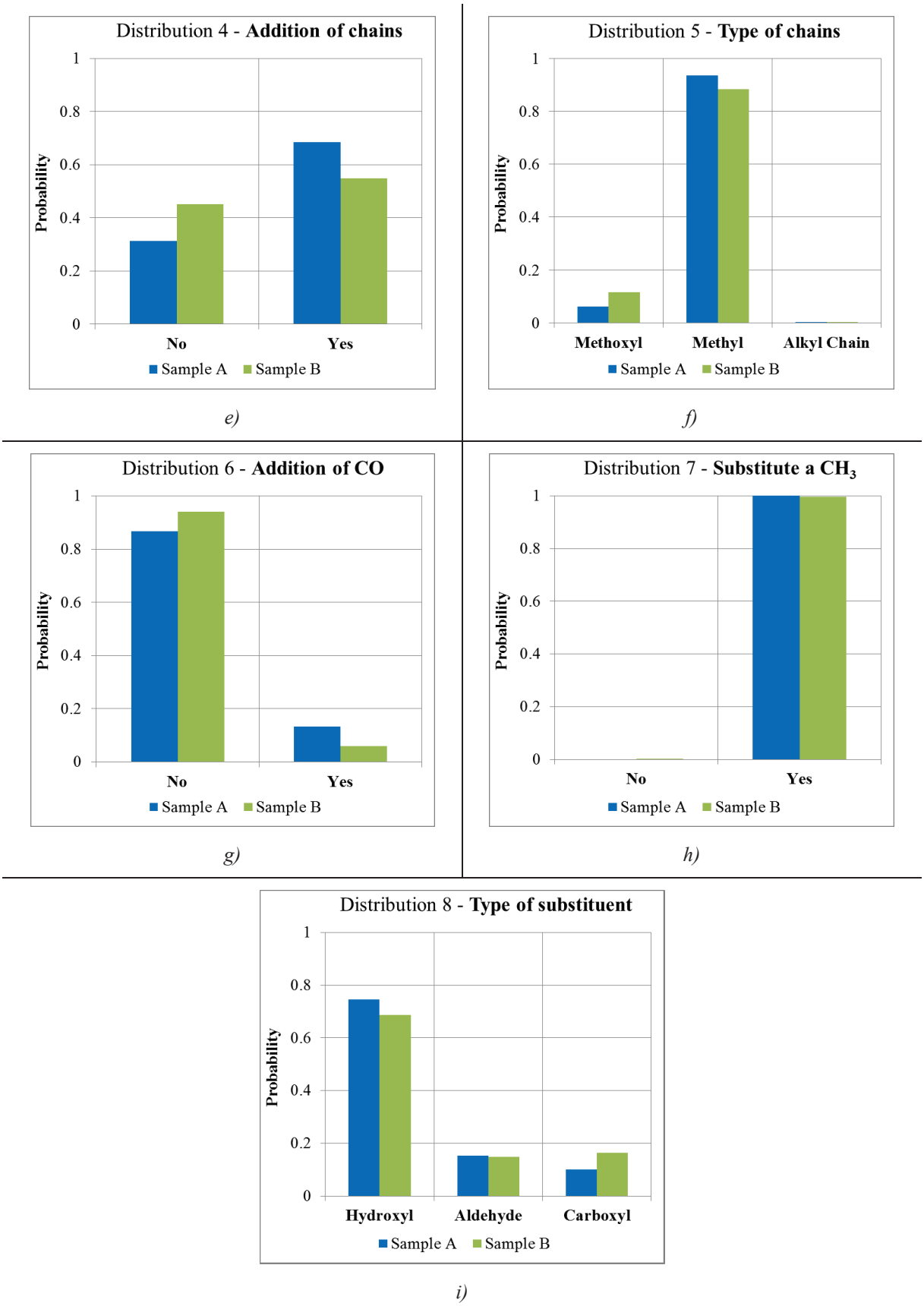


Figure X-9 - Resultant probability distribution functions of building diagram #1 for sample A and B: a) Distribution 0 – Number of cores; b) Distribution 1 – Length of chains; c) Distribution 2 – Total number of rings; d) Distribution 3 – Type of

ring; e) Distribution 4 – Addition of alkyl chains; f) Distribution 5 – Type of chain; g) Distribution 6 – Addition of a carbonyl to a CH_2; h) Distribution 7 – Substitute a CH_3; i) Distribution 8 – Type of substituent.

For both samples, the ***number of cores*** presents similar results. The algorithm predicts three cores as the maximum for the molecules. In both cases, there is a higher probability of generation linear molecules (paraffins), than cyclic compounds. Amongst these, monomer molecules are also more probable than dimers. Furthermore, sample B has a higher probability to create paraffins than sample A. This may justify by the higher content of saturated carbons in sample B than sample A. In summary, the algorithm reconstructs mainly *n*-paraffins, follow by monomers and dimers. Although these structures can be found in bio-oil, this type of mixtures has as well other much larger molecules (Bayerbach and Meier 2009).

For the ***length of chains***, both samples share an identical behavior. Chains with two atoms are the most probably to be reconstructed, followed by chains with three and four atoms. Larger chains with five or more atoms can also be generated, but their probability is lower than the formers. In summary, the algorithm reconstructed more paraffins with two, three and four atoms and rings with sidechains equally small. These type of structures are expected for typical molecules of bio-oil (Bayerbach and Meier 2009) (Staš et al. 2014).

Concerning the ***number of rings***, the probability of most cores only presenting a ring is higher than the probability of having two or three rings. Be that as it may, sample B has a higher probability than sample A regarding the generation of two or three rings per cores. This difference can be justified by the SEC experimental results, which state that sample B has more high molecular weight compounds than sample A (see Figure X-6). To conclude, the algorithm reconstructs mainly cores (monomers and dimers) with a single ring and occasionally with two rings. These molecular structures, with one/two rings per core, are in agreement Bayerbach and Meier (Bayerbach and Meier 2009) (see Figure X-2).

For each ring, the most probable type is benzene, followed by furan, then THF and THP. These last three ***types of ring*** have a much lower probability than benzene. Although the amount of the several types of rings is not known, it is expected a higher amount of benzene, due to the amount of pyrolytic lignin present in bio-oil (report to Figure I-2 in the State of the Art chapter). In sum, the algorithm reconstructs mostly benzene rings, which is in agreement with the structures of bio-oil (Bayerbach and Meier 2009) (Staš et al. 2014).

In the case of the ***addition of chains***, both samples have a higher probability of accepting sidechains. Even so there is a difference in the probability. As mentioned before, such deviations are caused by the experimental data, which is not the same for both samples. For this case, the sample B has less hydroxyls and methoxyls groups than sample A. Therefore, by reducing the number of sidechains in the rings, the algorithm can also reduce the amount of these two groups. In summary, the algorithm generates mostly rings with sidechains. As seen in before in Figure X-1 and Figure X-2, the structure reconstructed by

the algorithm are within the usual representation of bio-oil molecules (Bayerbach and Meier 2009) (Staš et al. 2014).

Similar to the rings, for each sidechain, there is a ***type of chain***. For both samples, the methyl sidechains are the most probable type, followed by methoxyls and then alkyl chains. These last types have a much lower probability than the methyls, in particularly the alkyl chains which are practically impossible to generate, as its probability is extremely low. Although sample B has less hydroxyls and methoxyls groups than sample A, as mentioned before, the algorithm predicts a higher probability for selecting methoxyls for sample B than sample A. One of the reasons maybe that instead of sampling too much methyls, which will eventually become hydroxyls (also in lower quantity in sample B), the algorithm prefers to occasionally sample methoxyls. In this way the number of hydroxyls will be kept low and number of saturated carbons will increase (in higher quantity in sample B), due to the presence of the CH_3 in the methoxyl group. In summary, the algorithm will mostly generate rings with methyl sidechains, which may eventually become hydroxyls. When compare with the expected molecules of bio-oil, these results are in accordance. However, it is also quite possible to find molecules with one or more methoxyls or even alkyls chains in bio-oil mixtures (Wang et al. 2013) (Bayerbach and Meier 2009).

Regarding the heteroatom groups, both samples have more probability to not add a carbonyl group. Also, the ***addition of carbonyl*** (CO) varies between both samples. A higher probability of adding a CO is seen in sample A. As it happens, this sample has a higher number of aldehydes and ketones, thus the probability for the sampling of the yes is higher for sample A than sample B. In summary, the algorithm generates mostly chains without or with few carbonyls. These results are in agreement with the molecules expected for bio-oil, as in ketones, the amount of CH_2 is always higher than CO (e.g. acetol and 2-oxobutanol) (Staš et al. 2014).

When it comes to terminal groups, the probability of ***substituent a CH_3*** is almost 100 % for both samples. In practice, this means that all (the big majority) of the methyl groups are replaced by an oxygenated group, as hydroxyl, aldehyde and carboxylic. This distribution is in the accordance with the expected molecules, because their terminal groups are mostly (if not all) oxygenated functions (Wang et al. 2013) (Bayerbach and Meier 2009) (Staš et al. 2014). Furthermore, and since the probability is almost 100 %, the presence of this diagram could be neglected. If all methyls were forced by the algorithm to be replaced, the final results would be similar to the one presented here. Furthermore, the algorithm would have less a diagram to go through, meaning that the noise associated to the stochastic nature would be reduced, as the number of parameters to estimate.

Finally, and concerning the ***type of the substituent***, hydroxyls are the most probable groups, followed by aldehydes and carboxylic groups. When comparing both samples, sample B has a lower probability in the case of hydroxyls, and a higher probability in the case of carboxylic than sample A. As mentioned

before, sample B presents less hydroxyls and methoxyls, but more carboxylics than sample A, hence justifying the deviations between the columns of the diagram. When compare with the expected molecules of bio-oil, these results are not totally in accordance. Although a high amount of hydroxyls exist in bio-oil, there is also a great quantity of carboxylic acids (Staš et al. 2014).

Estimated properties

The parameters presented in Table X-4 were optimized thanks to the experimental properties used as input: Elemental analysis (CHONS), ^{13}C NMR analysis and SEC analysis (report to Table X-2 and Figure IX-6). Here are presented the results and a discussion about the estimated properties. These results are presented in a parity plot with an error area of 10 %, similar to the model molecules. The values and errors associated to the following properties are present in appendix D.

For the **elemental analysis**, both samples present similar and reasonable results, as seen in Figure X-10. With the optimized parameters, the models can estimate the hydrogen amount and the correct trend for the oxygen and carbon content. The deviations associated with these last two elements are within the 10 % error area.

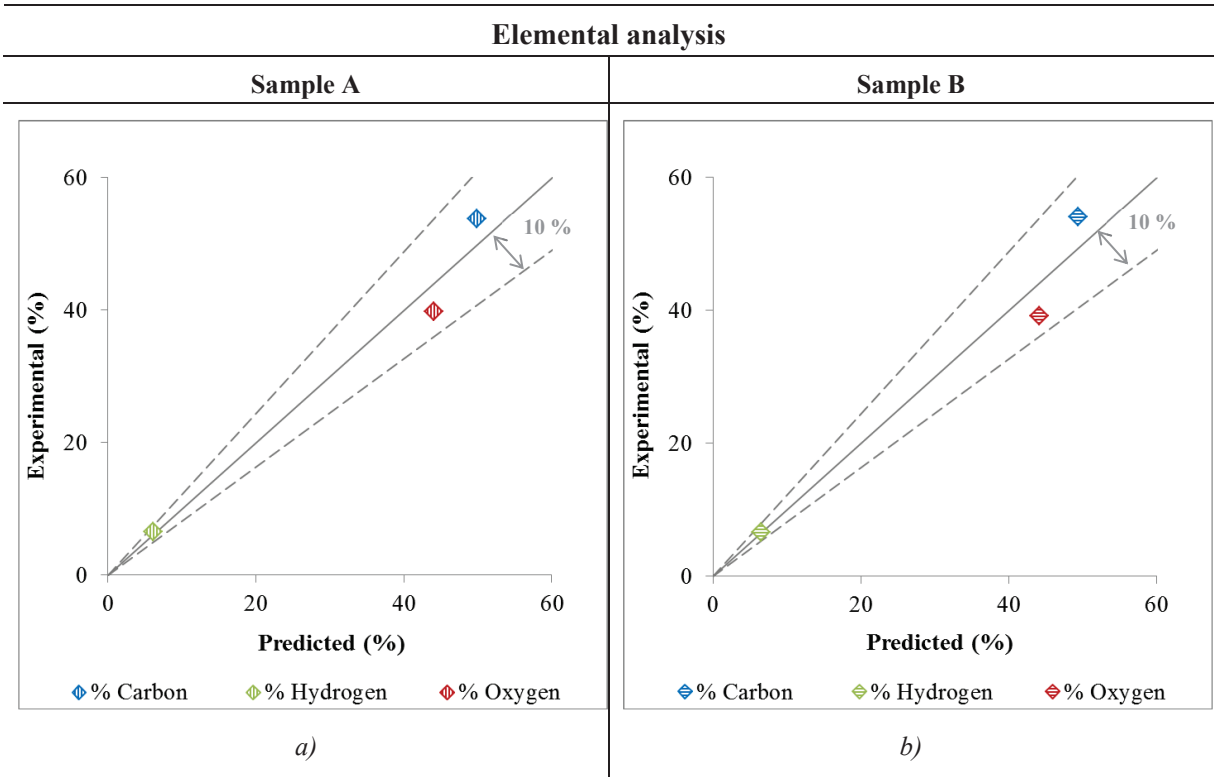


Figure X-10 - Results for the elemental analysis from building diagram #1 of: a) Sample A and b) Sample B.

For the **^{13}C NMR analysis**, and as shown in Figure X-11, both samples present similar and reasonable results for the aldehydes and ketones and finally the carboxylic acids fractions. However, sample B can estimate with more accuracy the fractions for aromatics, saturated and hydroxyl and methoxyls than sample A. As mentioned before, BD#1 generates much more paraffins than cyclic molecules, which are

mainly aromatics. In the case of paraffins, the replacement of atoms by methoxyls is impossible (report to Figure X-7), unlike the hydroxyls. Even so, the replacement of a CH_3 by a OH can only occur twice per molecule, as the algorithm only creates n-paraffins which just have two methyl terminal groups. In the end, the algorithm reconstructs mostly n-paraffins with low oxygen content and some aromatic molecules. For this reason, when faced with samples with low saturated and aromatic content and high hydroxyl and methoxyl content, as sample A (report to Table X-2), the algorithm battles to generate a library to match the experimental data. However, when the samples have a high saturate and aromatic content and low hydroxyl and methoxyl content, as sample B (report to Table X-2), the algorithm can easily reproduce a suitable library.

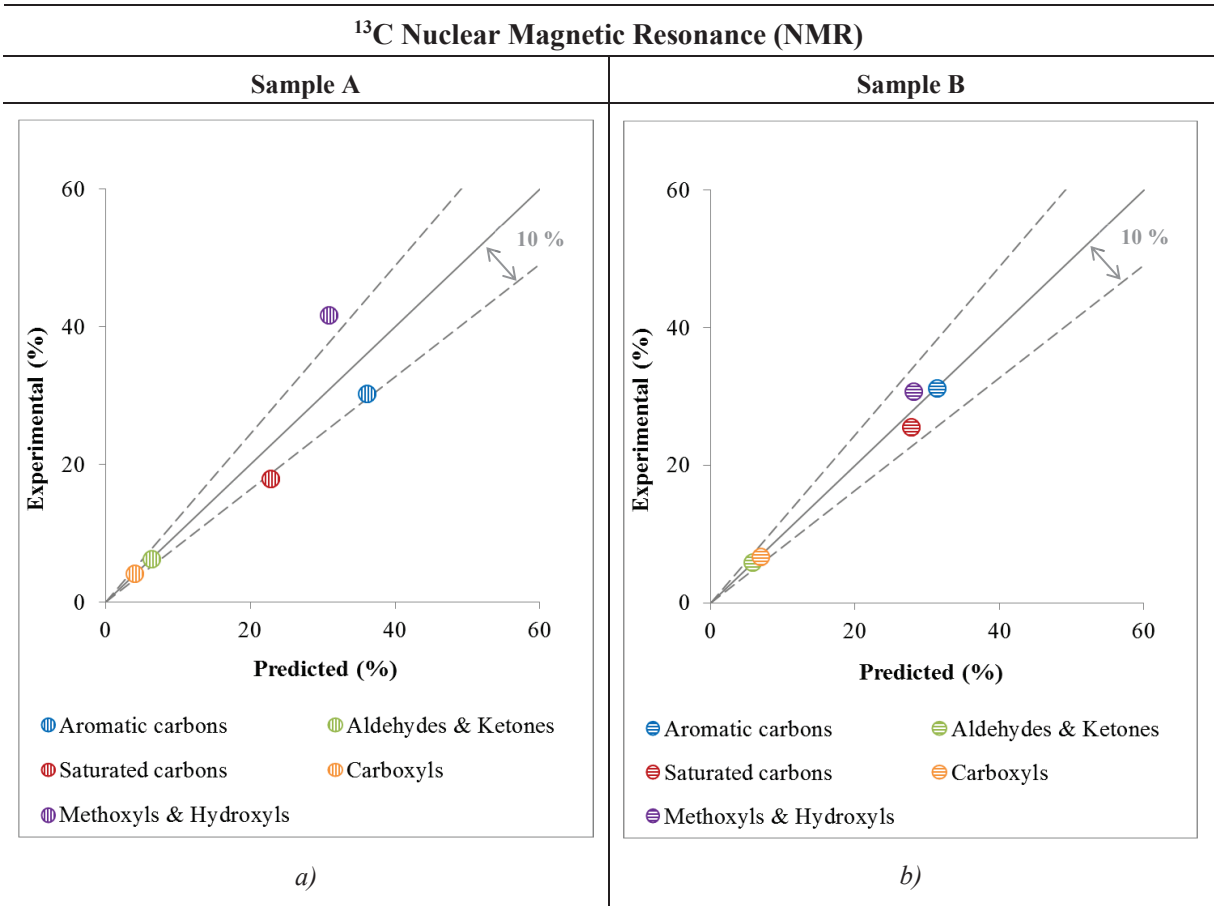


Figure X-11 - Results for the ^{13}C NMR analysis from building diagram #1 of: a) Sample A and b) Sample B.

For the **SEC analysis**, both samples present similar trend, as seen in Figure X-12. In the figure, the compared values are in polystyrene equivalent molecular weight and are divided by molar fractions, meaning that X % of the molecules have a Mw_{eq} PS equal or inferior to an Y value. For the initial experimental points of the SEC analysis (10 % - 50 %), the compositional model fits rather well the experimental data. However, after these, the model presents deviations, which grow with the fraction category. These deviations can originate either from the correlation implemented with the algorithm (report to Equation IX-9), or from the building diagram. Although all correlations are associated with non-conformities, in the present case, the main reason is the building diagram. As see before, BD#1

does not generates molecules with higher molecular weight than dimers. However, it is possible to find in bio-oil larger molecules than these (Bayerbach and Meier 2009) (Xu et al. 2014). For this reason, the model has divergences for the higher molecular weight molecules.

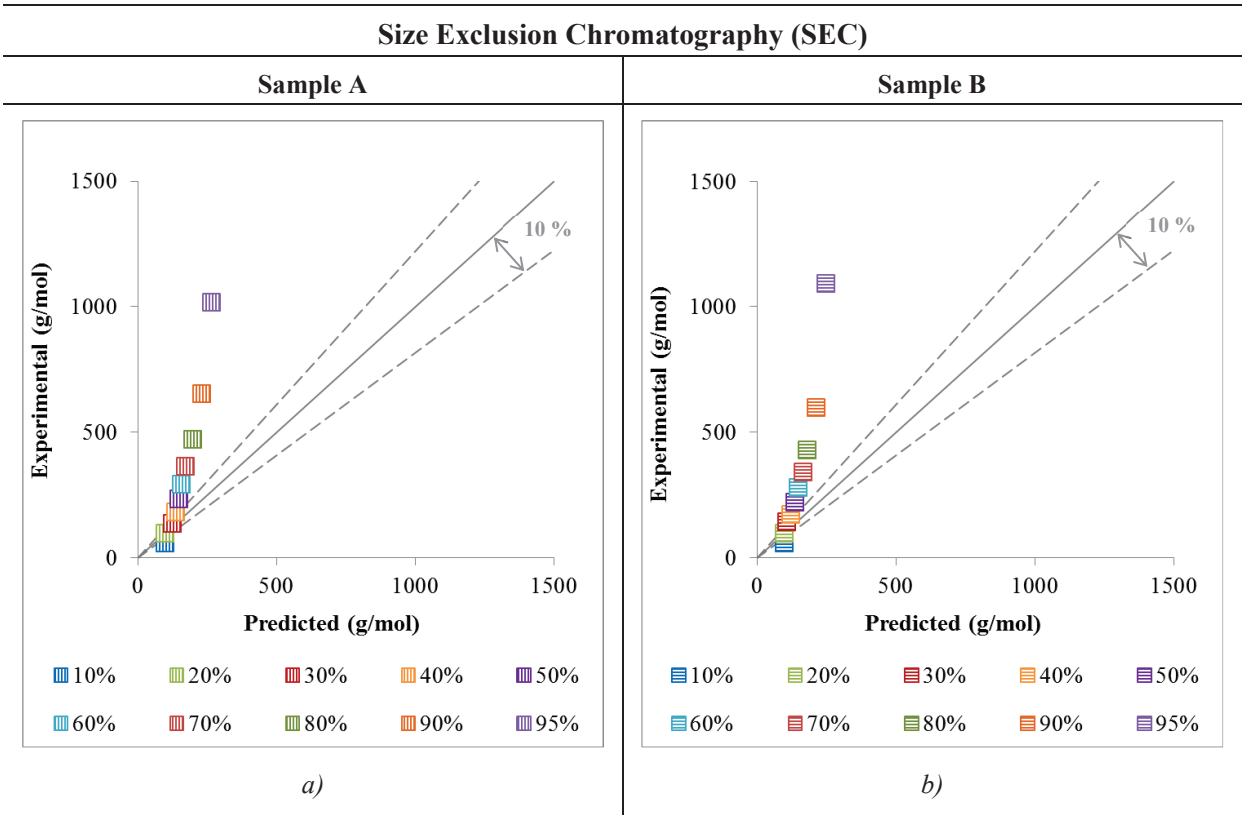


Figure X-12 - Results for the SEC analysis from building diagram #1 of: a) Sample A and b) Sample B.

Predicted properties

Besides the properties used in the estimation of the parameters, the SR can predict furthermore properties. When analytical data is available, these can be used to evaluate the model performance, allowing its validation. For sample A and sample B, the FT-ICR/MS results and the density (at 20 °C) were chosen to assess the model. Both experimental results were given by Figure X-5 and Table X-1, respectively. Here are presented the results and a discussion about the predicted properties. The values and errors associated to the following properties are presented in appendix D.

For the **FT-ICR/MS analysis**, both samples present an equal trend, as shown in Figure X-13. However, this trend does not correspond to the experimental results of FT-ICR/MS in any of the cases. As shown in Figure X-13, the experimental results follow the shape of a normal distribution, whereas the predicted results tend more to the shape of a gamma distribution. In other words, the reconstructed libraries have mainly molecules with one to three oxygens, while the experimental data has mainly a homogenous mixture, where most molecules have between four to ten oxygens. The absence of oxygen in the generated libraries maybe associated to high fraction of n-paraffins. As mentioned before, the algorithm reconstructs mostly n-paraffins, which even when oxygenated never present more than one, two or three

oxygens. Be that as it may, both samples also have molecules with four to fourteen (A) / twelve (B) oxygens. These correspond to the cyclic compounds reconstructed by the algorithm.

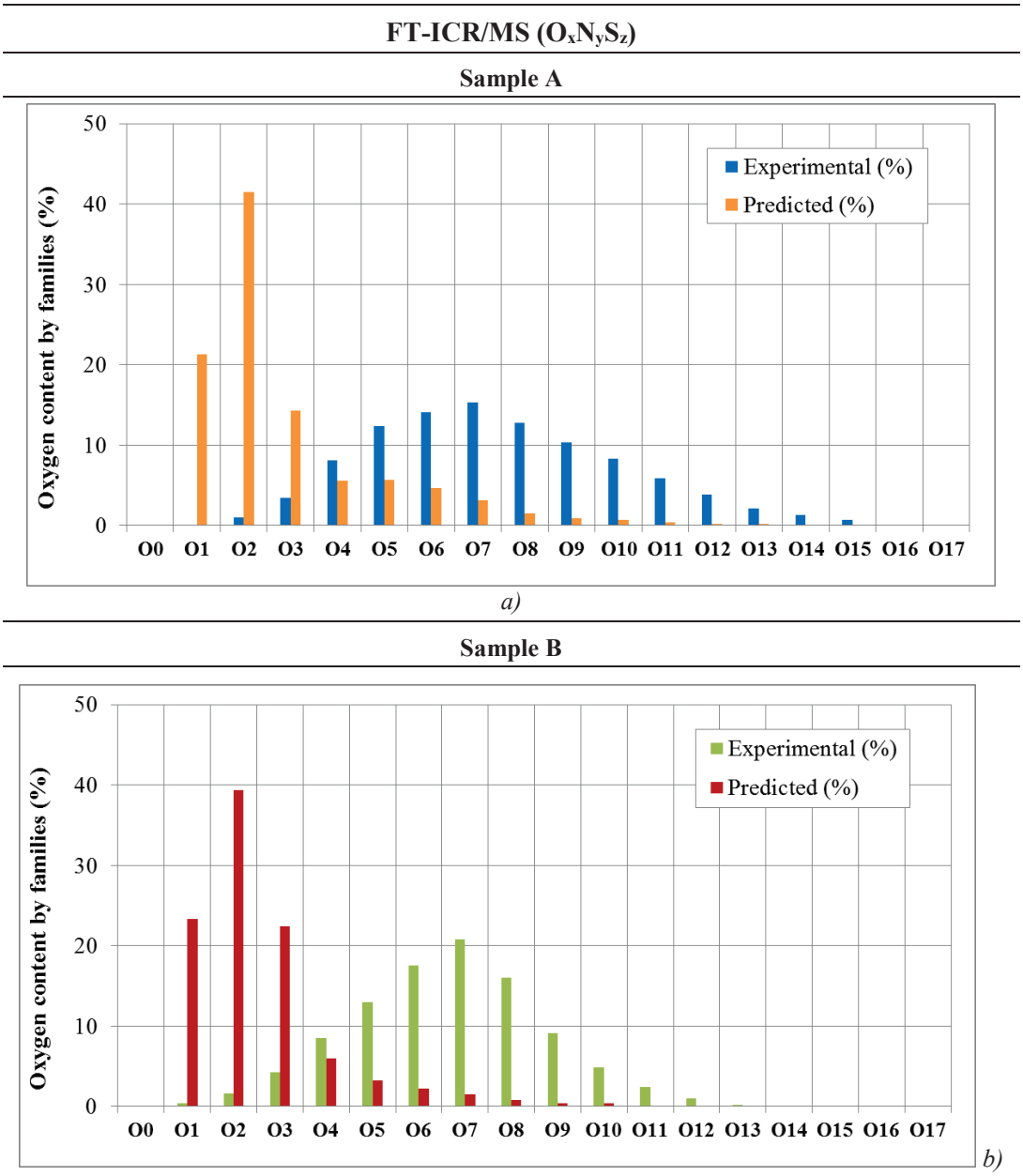


Figure X-13 - Results for the FT-ICR/MS analysis(%mol/mol) from building diagram #1 of: a) Sample A and b) Sample B.

For the **density at 20 °C**, the algorithm predicted for both samples a higher density value than the experimental data. As shown in Table X-5, the density values have a 5 % and 3 % relative error for sample A and sample B, respectively.

Table X-5 - Density and molecular weight predicted by building diagram #1 for sample A and sample B.

Building diagram #1			
Property	**Value type**	**Sample A**	**Sample B**
Density (g/cm^3)	Experimental	1.24	1.19
	Predicted	1.31	1.23
	Relative error (%)	5	3
Molecular weight (g/mol)	Predicted	103.4	90.4

In present work, the density is calculated by a method proposed by Hudebine and Wahl (Hudebine 2003), here given by Equation X-1. For both samples, the predicted density was beyond the statistical errors of the method (average relative error of 1.35 % and standard deviation of 0.04 g/cm^3).

$$d = \frac{Mw}{V_m} = \frac{Mw}{\sum_i \left(n_i \cdot \Delta V_{m,i} + F_{V_m}\right)} \qquad \textit{Equation X-1}$$

Where d is the density, Mw is the molecular weight, V_m is the molar volume, n_i is the number of contribution groups for structure group i, $\Delta V_{m,i}$ is the value of the contribution group of structural group i for the molar volume, F_{V_m} is a molar volume corrective factor which is obtained by Equation X-2, where N_R is the number of cycles in the molecule.

$$F_{V_m} = 25 \cdot N_R \qquad \textit{Equation X-2}$$

Besides non-conformities that may be associated with the method, the most plausible hypothesis for the deviations seen in Table X-5, is the preferential reconstructed of n-paraffins and the small cyclic molecules. As discussed before, the algorithm generates mostly n-paraffins, then monomers and finally a few dimers. In the cyclic molecules, the number of rings never overcomes three, even though is usually just one ($N_R = 1$). For this reason, the value of F_{V_m} is also small, which leads to an increase in the density of pure compounds. Besides, in the linear molecules, N_R is zero as the F_{V_m}value, which also promotes a higher density value. In the end, small cyclic molecules with an over predicted density and a high amount of n-paraffins increase the average density above the desired value. To finish, a high average molecular weight, retrieved from the pure compounds molecular weight, could also be problematic. However, since the algorithm generates smaller molecules than the real bio-oil feed (report to Figure X-12), the molecular weight of the mixture (see Table X-5) is also quite small. Therefore, this could not be the source of the problem.

Chemical structures

Besides the information presented in this section, it is crucial to analyze the structure of the generated molecules. As presented before, the SR reconstructs mostly paraffins (73 % for sample A and 82 % for sample B) and few cyclic molecules, from which the great majority have an aromatic character.

As presented in Figure X-14, the library of molecules of sample A and sample B has monoalcohols (one OH), dialcohols (two OH), dialdehydes (two COH), diacids (two COOH), and polyalcohols (more than two OH). Furthermore, the mixture also has "Di_Oxygen_FG" and "Poly_Oxygen_FG" families, which represent molecules with two or more than two distinctive functional groups, respectively. Molecules with just one or two hydroxyl and with two distinct functional groups are the most common in the both samples. As mentioned before, both samples are majority composed by n-paraffins, which can only be replaced maximum twice. Therefore, it would be wise to assume that most of these monoalcohols, dialcohols and molecules with two FG are indeed *n*-paraffins. The residual traces of dialdehydes and diacids are included as well. As for the molecules with three or more FG, these can only be cyclic compounds, for these accepted further than two FG.

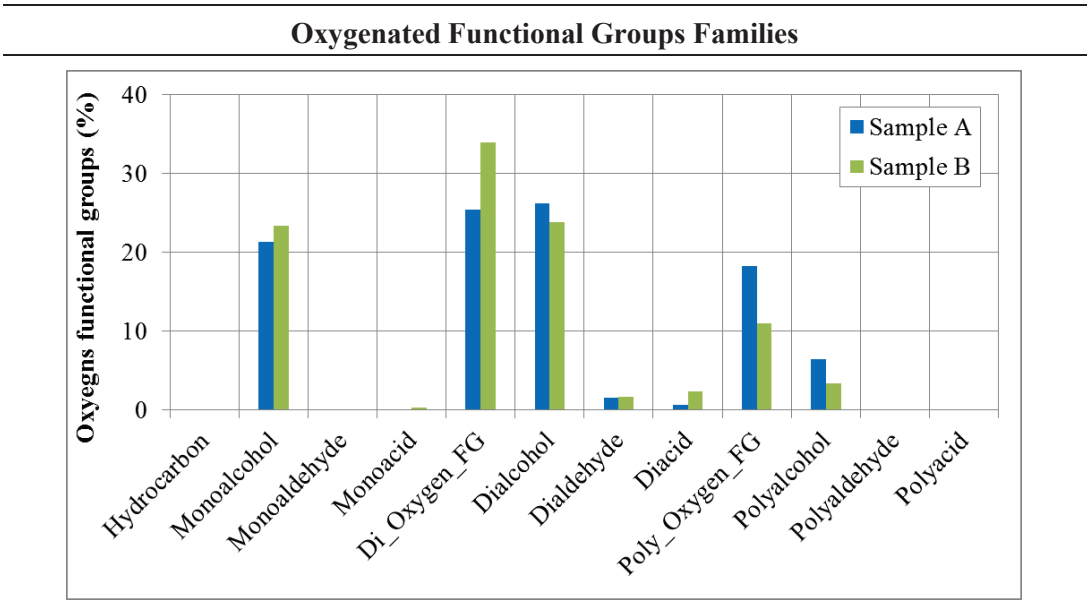


Figure X-14 – Predicted oxygenated families (%mol/mol) for building diagram #1 from sample A and sample B.

Figure X-14 also shows that sample A has more dialcohols and polyalcohols and less diacids than sample B, which is in accordance with the ^{13}C NMR data (report to Table X-1). However, it should be expected a higher amount of monoalcohols in sample A than B. Regarding the "Di_Oxygen_FG" and "Poly_Oxygen_FG" families, it is difficult to analyze, as the exact FG are unknown. However, the first family may include molecules with carboxylic acids, which would justify the higher amount of these in sample B than in sample A (report to ^{13}C NMR data in Table X-1). As for the "Poly_Oxygen_FG" family, the inclusion of methoxyls in the cores, would lead to an increase of this family for sample A, since it has a higher amount of these than sample B (report to ^{13}C NMR data in Table X-1).

To conclude this section, Figure X-15 presents several examples of molecules reconstructed by BD#1 for both samples. Methanol (*a)* is the smaller molecule generated by the SR, while the compounds named of *l)* are one of the largest. As seen before, BD#1 generates small paraffins, in addition to monomers and dimers with maximum three cycles. This diagram can also recreate several functional groups and the intense presence of oxygen in the molecules. The low molecular weight molecules (a, b, c, d, e, f, h and i) were already seen in the open literature (Staš et al. 2014) and are in agreement with the type of molecules expected for bio-oil. The high molecular weight molecules (j, k and l), although extremely difficult to identify, are in accordance with the structures proposed for the macromolecules existent in a bio-oil mixture (Bayerbach and Meier 2009) (Xu et al. 2014).

In summary, and from a mixture point of view, both libraries (sample A and sample B) are composed mainly of n-paraffins and small aromatic cyclic compounds. Furthermore, most of these molecules have only one to three oxygen atoms, which reduces the oxygen diversity of the typical functional groups of bio-oil. Although most of the features predicted by both compositional models correspond to the information in the open literature (Bayerbach and Meier 2009) (Staš et al. 2014), certain characteristics, as the size of the molecules and the diversity of the oxygen groups, are still missing. A desired library should present more high molecular weight compounds, which by increasing the number of potential substituted carbons, would increase the number of oxygen (better FT-ICR/MS results) and the diversity of the functional groups (better oxygen functional group families results) per molecules. For these reasons, the libraries created by the BD#1 for sample A and sample B do not represent the totality of a bio-oil mixture.

X.3.1.1 Reconstruction by entropy maximization

As presented before, the libraries created by the SR method for BD#1 do not represent the totality of a bio-oil fraction. For this reason, it was concluded that the application of the reconstruction by entropy maximization would not be appropriate.

In the present conditions, if this method would be applied, the molar fractions would probably be overestimated for certain molecules and underestimated for others. As presented in section IX.2.2, the method used an optimization technique which will force the composition to vary and fit in the analytical data. If the library of molecules is not the most suitable for the mixture (as the present case), the reconstruction by entropy maximization would be also generate an unsuitable synthetic mixture.

Be that as it may, if the library of molecules was suitable for the mixture, the reconstruction by entropy maximization could easily generate a correct synthetic mixture, as already seen for heavy oil fractions. Even if only the library of molecule of one of the bio-oil samples was appropriate, the reconstruction by entropy maximization could recreate both synthetic mixtures.

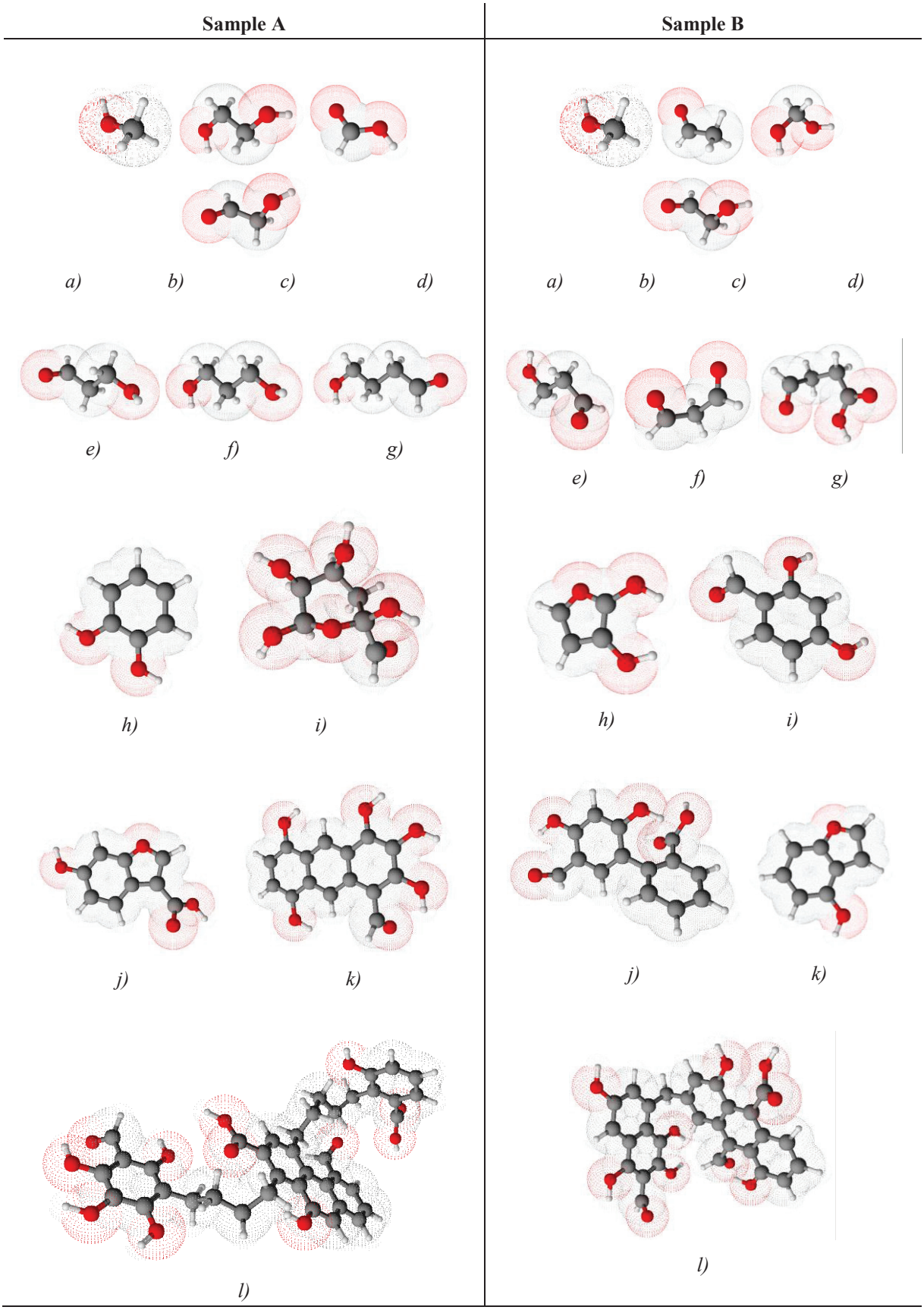


Figure X-15 - Example of reconstructed molecules from BD#1 for sample A and sample B: black – carbon, white – hydrogen, red - oxygen (report to appendix E for the molecules names)

X.3.2 Second hypothesis of reconstruction

As seen before, the first hypothesis of reconstruction presents certain shortcomings, in particular at the level of the size of the molecules and the ^{13}C NMR. For this reason, a second approach was created. Here, it will be presented the second hypothesis of reconstruction (BD#2). Furthermore, the present section includes the results for both bio-oil samples.

X.3.2.1 Building diagram

On contrary to the first building diagram, the second diagram here presented starts with the sampling of the ***type of molecules***, as shown in Figure X-16. It was assumed that the addition of this new attribute could help the algorithm manage the division of the two types of molecules. By forcing this selection, the algorithm could increase the size of the molecules, by sampling more cyclic molecules, and reestablish the balance between the aromatic and saturated carbons by finding an equilibrium between the paraffin and the cyclic molecules. Furthermore, the addition of this attribute would also force the algorithm to increase the number of cores, because, once a cyclic molecule is sampled, its core number must mandatorily be at least one.

Concerning just the molecular attributes, and apart the addition of a new one, the second diagram is similar to the first, as shown in Figure X-7 and Figure X-16. However, at the probability distributions function, the second diagram still has two differences. In order to increase of the molecules size, the distribution of the ***number of cores*** was changed from an exponential to a gamma distribution. As discussed in section IX.1.2.1, a gamma distribution allows the algorithm to sample higher values for the attributes. In this case, the idea is to sample a higher number of cores. Besides the number of cores, the distribution of the ***number of rings*** was changed as well. Since the aim is to increase the size of molecules, the algorithm requires a distribution that allows more than three rings per core. Therefore, this distribution could be replaced by an exponential or gamma function. As seen before, exponential distributions are mostly use for attributes with a wider range, but with a higher probability to sample smaller attributes values. The gamma distribution is also for wider ranges, but allows the algorithm to sampled larger attributes values. For these reasons, the distribution for the number of rings was changed for a gamma distribution as well.

In Table X-6 is presented a summary of the molecular attributes, together with their probability function distributions and their respective number of parameters. In summary, the building diagram here proposed, and present in Figure X-16, has ten distributions and fourteen parameters to be estimated.

Table X-6 - Molecular attributes and the respective PDF for building diagram 2.

PDF Number	Molecular attribute	PDF Type	Parameters Number
0	**Type of molecule**	**Histogram**	**1**
1	**Number of cores**	**Gamma**	**1**
2	Length of chains	Gamma	1
3	**Number of rings**	**Gamma**	**1**
4	Type of ring	Histogram	3
5	Addition of alkyl chains	Histogram	1
6	Type of chains	Histogram	2
7	Addition of carbonyl to CH_2	Histogram	1
8	Substitute a CH_3	Histogram	1
9	Type of substituent	Histogram	2

X.3.2.1 Stochastic reconstruction

For sample A and sample B, a library of molecules was generated through building diagram 2 (BD#2). Here, the results will be presented and discussed. These were divided into five sections: Objective function, probability distribution diagrams, estimated properties, predicted properties (those not used in the fitting of the parameters) and finally the molecular structures.

Objective function

For building diagram #2, the minimum objective function value was of 119.51 and 111.35 for sample A and sample B, respectively. Similar to building diagram#1, these values are far from the desired result. However, these were obtained after a series of two individual optimizations processes. As previously explained, this procedure intended to force the reduction of the objective function. In the case of BD#2, and ass shown in Figure X-17, after iteration 60, for sample A, and 20, for sample B, the OF value is practically constant, meaning that this can hardly decrease furthermore. Thus, and as for BD#1, the optimization procedure was stopped.

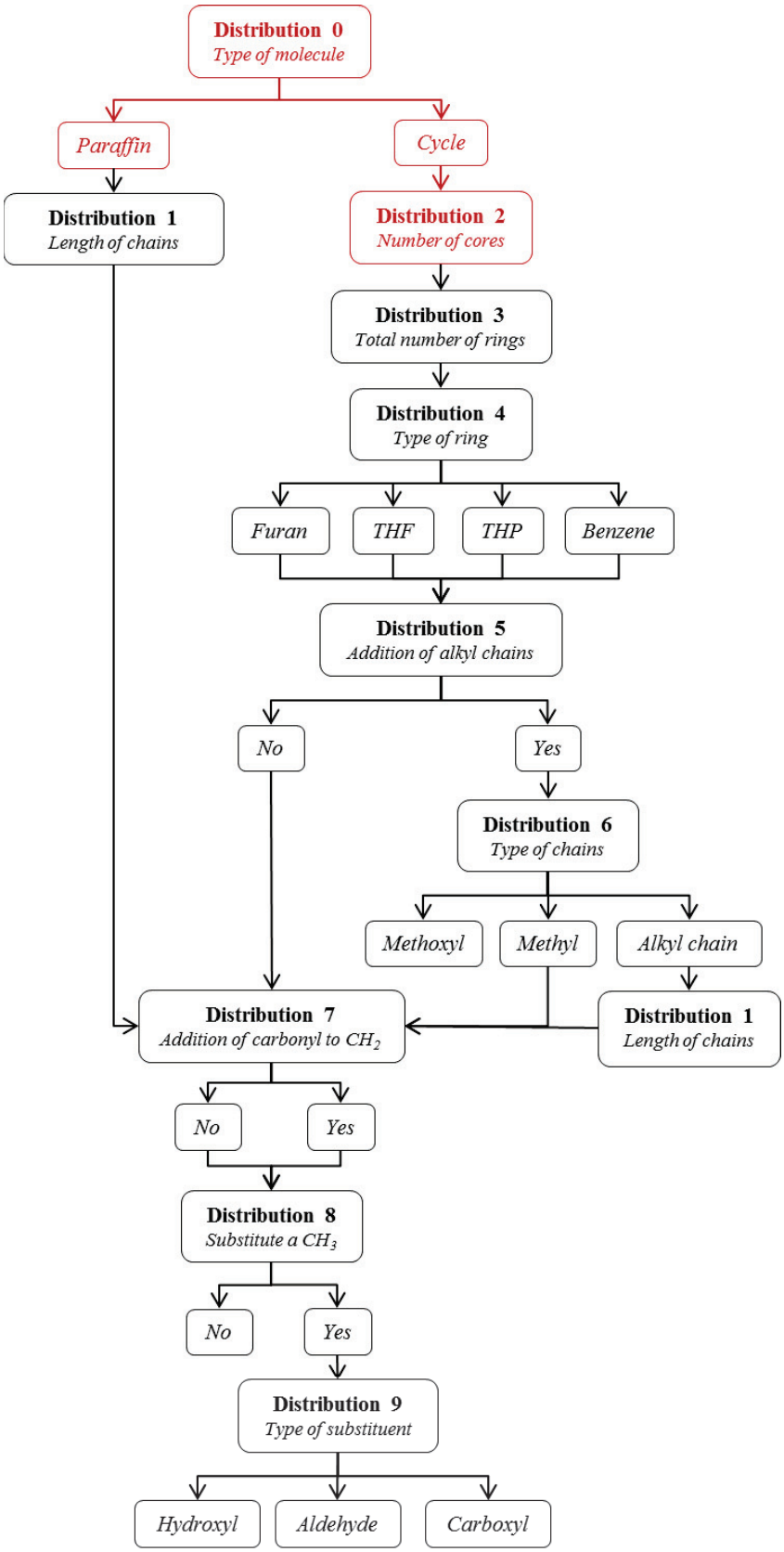


Figure X-16 - Diagram number 2 for the reconstruction of bio-oils.

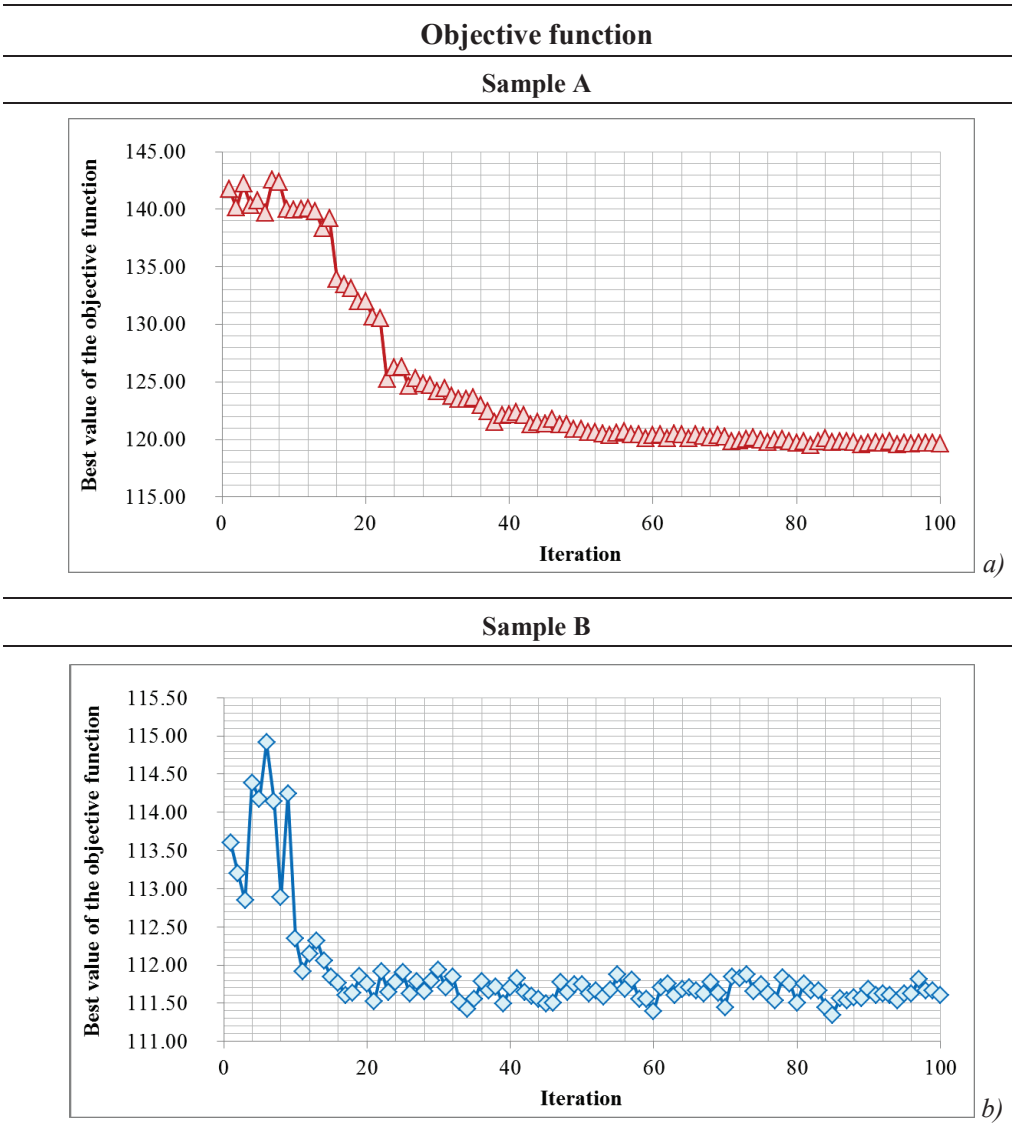


Figure X-17 - Evolution of the nest value of the objective function with the iterations: a) Sample A; b) Sample B.

Probability distribution functions

For the previous reason, it was considered that the parameters obtained after the 2nd optimization run were the optimum values. These are present in Table X-7, together with their respective molecular attribute and probability distribution function (PDF). For a better analysis of the optimized parameters, the results of their discretized PDF are illustrated in Figure X-18.

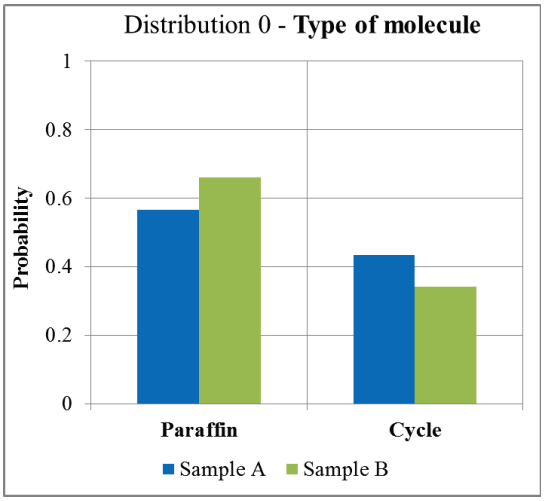
Table X-7 - Molecular attributes and the respective PDF and optimized parameters for building diagram 2.

PDF Number	Molecular attribute	PDF Type	Parameters Number	Optimized parameter value	
				Sample A	Sample B
0	**Type of molecule**	**Histogram[2]**	**1***	**0.566**	**0.659**
				0.434	**0.341**
1	**Number of cores**	**Gamma**	**1**	**0.503**	**0.512**
2	Length of chains	Gamma	1	1.008	1.018
3	**Number of rings**	**Gamma**	**1**	**0.572**	**0.888**
4	Type of ring	Histogram	3*	0.980	0.937
				0.014	0.026
				0.000	0.000
				0.006	0.037
5	Addition of alkyl chains	Histogram	1*	0.381	0.398
				0.619	0.602
6	Type of chains	Histogram	2*	0.130	0.064
				0.867	0.632
				0.003	0.304
7	Addition of carbonyl to CH_2	Histogram	1*	0.709	0.929
				0.291	0.071
8	Substitute a CH_3	Histogram	1*	0.005	0.029
				0.995	0.971
9	Type of substituent	Histogram	2*	0.735	0.647
				0.161	0.164
				0.104	0.188

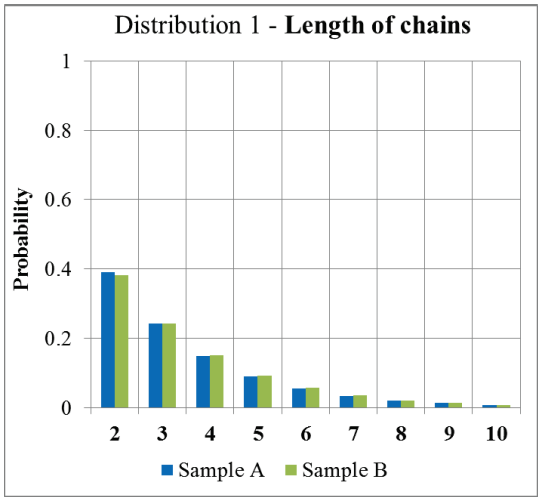
As seen in Figure X-18, the PDFs follow the same trends for both samples (except the attributes: number of rings and addition of a carbonyl). Even so, there is certain differences between both (also seen in Table X-7). These are originated from the distinct experimental data of sample A and sample B used by the SR for the estimation of the parameters.

[2] Once again, in histograms, the extra parameter (*) is always obtained through difference (1 – the other parameters).

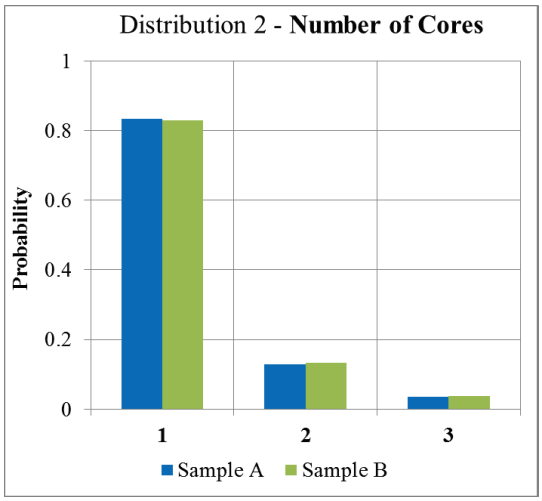
Building diagram # 2 - Sample A and Sample B



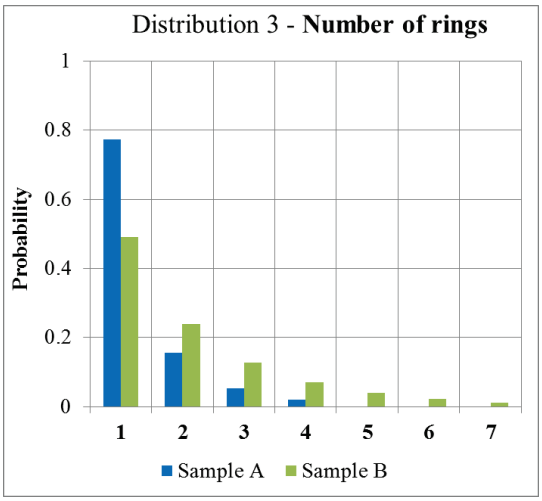
a)



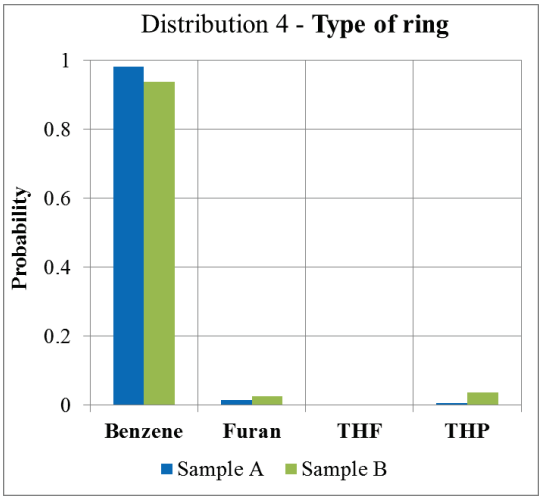
b)



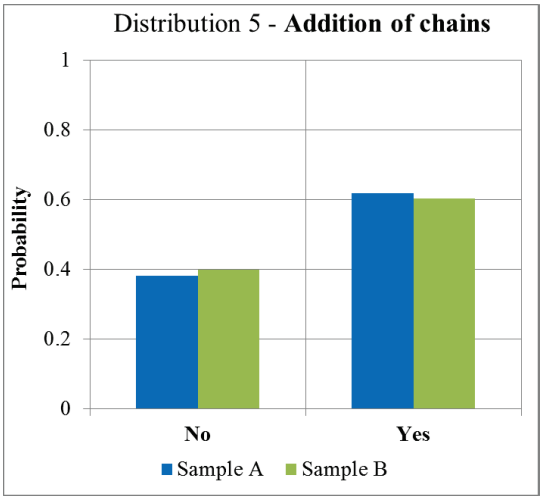
c)



d)



e)



f)

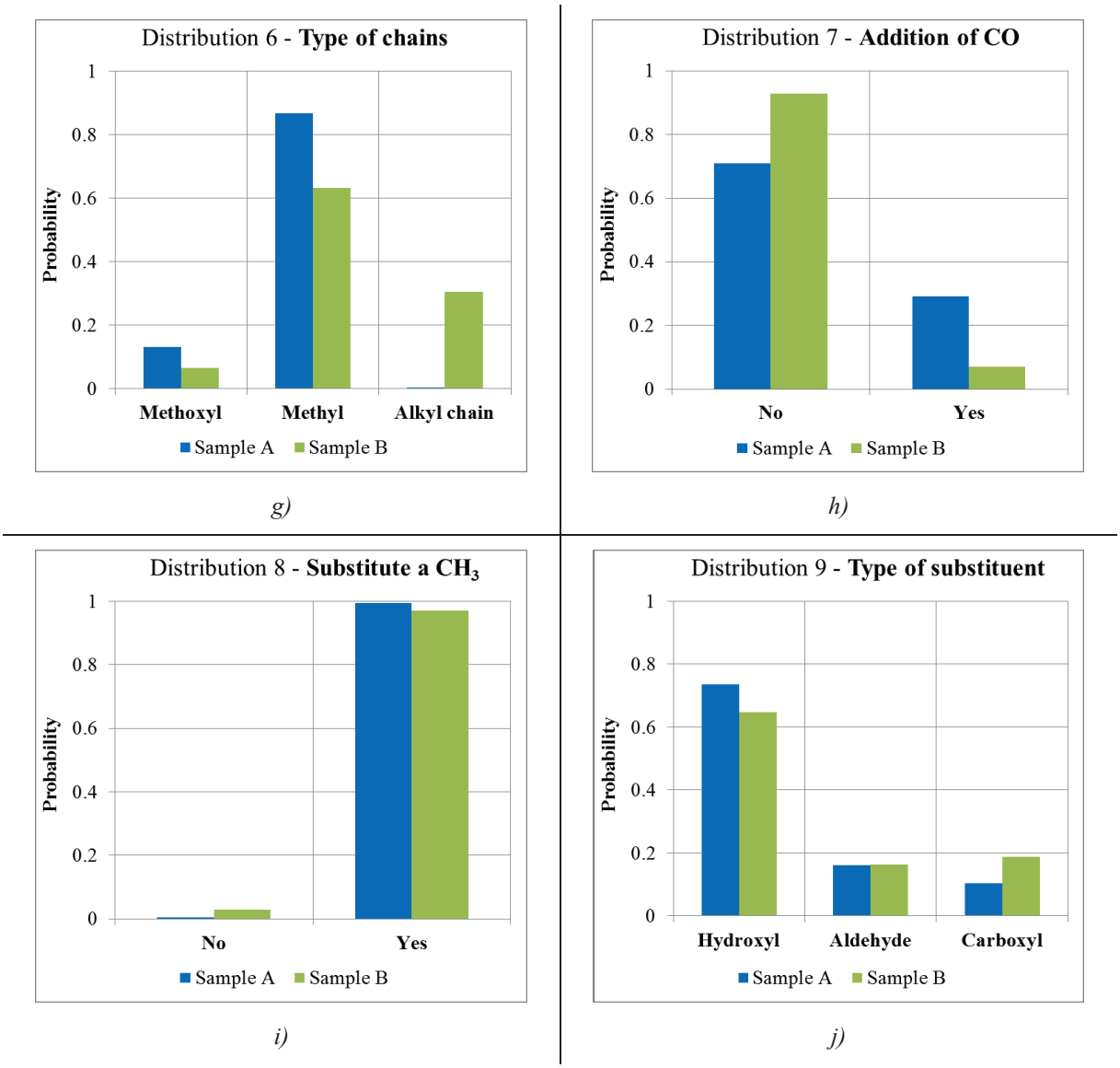


Figure X-18 - Resultant probability distribution functions of building diagram #2 for sample A and B: a) Distribution 0 – Type of molecule; b) Distribution 1 – Length of chains; c) Distribution 2 – Number of cores; d) Distribution 3 - Total number of rings; e) Distribution 4 – Type of ring; f) Distribution 5 – Addition of alkyl chains; g) Distribution 6 – Type of chain; h) Distribution 7 – Addition of a carbonyl to a CH_2; i)Distribution 8 – Substitute a CH_3; j) Distribution 9 – Type of substituent.

For both samples, the ***type of molecules*** PDFs present similar results. In both cases, there is a higher probability of generation linear molecules (paraffins), than cyclic compounds. Furthermore, sample B has a higher probability to create paraffins than sample A. This may justify by the higher content of saturated carbons in sample B than sample A. In summary, the algorithm reconstructs usually more *n*-paraffins, but may also generate an important number of cyclic molecules. This results are coherent with the structures found in bio-oil (Staš et al. 2014).

For the ***length of chains***, both samples share an identical behavior, also seen for BD#1. Chains with two atoms are the most probably to be reconstructed, followed by chains with three and four atoms. Larger chains with five or more atoms can also be generated, but their probability is lower than the formers. In summary, the algorithm reconstructed more paraffins with two, three and four atoms and rings with

sidechains equally small. These type of structures are expected for typical molecules of bio-oil (Bayerbach and Meier 2009) (Staš et al. 2014).

For both samples, the ***number of cores*** present similar results. The algorithm predicts three cores as the maximum for the molecules. In both cases, there is a higher probability of generation monomers, than dimers or trimers. Amongst these, dimers are also more probable than trimers. In summary, the algorithm reconstructs mainly monomers, follow by dimers and trimers. Although these structures can be found in bio-oil, these mixtures have also larger molecules (Bayerbach and Meier 2009).

Concerning the ***number of rings***, the algorithm predicts four rings per core as the maximum for the molecules in sample A, whereas for sample B this value can achieve seven rings per core. Be that as it may, for both cases one ring per core is the most probable event, followed by two rings, then three rings and so on. The difference between the PDF of sample B and sample A can be justified by the SEC experimental results, which state that sample B has more high molecular weight compounds than sample A (see Figure X-6). In summary, the algorithm reconstructs mainly cores (monomers) with one or two ring and occasionally with three or more. These molecular structures, with one/two rings per core, are in agreement Bayerbach and Meier (Bayerbach and Meier 2009) (see Figure X-2).

For each ring, the most probably type is benzene, follow by furan, then THP. As shown in Table X-7 for the ***types of ring***, THF do not have any probability of being sampled. Both furan and THP have a much lower probability than benzene. As mentioned before, the amount of each type of rings is unknown. However, a higher amount of benzene can be expected, due to the presence of pyrolytic lignin on bio-oil (report to Figure I-2 the State of the Art chapter). In summary, the algorithm reconstructs mostly benzene rings and occasionally furans or THP, which is in agreement with the structures of bio-oil (Bayerbach and Meier 2009) (Staš et al. 2014).

In the case of the ***addition of chains***, both samples have a higher probability of accepting sidechains. Even so there is a difference in the probability. As mentioned for BD#1, such deviations are caused by the experimental data, which is not the same for both samples. For this case, the sample B has less hydroxyls and methoxyls groups than sample A. Therefore, by reducing the number of sidechains in the rings (accepting more “no”), the algorithm can also reduce the amount of these two groups. In summary, the algorithm generates mostly rings with sidechains. As seen in before in Figure X-1 and Figure X-2, the structure reconstructed by the algorithm are within the usual representation of bio-oil molecules (Bayerbach and Meier 2009) (Staš et al. 2014).

Similar to the rings, for each sidechain, there is a ***type of chain***. For both samples, the methyl sidechains are the most probable type. These are followed by methoxyls and then alkyl chains, in the case of sample A, and by alkyl chains and the methyls in the case of sample B. These last types have a much lower probability than the methyls, in particularly the alkyl chains in sample A which are practically

impossible to generate, as its probability is extremely low. Although sample B has less hydroxyls and methoxyls groups than sample A, as mentioned before, the algorithm predicts a higher probability for selecting methoxyls for sample B than sample A. One of the reasons maybe that instead of sampling too much methyls, which will eventually become hydroxyls (also in lower quantity in sample B), the algorithm prefers to occasionally sample methoxyls. In this way the number of hydroxyls will be kept low and number of saturated carbons will increase (in higher quantity in sample B), due to the presence of the CH_3 in the methoxyl group. Regarding the alkyl chains, this may be justified by the SEC experimental results. By sampling more alkyl chains, the algorithm increases the number of carbon atoms in the molecules. This can the lead to the increase of the molecule molecular weight. Since sample B has more high molecular weight compounds than sample A (see Figure X-6), it is justified why sample B presented a higher probability for alkyl chains than sample A. In summary, the algorithm will mostly generate rings with methyl sidechains, which may eventually become hydroxyls. When compare with the expected molecules of bio-oil, these results are in accordance. However, it is also quite possible to find molecules with one or more methoxyls or even alkyls chains in bio-oil mixtures (Wang et al. 2013) (Bayerbach and Meier 2009).

Regarding the heteroatom groups, both samples have more probability to not add a carbonyl group. Also, the ***addition of carbonyl*** (CO) varies between both samples. A higher probability of adding a CO is seen in sample A. As it happens, this sample has a higher number of aldehydes and ketones, thus the probability for the sampling of the yes is higher for sample A than sample B. In summary, the algorithm generates mostly chains without or with few carbonyls. These results are in agreement with the molecules expected for bio-oil, as in ketones, the amount of CH_2 is always higher than CO (e.g. acetol and 2-oxobutanol) (Staš et al. 2014).

When it comes to terminal groups, the probability of ***substituent a CH_3*** is almost 100 % for both samples, as for BD#1. As mentioned before, most methyl groups are replaced by an oxygenated group, as hydroxyl, aldehyde and carboxylic. This distribution is in the accordance with the expected molecules, because their terminal groups are mostly (if not all) oxygenated functions (Wang et al. 2013) (Bayerbach and Meier 2009) (Staš et al. 2014). AS seen also before, the presence of this PDF could be neglected, as it is basically the same as forcing the algorithm to replace methyls. The suppression of this PDF would allow the reduction of the noise associated to the stochastic nature and the number of parameters to be estimated.

Finally, concerning the ***type of the substituent***, hydroxyls are the most probable groups, followed by aldehydes and carboxylic groups. As presented for BD#1, sample B has a lower probability in the case of hydroxyls, and a higher probability in the case of carboxylic than sample A. As mentioned before, sample B presents less hydroxyls and methoxyls, but more carboxylics than sample A, thus justifying the deviations between the columns of the diagram. When compared to the expected molecules of bio-

oil, these results are not totally in accordance. Although a high amount of hydroxyls exist in bio-oil, there is also a great quantity of carboxylic acids (Omais et al. 2013).

Estimated properties

The parameters present in Table X-7 were optimized thanks to the experimental properties used as input: Elemental analysis (CHONS), ^{13}C NMR analysis and SEC analysis (report to Table X-2 and Figure IX-6). Here is presented the results and a discussion about the estimated properties. These results are presented in a parity plot with an error area of 10 %, similar to the model molecules. The values and errors associated to the following properties are present in appendix D.

For the **elemental analysis**, both samples present similar and reasonable results, as seen in Figure X-19. With the optimized parameters, the models can estimate the hydrogen content and the correct trend for the oxygen and carbon content. The deviations associated with these last two elements in sample A are within the experimental error of the 10 % error area.

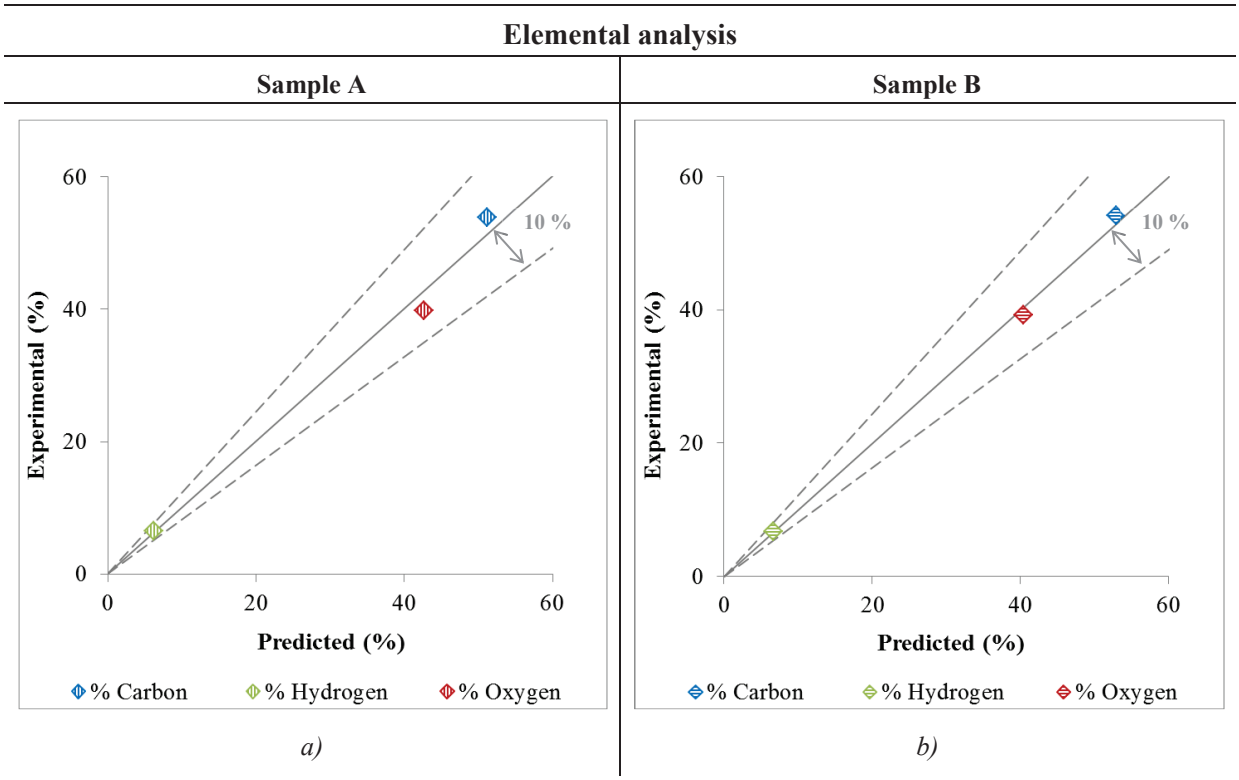


Figure X-19 - Results for the elemental analysis from building diagram #2 of: a) Sample A and b) Sample B.

For the **^{13}C NMR analysis**, and as shown in Figure X-20 , both samples present similar and reasonable results for the aldehydes and ketones and finally the carboxylic acids fractions. In the case of the aromatics and hydroxyl and methoxyls, sample B can predict them more accuracy than sample A. The deviation of the saturated fraction in sample A and the deviation of the saturated and the hydroxyl and methoxyl fractions in sample B can be justified by the experimental error of the ^{13}C NMR. Regarding the aromatics and the hydroxyl and methoxyls, and as seen before, BD#2 generates much more paraffins

than cyclic molecules, which are mainly aromatics. In the case of paraffins, the replacement of atoms by methoxyls is impossible (report to Figure X-7), unlike the hydroxyls. Even so, the replacement of a CH_3 by a OH can only occur twice per molecule, because the algorithm only creates n-paraffins which just have two methyl terminal groups. In the case of the cyclic molecules, the algorithm generates more rings, mostly aromatic, for sample B than sample A. In the end, the algorithm is suitable for the edification of n-paraffins with low oxygen content and monomers with a high aromatic content. For these reasons, when faced with samples with low saturate and aromatic content and high hydroxyl and methoxyl content, as sample A (report to Table X-2), the algorithm battles to generate a library to match the experimental data. However, when the samples have a high saturate and aromatic content and low hydroxyl and methoxyl content, as sample B (report to Table X-2), the algorithm can easily reproduce a suitable library.

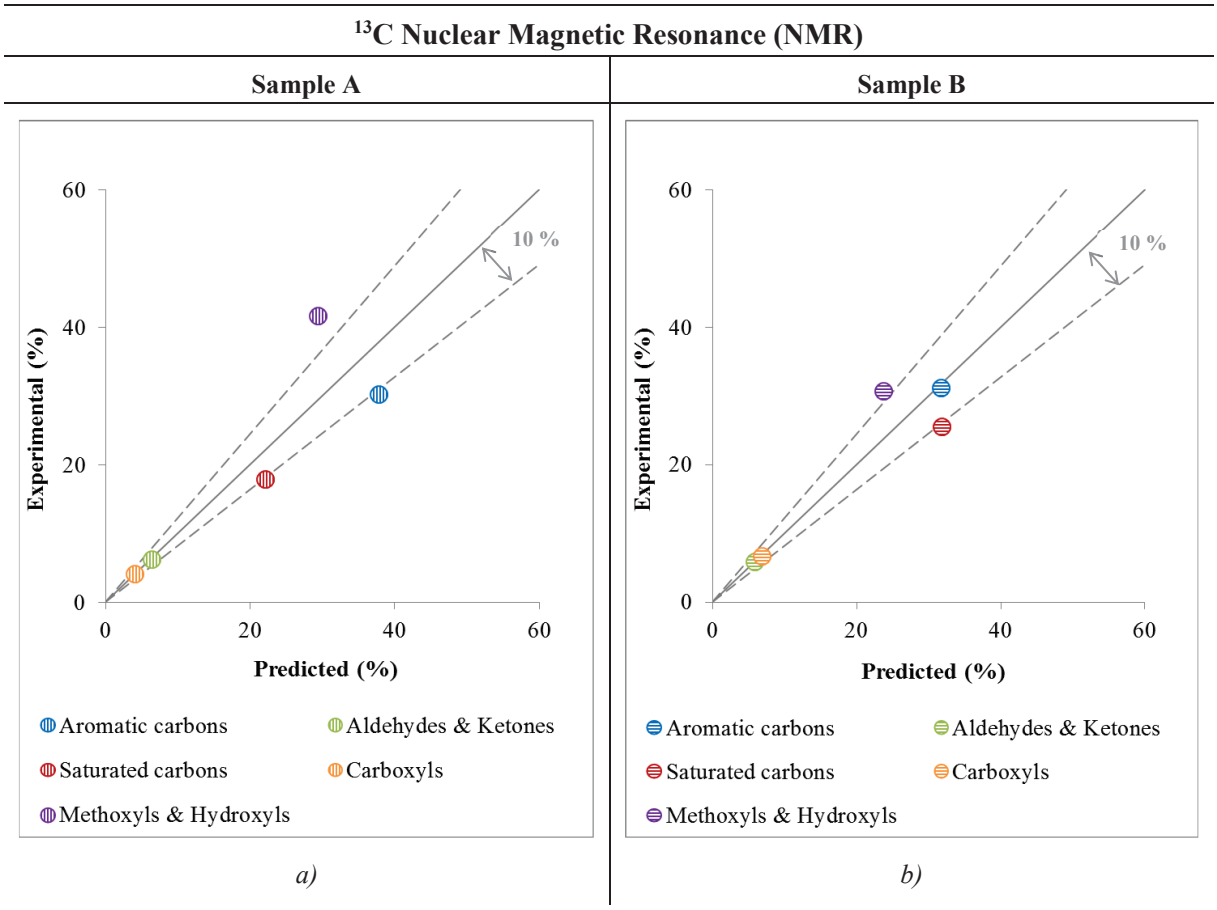


Figure X-20 - Results for the ^{13}C NMR analysis from building diagram #2 of: a) Sample A and b) Sample B.

For the **SEC analysis**, both samples present similar trend, as seen in Figure X-21. This trend is similar to the trend generated by BD#1, indicating that is associated with the correlation. For the initial experimental points of the SEC analysis (10 % - 50 %), the compositional model fits rather well the experimental data. However, after these, the model presents deviations, which grow with the fraction category. These deviations can originate either from the correlation implemented with the algorithm (report to Equation IX-9), or from the building diagram. Although all correlations are associated with

non-conformities, in the present case, the main reason is the building diagram. As see before, BD#2 generates mostly paraffins with low molecular weight and cyclic molecules not larger than trimers. However, it is possible to find in bio-oil larger molecules than these (Bayerbach and Meier 2009) (Xu et al. 2014). For this reason, the model has divergences for the higher molecular weight molecules.

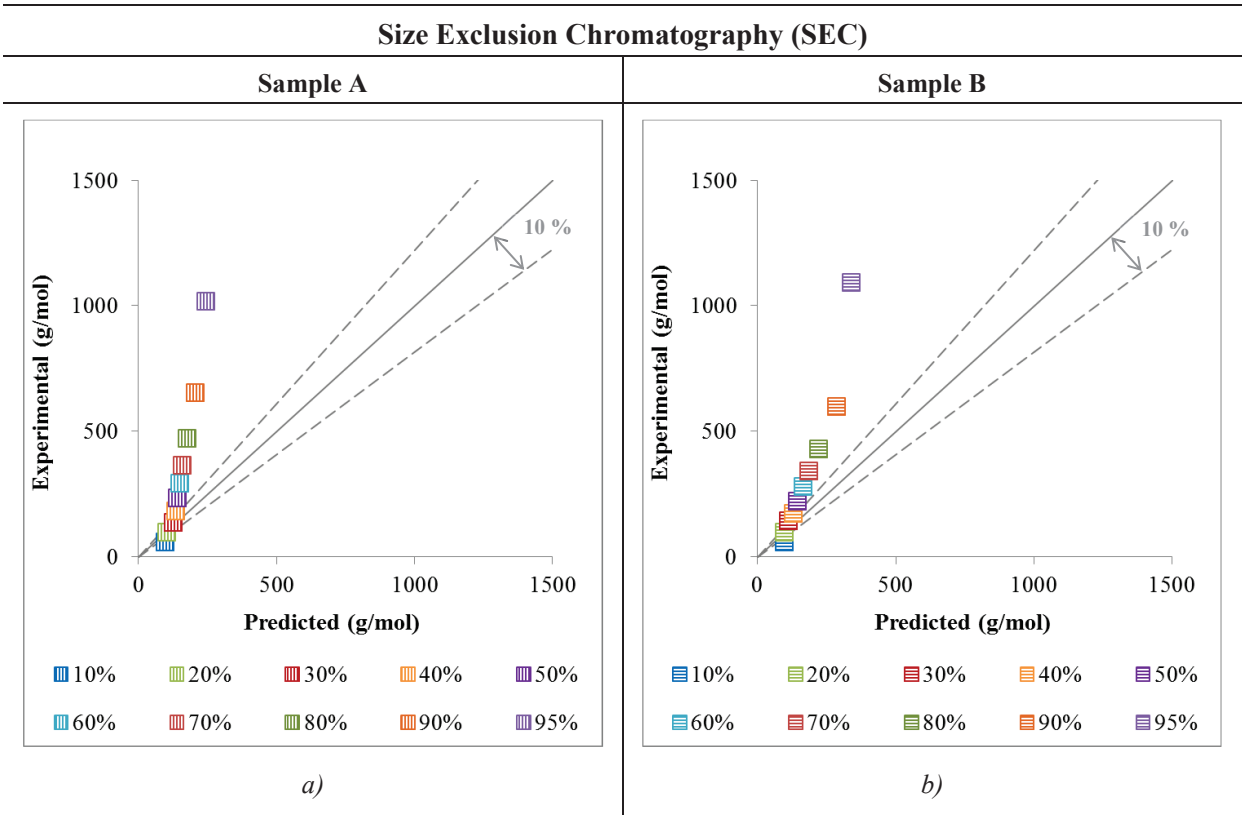


Figure X-21 - Results for the SEC analysis from building diagram #2 of: a) Sample A and b) Sample B.

Predicted properties

Similar to BD#1, beside the properties used in the estimation of the parameters, the SR can predict furthermore properties. When analytical data is available, these can be used to evaluate the model performance, allowing its validation. For sample A and sample B, the FT-ICR/MS results and the density (at 20 °C) were chosen to assess the model. Both experimental results were given by Figure X-22 and Table X-8, respectively. Here is presented the results and a discussion about the predicted properties. The values and errors associated to the following properties are present in appendix D.

For the **FT-ICR/MS analysis**, both samples present an equal trend, as shown in Figure X-22. As for the SEC analysis, this trend is similar to the trend generated by BD#1. Although both samples present similar results, the trend does not correspond to the experimental results of FT-ICR/MS in any of the cases. As shown in Figure X-22, the experimental results follow the shape of a normal distribution, whereas the predicted results tend more to the shape of a gamma distribution. In other words, the reconstructed libraries have mainly molecules with one to three oxygens, while the experimental data has mainly a homogenous mixture, where most molecules have between four to ten oxygens. The absent

of oxygen in the generated libraries maybe associated to high fraction of n-paraffins. As mentioned before, the algorithm reconstructs mostly n-paraffins with low oxygen content. Be that as it may, both samples also have molecules with four to twelve (A) / thirteen (B) oxygens. These correspond to the cyclic compounds reconstructed by the algorithm.

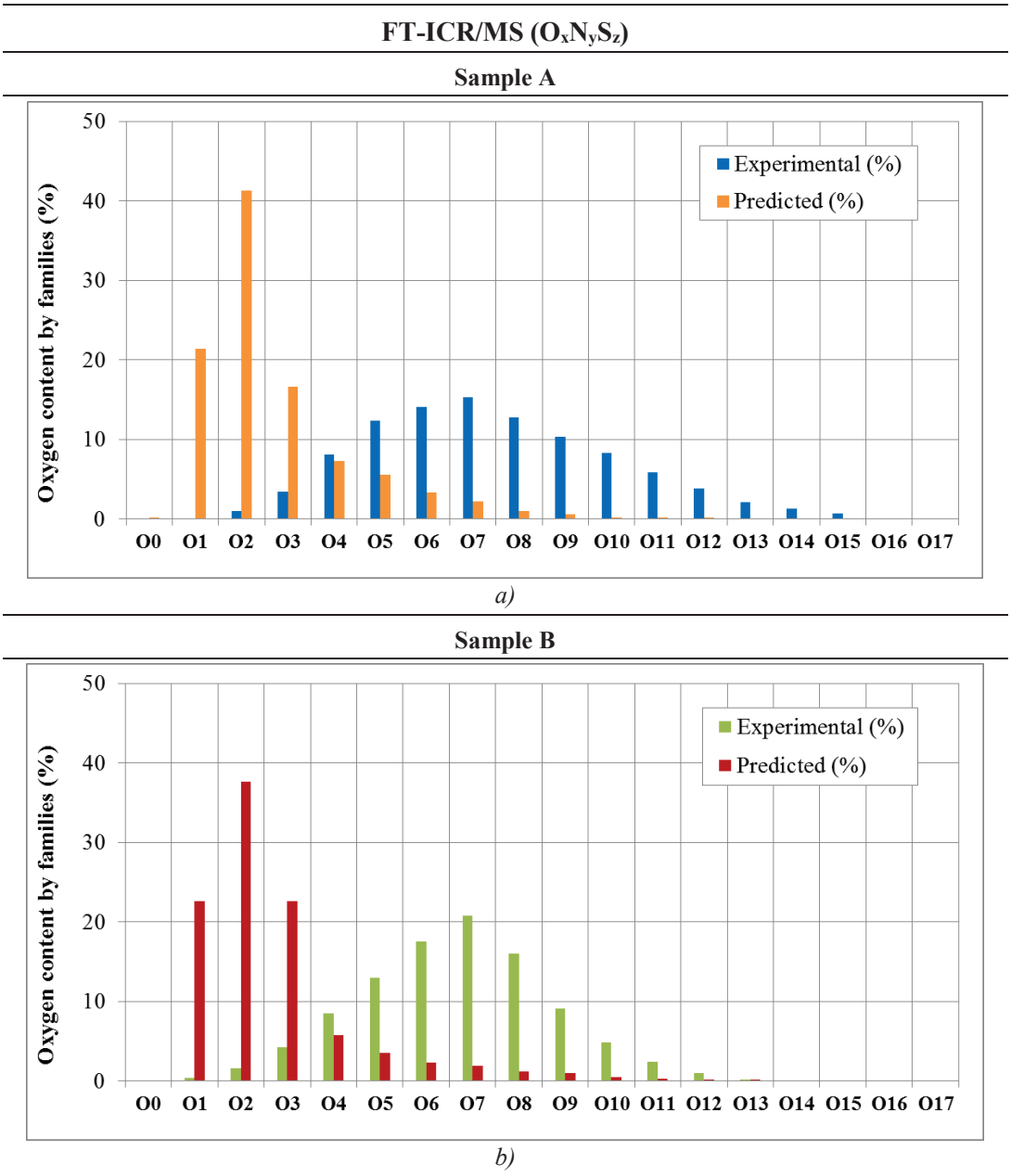


Figure X-22 - Results for the FT-ICR/MS analysis (%mol/mol) from building diagram #2 of: a) Sample A and b) Sample B.

For the **density at 20 °C**, the algorithm predicted for both samples a higher density value than the experimental data. As shown in Table X-8, the density values have a 3 % and 4 % relative error for sample A and sample B, respectively.

Table X-8 - Density and molecular weight predicted by building diagram #2 for sample A and sample B.

Building diagram #2			
Property	**Value type**	**Sample A**	**Sample B**
Density (g/cm^3)	Experimental	1.24	1.19
	Predicted	1.28	1.24
	Relative error (%)	3	4
Molecular weight (g/mol)	Predicted	99.8	106.5

As mentioned before, in present work, the density is calculated by a method proposed by Hudebine and Wahl (Hudebine 2003), here given by Equation X-1. For both samples, the predict density was beyond the statistical errors of the method (average relative error of 1.35 % and standard deviation of 0.04 g/cm^3). Besides non-conformities that may be associated with the method, the most plausible hypothesis for the deviations seen in Table X-8, is the preferential reconstruct of n-paraffins and the small cyclic molecules, as in BD#1. As discussed before, the algorithm generates mostly n-paraffins, then monomers and finally a few dimers. In the cyclic molecules, the number of rings never overcomes three, even though is usually just one ($N_R = 1$). For this reason, the value of F_{V_m} is also small, which leads to an increase in the density of pure compounds. Besides, in the linear molecules, N_R is zero as the F_{V_m} value, which also promotes a higher density value. In the end, small cyclic molecules with an over predicted density and a high amount of n-paraffins, increase the average density above the desired value. To finish, a high average molecular weight, retrieved from the pure compounds molecular weight, could also be problematic. However, since the algorithm generates smaller molecules than the real bio-oil feed (report to Figure X-12), the molecular weight of the mixture (see Table X-8) is also quite small. Therefore, this could never be the source of the problem.

Chemical structures

Besides the information presented in this section, it is crucial to analyze the structure of the generated molecules. As presented before, the SR reconstructs mostly paraffins (71 % for sample A and 80% for sample B) and few cyclic molecules, which all have an aromatic character (report to Figure X-18).

As presented in Figure X-23, and unlike the results of BD#1, the library of molecules of sample A and sample B has monoalcohols (one OH), monoaldehydes (one COH), monoacids (one COOH), dialcohols (two OH), dialdehydes (two COH), diacids (two COOH), and polyalcohols (more than two OH). Furthermore, the mixture also has "Di_Oxygen_FG" and "Poly_Oxygen_FG" families. Molecules with just one or two hydroxyl and with two distinct functional groups are the most common in the both samples. As mentioned before, and similar to the results of BD#1, both samples are majority composed by n-paraffins, which can only be replaced maximum twice. Therefore, it would be wise to assumed that most of these monoalcohols, dialcohols and molecules with two FG are indeed *n*-paraffins. The residual

traces of mono and dialdehydes and mono and diacids are included as well. As for the molecules with three or more FG, these can only be cyclic compounds, for these accept further than two FG.

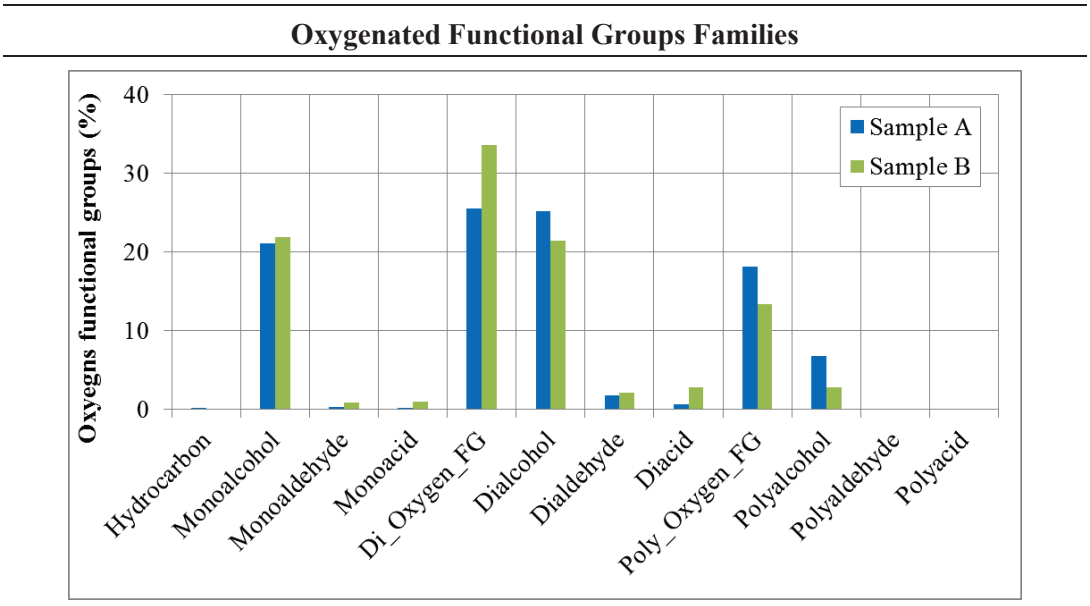


Figure X-23 – Predicted oxygenated families for (%mol/mol) building diagram #2 from sample A and sample B.

When comparing both samples, Figure X-23 also shows that sample A has more dialcohols and polyalcohols and less monoacids and diacids than sample B, which is in accordance with the ^{13}C NMR data (report to Table X-1). However, and like the results of BD#1, it should be expected a higher amount of monoalcohols, monoaldehydes and dialdehydes in sample A than sample B. Regarding the "Di_Oxygen_FG" and "Poly_Oxygen_FG" families, it is difficult to analyze, because the exact FG are unknown. However, and as mentioned previously, the first family may include molecules with carboxylic acids, which would justify the higher amount of these in sample B than in sample A (report to ^{13}C NMR data in Table X-1). As for the "Poly_Oxygen_FG" family, the inclusion of methoxyls in the cores, would lead to an increase of this family for sample A, as it has a higher amount of these than sample B (report to ^{13}C NMR data in Table X-1).

To conclude this section, Figure X-24 presents several examples of molecules reconstructed by BD#2 for both samples. Methanol (*a)*) is the smaller molecule generated by the SR, while the *l)* compounds are one of the largest. As seen before, BD#2, generates small paraffins, in addition to monomers, dimers and trimers with maximum four (A) / seven (B) cycles per core. This diagram can also recreate several functional groups and the intense presence of oxygen in the molecules. The low molecular weight molecules (a, b, c, d, e, f, h and i) were already seen in the open literature (Staš et al. 2014) and are in agreement with the type of molecules expected for bio-oil. The high molecular weight molecules (j, k and l), although extremely difficult to identify through analytical techniques, are in accordance with the structures proposed for the macromolecules existent in a bio-oil mixture (Bayerbach and Meier 2009) (Xu et al. 2014).

In summary, and from a mixture point of view, both libraries (sample A and sample B) are composed mainly of n-paraffins and small aromatic cyclic compounds. Furthermore, most of these molecules have only one to three oxygen atoms, which reduces the oxygen diversity of the typical functional groups of bio-oil. Although most of the features predicted by both compositional models correspond the information in the open literature (Bayerbach and Meier 2009) (Staš et al. 2014), certain characteristics, as the size of the molecules and the diversity of the oxygen groups, is still missing. A desired library should present more high molecular weight compounds, which by increasing the number of potential substituted carbons, would increase the number of oxygen (better FT-ICR/MS results) and the diversity of the functional groups (better Oxygen functional group families results) per molecules. For these reasons, the libraries created by the BD#2 for sample A and sample B do not represent the totality of a bio-oil fraction.

X.3.2.1 Reconstruction by entropy maximization

Similar to the first hypothesis, the libraries created by the SR method for BD#2 do not represent the totality of a bio-oil fraction. As explained before, in the presence of an unsuitable library of molecules, the method would overestimate the molar fractions of certain molecules and underestimate others. Therefore, it was decided not to use the reconstruction by entropy maximization.

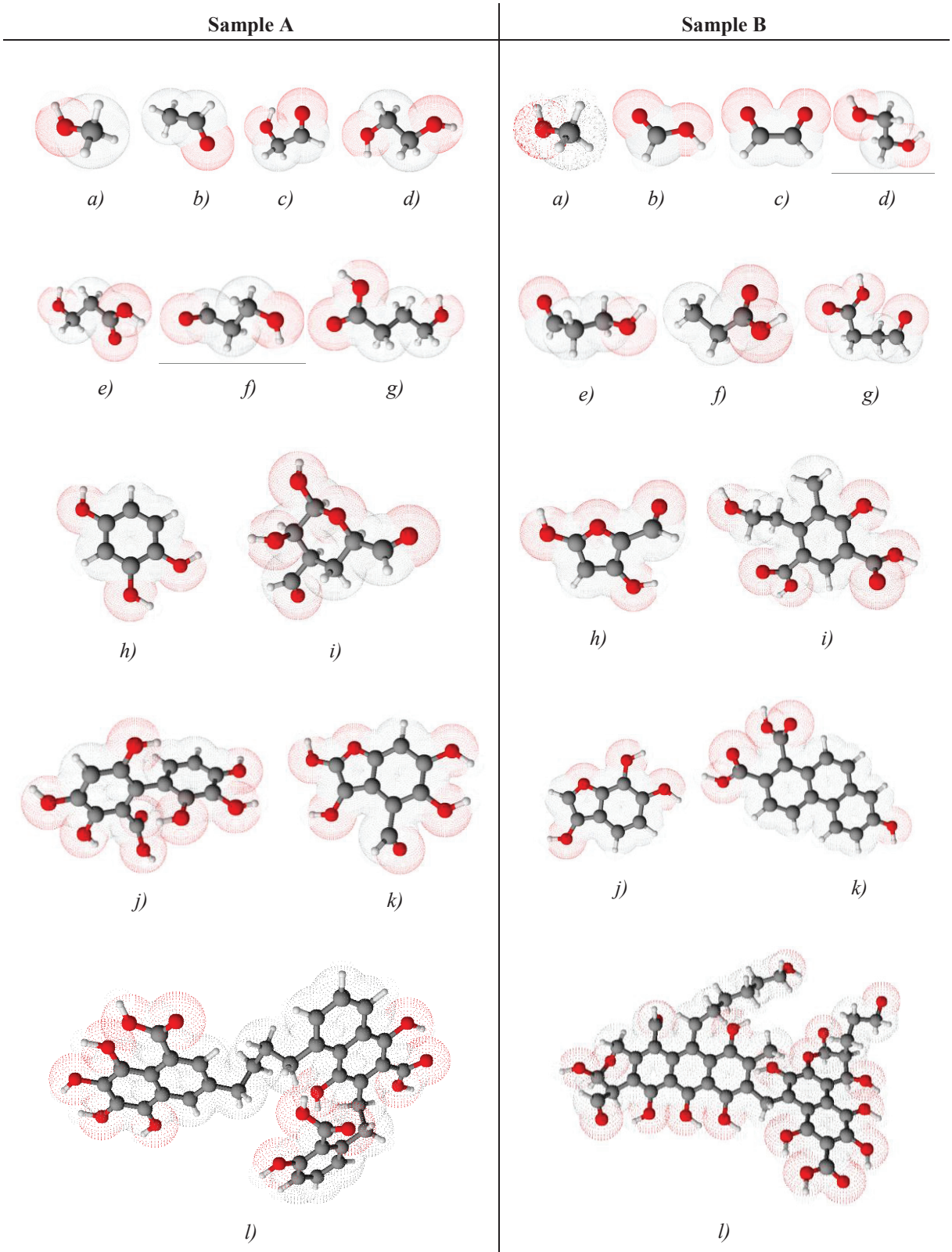


Figure X-24 - Example of reconstructed molecules from BD#2 for sample A and sample B: black – carbon, white – hydrogen, red - oxygen (report to appendix E for the molecules names)

X.3.3 Comparison of the building diagrams

Here both building diagrams will be compared for each sample. This section will focus on the estimated properties, the predicted properties and the chemical structures generated for each BD and sample. Through this comparison it is intended to defined a library of molecules, which may later on applied in the SSA for the simulation of the hydrotreating of bio-oil.

Estimated properties

The parameters present in Table X-4 and Table X-7 were optimized thanks to the experimental properties used as input: Elemental analysis (CHONS), ^{13}C NMR analysis and SEC analysis (report to Table X-2 and Figure IX-6). Here is presented a comparison of the fitting properties between both BDs for each sample.

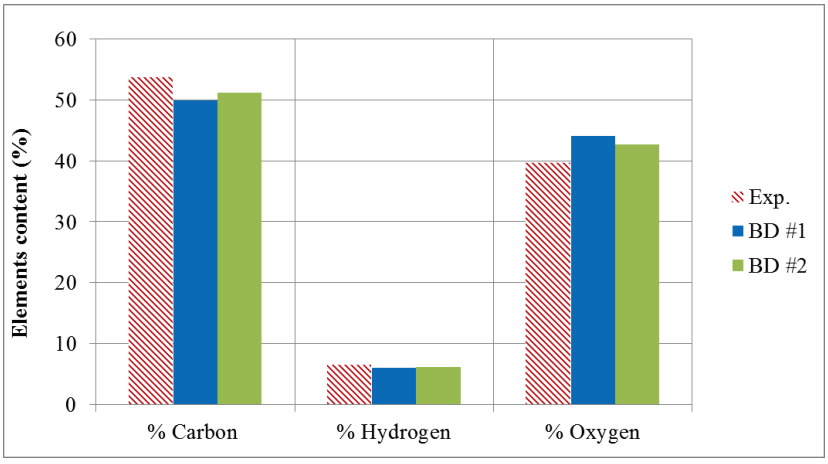
For the **elemental analysis**, both BDs present similar and reasonable results, as seen in Figure X-25. For **sample A**, both diagrams predict the same content of hydrogen and almost the same content of carbon and oxygen. As shown in Figure X-25 *a)*, the content of carbon is underestimated, while the content of oxygen is overestimated. However, and as seen in Figure X-10 and Figure X-19, these deviations are within 10% error area, as seen.

For **sample B**, BD#2 predicted quite well the content of all elements as shown in Figure X-25 *b)*. On the contrary, BD#1 underestimates the content of carbon and overestimates the content of oxygen, even if it able to predict the content of hydrogen. Be that as it may, and as seen in Figure X-10 and Figure X-19, these deviations are within 10% error area, as seen.

Between both BDs, BD#2 estimates reasonably better the elemental analysis than BD#1. This can be justified by the global error (report to appendix D), which is 8 % for BD#1 in both samples and 6 % and 2 % for BD#2 in sample A and sample B, respectively.

Elemental Analysis

Sample A



a)

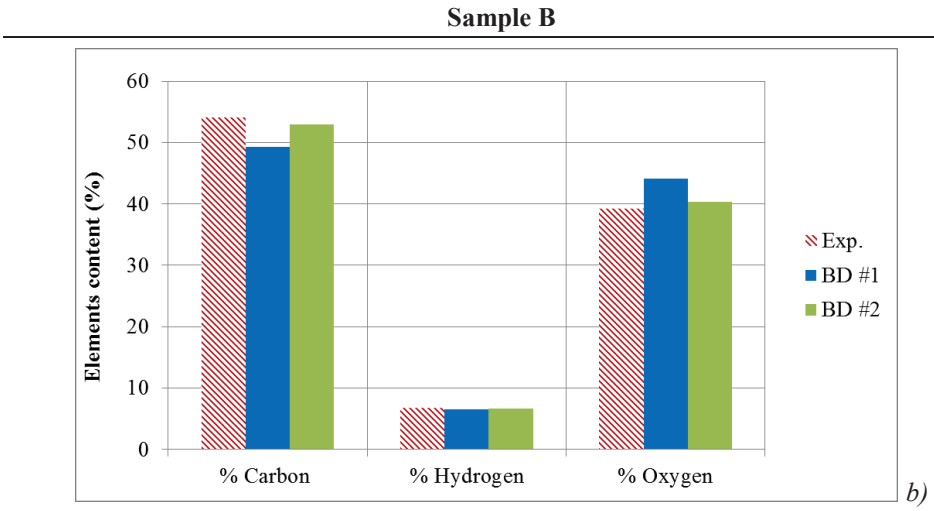


Figure X-25 – Comparison of the results of the elemental analysis for both diagrams: a) Sample A; b) Sample B.

For the **^{13}C NMR**, the precision of both BDs varies according to the sample as seen in Figure X-26. For **sample A**, both diagrams predict the same content of aldehydes and ketones, carboxyls and almost the same content of saturated carbons. However, BD#1 predicts more hydroxyls and methoxyls, while BD#2 predicts more aromatic carbons. As shown in Figure X-26 *a)*, the content of aldehydes and ketones and finally carboxyls are well predicted by both BDs. However, both diagrams also underestimate the content of hydroxyls and methoxyls and overestimate the content of aromatic and saturated carbons. Concerning the content of hydroxyls and methoxyls, and as also seen in Figure X-9 *f)* with Figure X-18 *g)*, the probability of sampling a methoxyl is the same, but the probability of sampling a methyl is higher in BD#1. This means that BD#1 can generate more hydroxyls than BD#2. For this reason, and as shown in Figure X-26 *b)*, the content of hydroxyls and methoxyls is higher for BD#1 than BD#2. Regarding the aromatic carbons, BD#2 generates larger molecules (report to Figure X-9 *c)* and Figure X-18 *d)*) , which are usually aromatic compounds. For this reason, the content of aromatic carbons is higher for BD#2 than BD#1. For the saturated carbons, although both diagrams overestimate their content, the slight difference between the BDs may be associated to the stochastic nature of the method, not being representative in here.

For **sample B**, both diagrams predict the basically same content of aldehydes and ketones and carboxyls. However, both diagrams overestimate the content of saturated carbons and underestimate the content of hydroxyls and methoxyls. As shown in Figure X-26 *b)*, the content of saturated carbons is higher for BD#2 than BD#1, unlike in sample A. This difference may be associated to the higher probability of sampling alkyl chains associated to BD#2 for sample B, as previously seen in Figure X-9 *f)* with Figure X-18 *g)*. Concerning the content of hydroxyls and methoxyls, and as also seen in Figure X-9 *f)* with Figure X-18 *g)*, the probability of sampling a methoxyl is the same, but the probability of sampling a methyl is higher in BD#1. This means that BD#1 can generate more hydroxyls than BD#2. For this reason, and as shown in Figure X-26 *b)*, the content of hydroxyls and methoxyls is higher for BD#1 than

BD#2. Regarding the aromatic carbons, the two BDs overestimate the carbons content. Be that as it may, both diagrams can almost predict accurately the content of the aromatic carbons. As shown in Figure X-11 and Figure X-20, the deviations associated to these carbons are within 10% error area, as seen.

Between both BDs, BD#1 estimates reasonably better the ^{13}C NMR analysis than BD#2. This can be justified by the global error (report to appendix D), which is 16 % and 5 % for BD#1 in sample A and sample B, respectively and 17 % and 11 % for BD#2 in sample A and sample B, respectively.

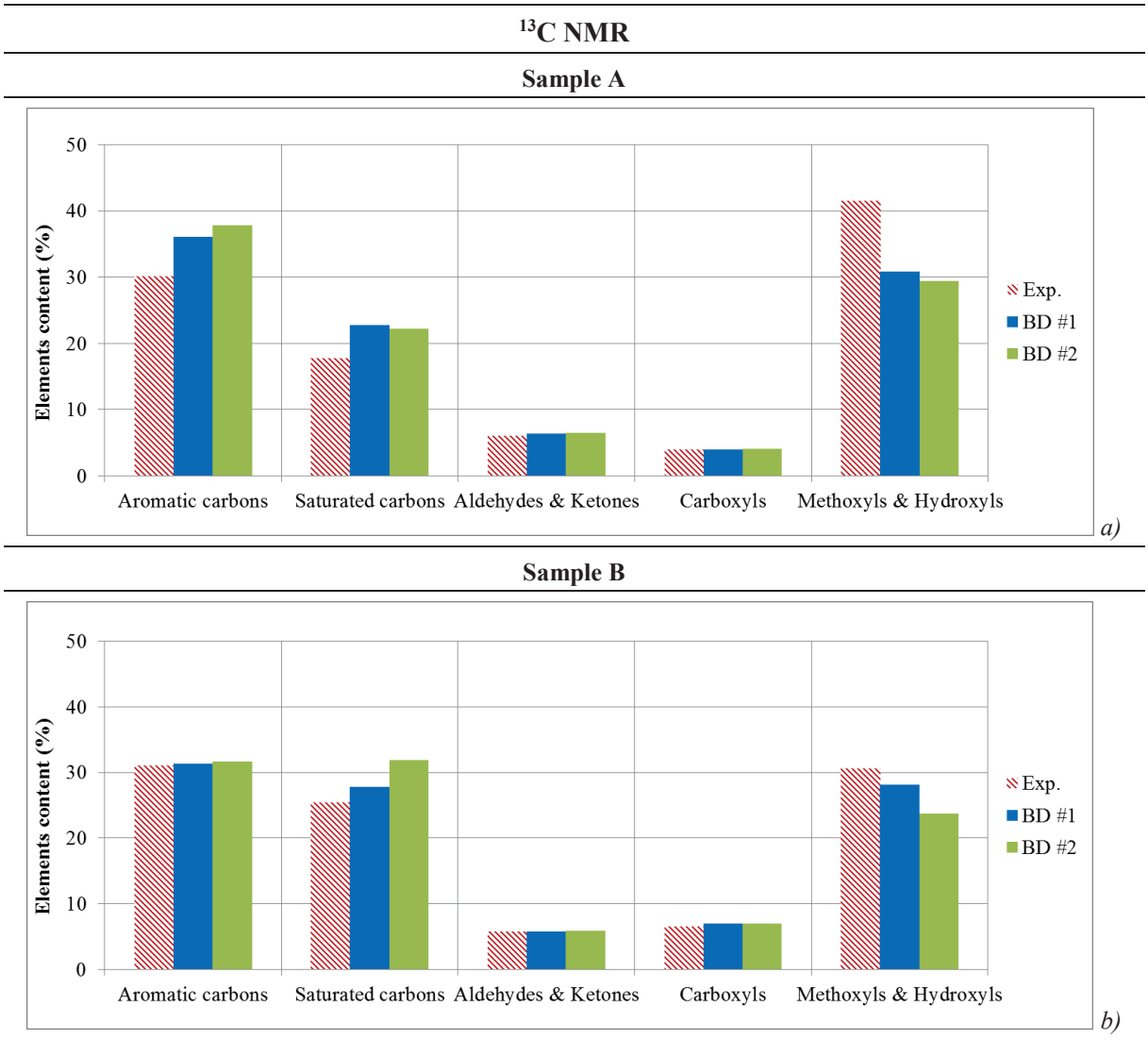


Figure X-26 – Comparison of the results of the ^{13}C NMR for both diagrams: a) Sample A; b) Sample B.

For the **SEC**, both BDs and both samples present similar results, as seen in Figure IV-1. As mentioned before, between 10 % and 50 % the results for both BDs and samples predict reasonably well the polystyrene equivalent molecular weight. The existent deviations may be associated to the experimental error and also to the correlation applied (report to Equation IX-9). For fraction higher than 60 %, both BDs underestimate the results for both samples. This divergence is mostly due to the library of molecules. As mentioned before, all the libraries so far created lack high molecular weight molecules. These libraries have mostly small paraffins and cores with a few numbers of rings.

Between both BDs, BD#2estimates reasonably better the SEC analysis than BD#1. This can be justified by the global error (report to appendix D), which is 44 % and 45 % for BD#1 in sample A and sample B, respectively and 46 % and 40 % for BD#2 in sample A and sample B, respectively.

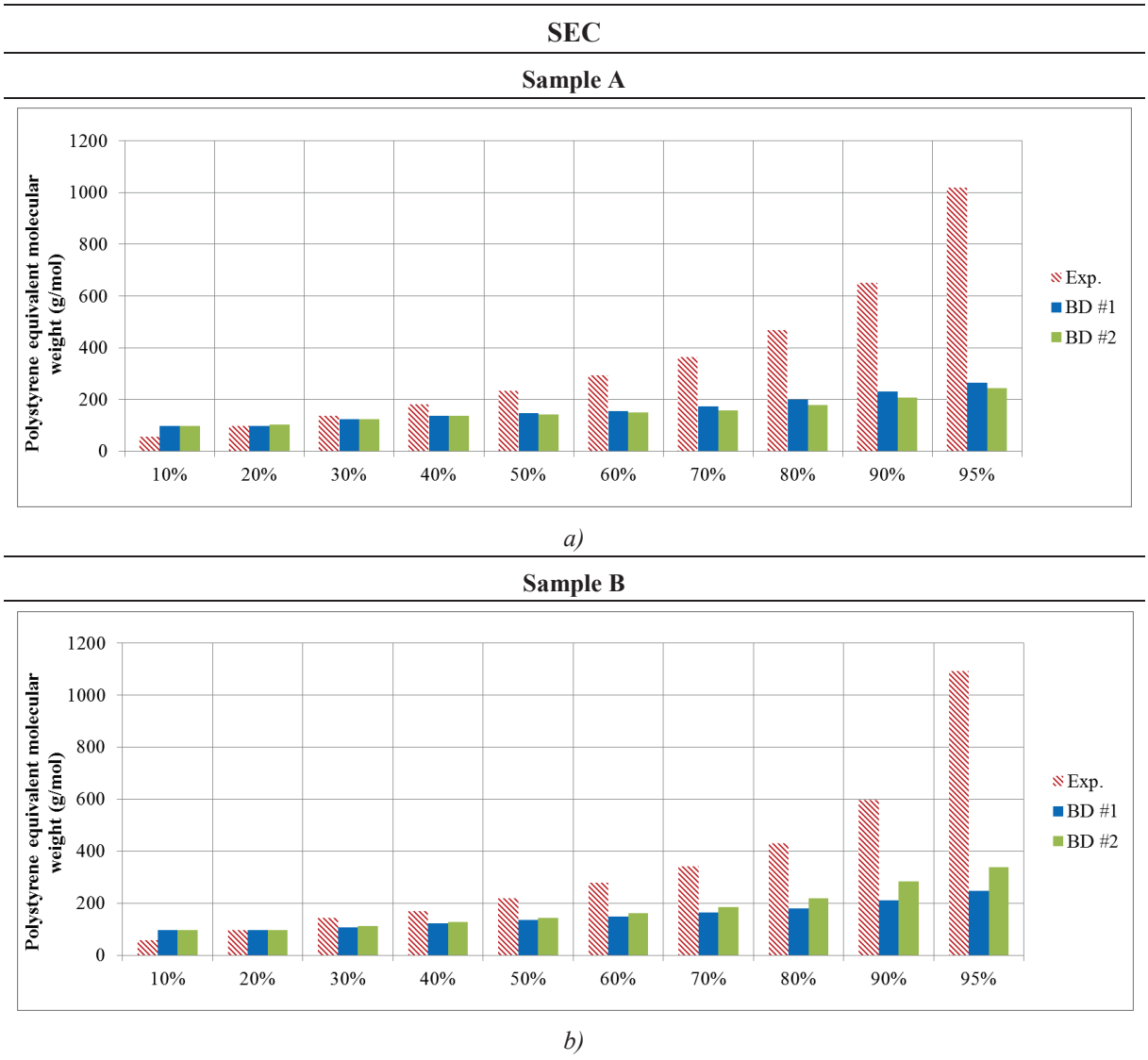


Figure X-27 – Comparison of the results of the ^{13}C NMR for both diagrams: a) Sample A; b) Sample B.

Predicted properties

As mentioned before, beside the properties used in the estimation of the parameters, the SR can predict furthermore properties. Here is presented a comparison of the predicted properties between both BDs for each sample.

For the **FT-ICR/MS analysis**, both BDS present an equal trend for both samples, as shown in Figure X-28. As seen before, although both samples present similar results, the trend does not correspond to the experimental results of FT-ICR/MS in any of the cases. As shown in Figure X-28, both BDs reconstructed mostly libraries with molecules with one to three oxygens, while the experimental data has mainly a homogenous mixture, where most molecules have between four to ten oxygens. As already

mentioned, the absent of oxygen in the generated libraries maybe associated to high fraction of n-paraffins. Even so, both BDS can generate molecules with four to thirteen (A) / twelve (B) oxygens for both samples. These correspond to the cyclic compounds reconstructed by the algorithm. The slight difference between both samples for each BD, may be associated to the stochastic nature of the method, not being representative in here.

Between both BDs, BD#2 predicts reasonably better the FT-ICR/MS analysis than BD#1. This can be justified by the global error. As report to appendix D, the global error of BD#2 for both samples is around about 25 % minor than the global error of BD#1 for both samples.

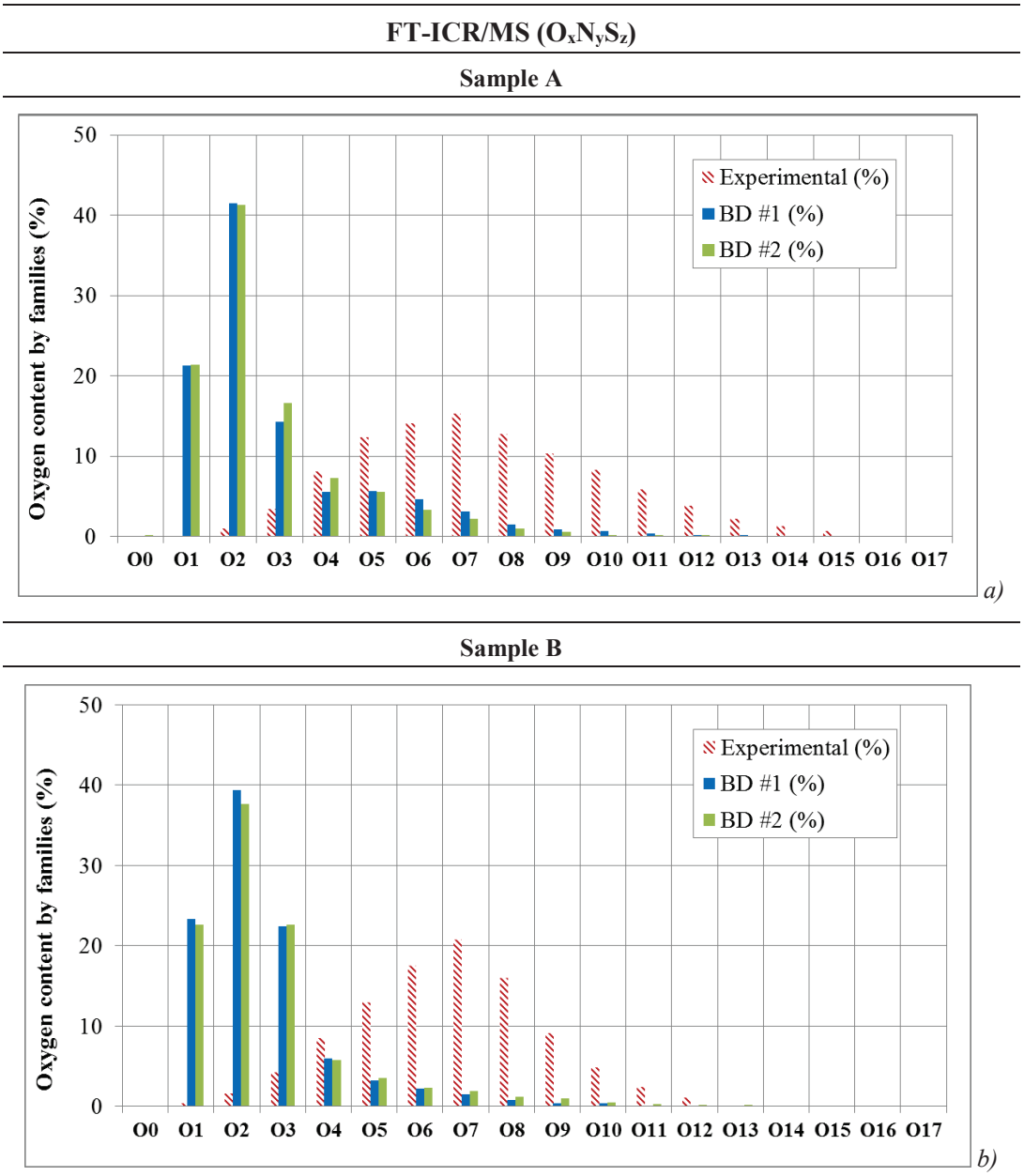


Figure X-28 - Results for the FT-ICR/MS analysis (%mol/mol) from both building diagram: a) Sample A and b) Sample B.

For the **density at 20 °C**, the both BDs predicted for both samples a higher density value than the experimental data. As seen before, all the predicted results are beyond the statistical errors of the method (average relative error of 1.35 % and standard deviation of 0.04 g/cm^3). Besides non-conformities that may be associated with the method, the most plausible hypothesis for the deviations is the libraries generated from both BDs for both samples. As discussed previously, the lack of cyclic molecules, especially macromolecules, increases the average density above the desired value. Be that as it may, and as shown in Table X-9, BD#2 predicts better the density, in the case of sample A, whereas BD#1 predicts better than density in the case of sample B. However, for sample B, the difference between the density predicted by both BDs is minor enough to be negligible. Between both BDs, BD#2 predicts reasonably better the density than BD#1. This can be justified by the global error, which as reported in Table X-5 and Table X-8, is globally inferior for BD#2.

Table X-9 - Density and molecular weight predicted by both building diagrams for sample A and sample B.

		Sample A		Sample B	
Property	**Value type**	**BD #1**	**BD #2**	**BD #1**	**BD #2**
Density (g/cm^3)	Experimental	1.24		1.19	
	Predicted	1.31	1.28	1.23	1.24
Molecular weight (g/mol)	Predicted	103.4	99.8	90.4	106.5

Chemical properties

As presented before, both BDs reconstruct mostly linear molecules and few cyclic molecules, which have mainly an aromatic character. In Figure X-29, is shown the PIONA families (Paraffins, Isoparaffins, Olefins, Naphthenes and Aromatics) predicted by the algorithm for each BD and each sample. As seen, sample B has more linear molecules than sample A for both diagrams. However, this may be justified by the higher amount of saturated carbons in sample B than sample A (see Table X-2), which forces the algorithm to generate more linear molecules. When comparing both building diagrams, Figure X-29 shows a slight decrease in the linear molecules content, together with an increase in the aromatic content. The content in other compounds (saturated monomers) is not only constants for in both BDs and samples, but also extremely low and thus, negligible. The reduction of the linear molecules content is associated with the addition of a new attribute (***type of molecule***) and the modifications in the ***number of cores*** and ***number of rings***. These forced the algorithm to reconstruct more cyclic molecules with a higher molecular weight. In summary, BD#2 promotes the increase of the molecules size as desired.

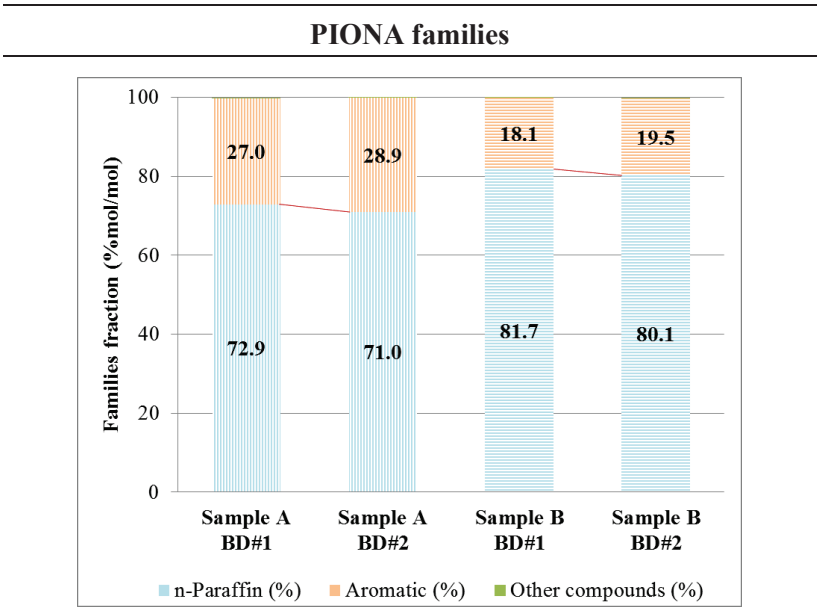


Figure X-29 – Predicted PIONA families from both building diagrams for sample A (vertical lines) and sample B (horizontal lines).

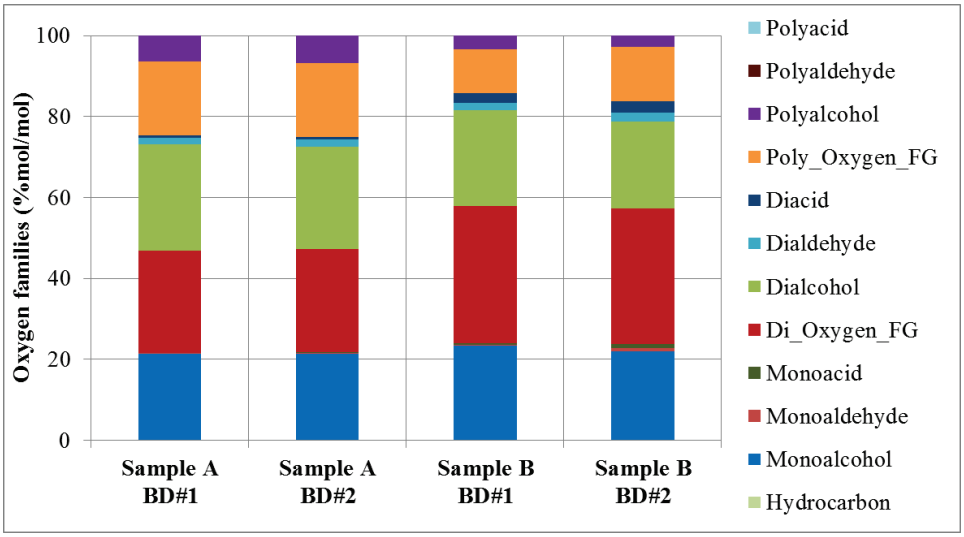


Figure X-30 - Predicted oxygenated families from both building diagrams for sample A and sample B.

Regarding the families of oxygens, monoalcohols, dialcohols, polyalcohols and molecules with two of more distinctive functional groups are the main compounds in the reconstructed mixtures. These also have a fraction of dialdehyes and diacids, in addition to a minor fraction of monoaldehydes and monoacids. The remaining oxygenated families fractions is extremely low and thus, negligible. As discussed before, both BDs generate for both samples mostly n-paraffins, which can only be replaced maximum twice. Therefore, it would be wise to assumed that most of these monoalcohols, dialcohols and molecules with two FG are indeed *n*-paraffins. The residual traces of mono and dialdehydes and

mono and diacids are included as well. As for the molecules with three or more FG, these can only be cyclic compounds, for these accept further than two FG. When comparing both building diagrams, Figure X-30 reveals the same pattern for both. The only minor difference resides in a higher content of mono and dialdehydes and mono and diacids in BD#2. However, these are mainly associated with the paraffins, which are the major compound fraction generated by BD#2.

X.4 Conclusion of Chapter X

For two bio-oil samples (A and B) and two different building diagrams, the stochastic reconstruction method was applied. The compositional models generated were fitted using the elemental analysis, ^{13}C NMR analysis and the SEC analysis. The performance of these was the evaluated thanks to the FT-ICR/MS analysis and the density of the synthetic mixtures. Furthermore, the molecular structures of synthetic mixture (library of molecules) were likewise analyzed and compared with the open literature in order to further validate the composition models.

For BD#1, the compositional model estimated reasonably well the elemental analysis, the ^{13}C NMR analysis and the initial fraction of the SEC analysis. The reaming experimental points of this analysis presented deviations, mostly associated to the lack of high molecular weight compounds in the libraries. Regarding its validation, although the mixtures present were few oxygens per molecule, the density of the mixtures was close to the expected value.

For BD#2, similar results of BD#1 were found. The compositional model estimated reasonably well the elemental analysis, the ^{13}C NMR analysis and the initial fraction of the SEC analysis. Regarding the higher fractions of SEC analysis, even if BD#2 has a higher probability of creating larger molecules, these were underestimated, still due to the lack of high molecular weight compounds in the libraries. Concerning the validation, BD#2 also reconstructed libraries with few oxygens, but with a density closed to the expected value.

Besides the lack of high molecular weight compounds, other features miss in the libraries. As presented for both diagrams, a larger diversity of functional groups should exist, in addition to more furan and THP rings. For these reasons, the entropy maximization method was not applied to none of the libraries. In the presence of an unsuitable library of molecules, the method would overestimate the molar fractions of certain molecules and underestimate others, which is not desirable.

Between both BDs, it was found that BD#2 had less deviations and could estimate and predict better the properties than BD#1. Furthermore, this diagram promotes the formation of high molecular weight molecules. These increase the number of potential substituted carbons, and therefore the number of oxygens per molecules and the diversity of the functional groups, as desirable.

To conclude, although both BDs present certain shortcomings, the compositional models here created can reconstruct a library which is able to represent the known and light fraction of bio-oil. For this reason, the results originated from BD#2 will be applied for the simulation of the hydrotreating of bio-oil.

Chapter XI. Simulation of bio-oil hydrotreating

Thanks to the kinetic model previously created by the SSA and the synthetic mixture of bio-oil generated by the SR, it is now possible to simulate the hydrotreating of a real bio-oil fraction. The simulation will be done by the SSA approach and intents to represented the trends of the deoxygenation of the mixture and its effluents. With the present chapter, the main objective of the thesis will finally be achieved.

For a better understanding of the simulation procedure and the application of the previous work, the following chapter is divided into five sections. The experimental data, which will present the experimental information of bio-oil hydrotreating; the molecular library, which will summarize the properties of the synthetic mixture of bio-oil and how will this library be applied; the reaction, which will present the reactions typical of the hydrotreating of bio-oil and those here considered; the kinetic model, which will summarize the kinetic parameters here estimated and their application in the reactions and finally, the results of the simulation of the hydrotreating of a real bio-oil fraction.

XI.1 Experimental data

As presented in the Experimental approach section, six experimental tests of the hydrotreating of bio-oil (at constant water content) were performed by Ozagac (Ozagac 2016). Similar to the model compounds, these were obtained in a batch reactor under 13 MPa of hydrogen over a reduced NiMo/γ-Al_2O_3 at a range of temperature between 200 °C to 300 °C and a range of reaction time between 1 h to 3 h. A blank test (0 h) at 250 °C was also done.

Contrary to the model molecules, the validation of the simulation results was done thanks to the ^{13}C NMR analysis performed by Ozagac (Ozagac 2016) to the effluents each hydrotreating test. In addition to this analysis, the author also did an elemental analysis to each test. The results of both analyses are presented in Table VI-1. It is important to highlight that each experimental test had an organic and an aqueous phase. At the moment, the SSA cannot reproduce this characteristic. For this reason, and for each test, the experimental data of the organic and aqueous phases were sum, in order to later compare with the results of the SSA for the hydrotreating of bio-oil.

Concerning the temperature profiles, and as explained before, the simulation process needs the temperature profile for each. Therefore, for all bio-oil tests, the temperature extracted from the pilot unit need to be implemented as a simulation input, together with the process conditions, as pressure (isobaric process) and the time of each test, given by the last value in the profile. The time of an experimental test is the time of the simulation.

Table XI-1 - Experimental ^{13}C NMR and CHONS results concentration used in the simulation of the hydrotreating of bio-oil.

Temperature (°C)	**200**	**250**				**300**
Reaction time (h)	**1**	**0**	**0.5**	**1**	**3**	**1**
^{13}C NMR	**Fraction (%w/w)**					
Saturated	32.0	26.5	33.0	32.3	30.7	50.0
Aromatics	18.6	22.0	23.0	27.7	32.5	17.0
Aldehydes & Ketones	4.6	8.3	0.4	1.8	1.9	2.0
Carboxylic acids & Esters	24.4	30.0	24.6	21.9	22.8	20.1
Hydroxyls, Ethers, Methoxys	20.3	13.2	18.9	16.3	12.1	10.9
CHONS	**Fraction (%w/w)**					
Carbon	57.7	54.4	57.9	58.0	59.6	58.9
Hydrogen	6.7	6.6	6.8	7.5	7.6	7.4
Oxygen	35.2	38.6	34.8	34.0	32.2	33.3
Nitrogen	0.4	0.4	0.4	0.5	0.5	0.4
Sulfur	0.0	0.0	0.0	0.0	0.0	0.0

XI.2 Molecular library

Concerning the number of molecules to simulate, the number of iterations and the number of registration points, the same values used for the model compounds were assumed: Ten thousand molecules (10 000), fifth iterations (50) and one hundred and ninety registration points (190). However, the molecular library generated by the SR only contains five thousand molecules (5 000). To overcome this obstacle, each molecule in the library was duplicate, thus generating a library with ten thousand molecules (10 000). This procedure does not change the properties of the synthetic mixture, since the duplication is homogenous for all molecules, meaning that the fraction of each molecule in the mixture is maintained.

For the simulation of bio-oil, the SSA will use the library created by BD#2 for sample B. As mentioned before in Chapter X, BD#2 is the best to represent a real bio-oil fraction. Regarding the samples, although BD#2 has a lower objective function for sample A (report to section X.3.1.2 and section X.3.2.1), there is not any information available for the hydrotreating of this sample. On the contrary, sample B underwent upgrading by hydrotreating (Ozagac 2016), thus this sample has useful information for the simulation of the hydrotreating of a bio-oil fraction. This information is summarized in Table VI-1, while the data of the synthetic mixture is presented in Table XI-2, together with the experimental data of the original bio-oil sample (B).

Table XI-2 - Experimental data for sample B of bio-oil and the reconstructed mixture of bio-oil B.

Property		Original bio-oil	Reconstructed bio-oil
Density (20 °C) (g/cm^3)		1.19	1.24
Elemental analysis (dry basis) (%)	Carbon	54.12	52.91
	Hydrogen	6.67	6.69
	Oxygen	39.21	40.40
	Nitrogen	0.00	0.00
	Sulfur	0.00	0.00
^{13}C NMR (%)	Saturated	25.50	31.70
	Aromatics	31.20	31.82
	Aldehydes & Ketones	5.90	5.84
	Carboxylic acids & Esters	6.70	6.94
	Hydroxyls, Ethers, Methoxys	30.70	23.70

XI.3 Reaction network

As presented in the State of the Art, bio-oil can undergo several reactions. These depend not only on the molecular structure of the compounds in the mixture, but also on the temperature of reaction, catalyst and the presence of a solvent, together with its type. Table I-19 in the State of the Art chapter summarized all the reactions accounted in the open literature for the hydrotreating of bio-oil and its model compounds.

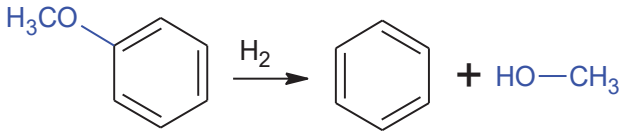
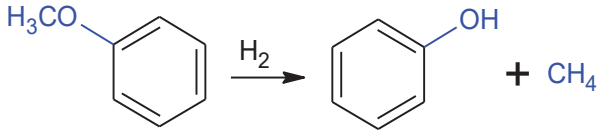
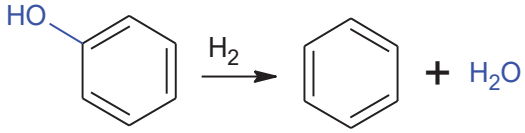
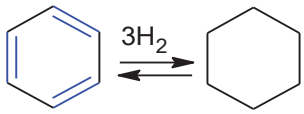
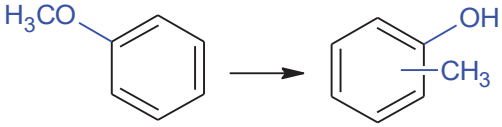
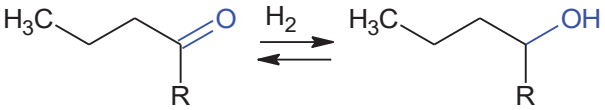
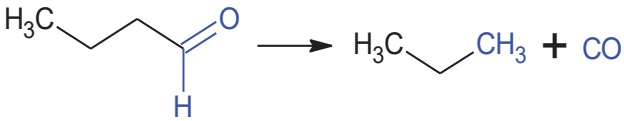
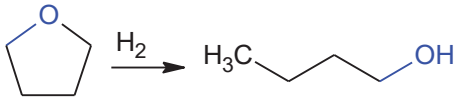
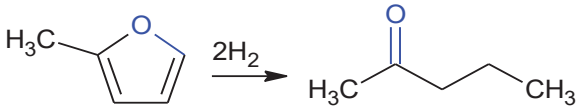
As shown in the unrolling of the present thesis, all the desirable reactions in the hydrotreating process were implemented in the source code (except hydrogenolysis). These are: Saturation, demethylation, demethoxylation, transalkylation, hydrodeoxygenation, ring opening and hydrogenation of the carbonyl, carboxyl and unsaturated bonds. Regarding the undesirable reactions, all gas reactions, decarbonylation, decarboxylation, retro-aldol, dehydration, hydration, aromatization and oxidation of a hydroxyl were also implemented in the source code. The remaining reactions were not added due to their bimolecular nature (not yet implemented in the SSA). Although several reactions were implemented in the source code, not all were used until now and were used for the hydrotreating of bio-oil. The simulation of the synthetic mixture of bio-oil was performed only with the reactions used in the molecule compounds. All the reactions used in the simulation of the hydrotreating of bio-oil are summarized in Table XI-3.

XI.4 Kinetic model

The kinetic parameters associated to the reaction applied in the hydrotreating of bio-oil are summarized in Table XI-3. These parameters correspond to those estimated via the SSA for the hydrotreating of guaiacol and furfural. Concerning the hydrodeoxygenation and saturation, both reactions exist in the

reaction network of both model compounds. These have different kinetic parameters for guaiacol and furfural. For the simulation of the synthetic mixture, it was selected the parameters estimated for the hydrotreating of guaiacol. As seen before, the library of molecules created for the bio-oil mainly benzene rings (report to Table X-7 in Chapter X). This is a characteristic of guaiacol, not of furfural, thus the parameters of guaiacol are more appropriate for the present simulation.

Table XI-3 - Reactions and their kinetic parameters for the HDT of bio-oil in the SSA

Name	Reaction	k_0 (/s)	Ea (kJ/mol)
Demethoxylation (DEMO)		1 500	75
Demethylation (DEM)		500	82
Hydrodeoxygenation (HDO)		50	82
Saturation (SAT)		100	70
Transalkylation (TRANS)		1 000	75
Carbonyl Hydrogenation (COH)		10 000	56
Decarbonylation (CO)		3 000	70
Ring Opening (RO)		600	60
		600	60

XI.5 Stochastic simulation results

As mentioned before, six experimental tests of the hydrotreating of bio-oil (at constant water content) were performed by Ozagac (Ozagac 2016). Similar to the model compounds, these were obtained in a batch reactor under 13 MPa of hydrogen over a reduced NiMo/γ-Al_2O_3 at a range of temperature between 200 °C to 300 °C and a range of reaction time between 1 h to 3 h. A blank test (0 h) at 250 °C was also done. All these tests were simulated via the SSA method according to the kinetic parameters previously obtained for the model compounds. The results obtained are presented in the following section divided into three: molecular detail of the effluent, hydrodeoxygenation performance and finally global properties of the effluent. The results of the molecular detail of the effluent are presented in a parity plot with an error area of 10 %, similar to the model molecules

Before advancing any further, it is important to note that at the time were the analyzes were done, the ^{13}C NMR instrument presented faults at the calibration level of the C-O bonds. This situation also included the analyses of the bio-oil sample A and B. However, both samples were re-analyzed after re-calibration, while the effluents of hydrotreating. Due to the nature of these (high reactivity and several phases), plus the unavailability of the NMR, it was decided not to re-analyses the samples. Therefore, the results here presented are associated to a slight calibration error.

XI.5.1 Molecular detail of the effluent

The results of the simulation were compared with the experimental data obtained by Ozagac (Ozagac 2016). For the elemental analysis, the stochastic simulator can predict with the error area all the elements content, as shown in Figure XI-1 a) and c). Be that as it may, the oxygen content is slight overestimated, which indicates a low removal of oxygen potentially caused by the absence or low impact of HDO reactions. To conclude, the model here applied can reasonably predict the element content. However, the slight deviations can be due to unsuitable kinetics or even the absence of reactions and the impact of the molecule structure and the catalyst in the reactivity of the molecules.

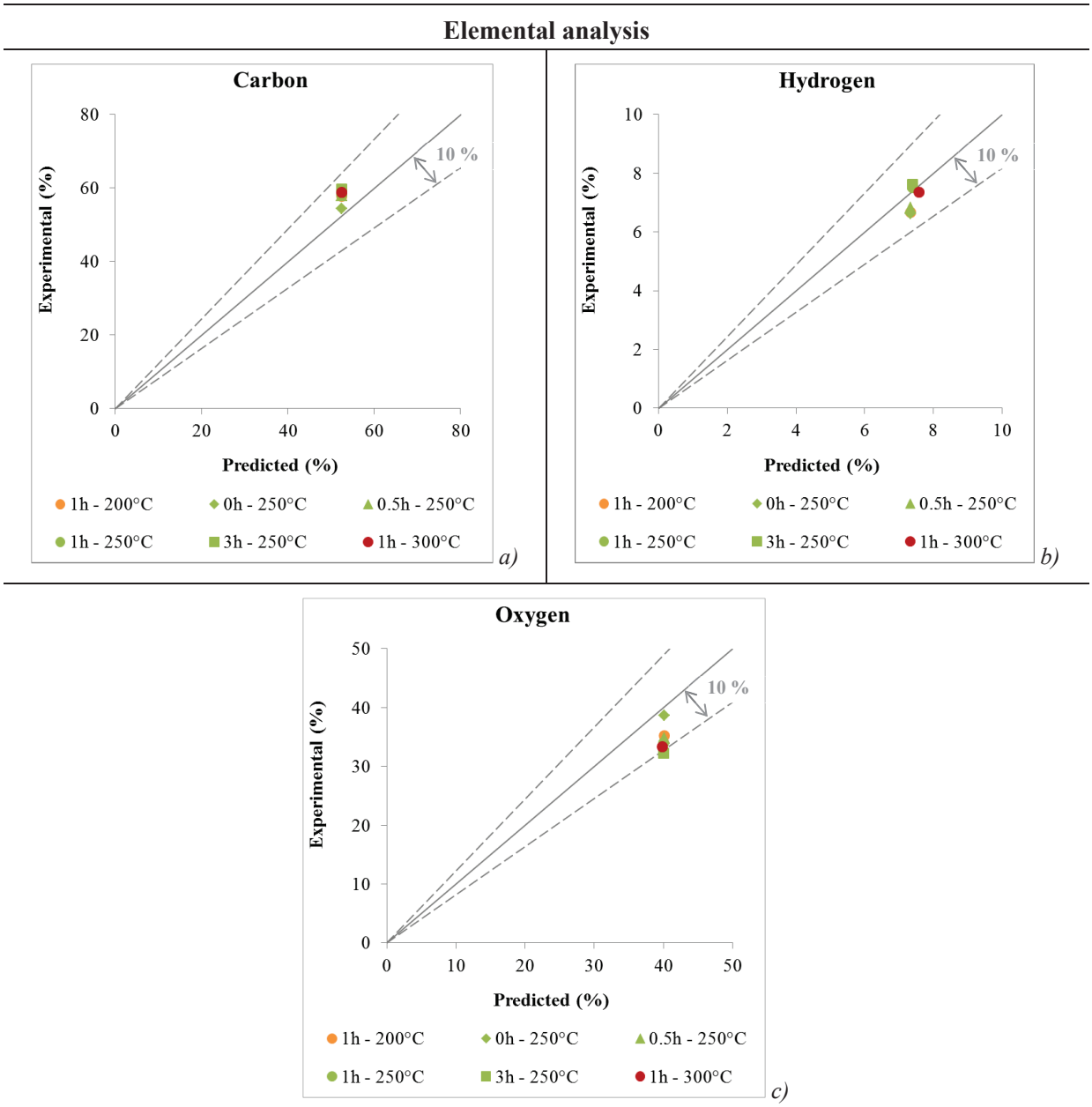


Figure XI-1 - Elemental analysis of the effluent of the simulation of the hydrotreating of bio-oil.

For the ^{13}C NMR, the SSA can predict reasonably the content of saturated and aromatic carbons, as shown by Figure XI-2. However, the simulator overestimates the hydroxyls and methoxyls, while underestimating the aldehydes, carbonyls and carboxylic acids.

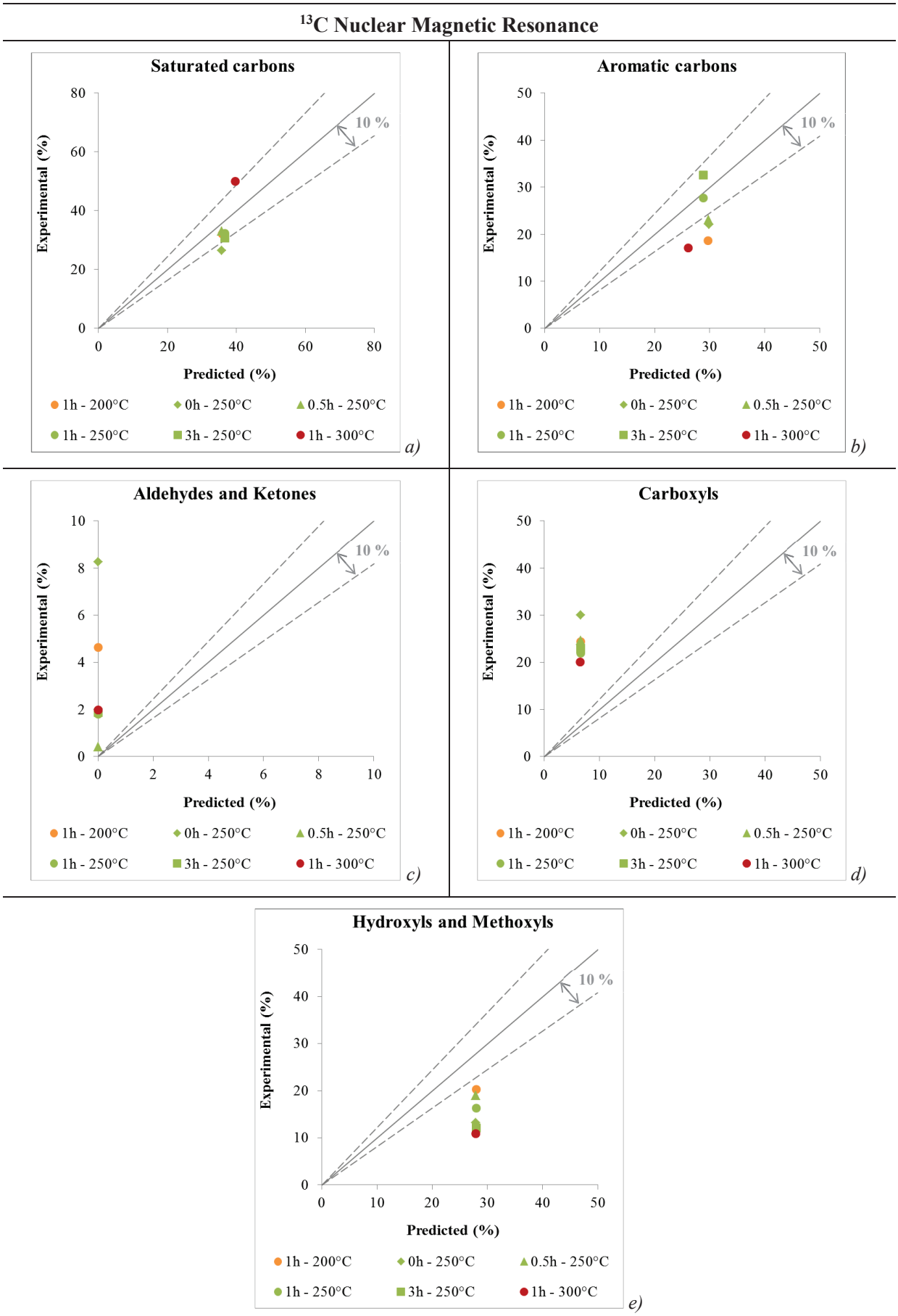


Figure XI-2 – ^{13}C NMR of the effluent of the simulation of the hydrotreating of bio-oil.

For the saturated carbons, the predicted data is within or near the area of the 10 % of error. Even so, the saturated carbons are slight overestimated at 200 °C and 250 °C. This impact is not shared by the aromatic carbons, as the content at 200 °C and 250 °C is not underestimated. For this reason, it possible to conclude that the saturation↔aromatization kinetic parameters are not the cause. In this case, the deviations may be generated by the library of molecules, which overestimated the content of saturated. On the contrary, at 300 °C, the saturated carbons are underestimated, while the aromatic carbons are overestimated. This may indicate a shortcoming in the kinetics of the saturation↔aromatization reaction. This reason, may also be responsible to the deviations found in the aromatic carbons.

For aldehydes, ketones, the SSA underestimated all the experimental data, as shown in Figure XI-2 c) and d). Although the absence of reactions can once again put into evidence together with the calibration shortcomings, the main reason for these deviations is the unsuitable kinetic parameters or even more in detailed the negligence of the molecule structure in the kinetic model. As shown in Figure XI-3, the equilibrium reaction between the hydrogenation of the carbonyl and the oxidation of the hydroxyl is the main pathway in the hydrotreating of bio-oil. This reaction represents more than 90 % in all experimental tests. As it happens both hydrogenation of carbonyl converts all the aldehydes and ketones into alcohols. These can be oxidized once again into a carbonyl. However, and as seen before (report to Figure V-4 in the Deterministic modeling methodology chapter), the equilibrium constant favor the oxidation more than the reduction (hydrogenation). Since the equilibrium of this reaction tends more to the formation of hydroxyls, the quantity of carbonyl in the effluent decreases and thus is underestimated.

The behavior explained above can also be seen in the hydroxyls content represented by Figure XI-2 e). Although these values could also be affected by the presence of methoxyls, it was seen before that the probability of sampling a methoxyls is low, so there is too few methoxyls in the mixture. For this reason, the content show in the figure corresponds basically to hydroxyls. As shown in Figure XI-2 e), the hydroxyls are overestimated due to the high reactivity of the hydrogenation of the carbonyl.

Regarding the carboxyls, the SSA underestimated all the experimental data, as shown in Figure XI-2. Since CO↔OH reaction does not affect carboxyls (code source constraint) and all the remain reaction have an extremely low impact (report to Figure XI-3), the low content of acids may be due to the shortcomings in the ^{13}C NMR calibration.

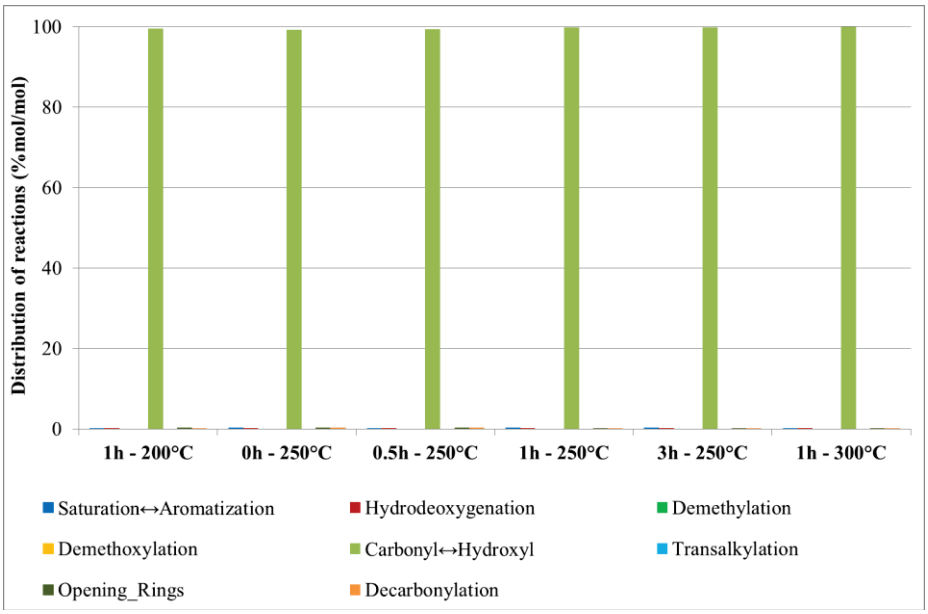


Figure XI-3 - Simulated reactions for all the experimental tests.

XI.5.2 Hydrodeoxygenation performance

Since the conversion of bio-oil cannot be analyzed, the performance of the simulator was analyzed through its hydrodeoxygenation performance. This can be given by the removal of oxygen from the molecules. As shown in Figure XI-4, the oxygen starts begin removed when the temperature hits the 150 °C during heating process. This means that bio-oil, similar to its model molecules, is extremely reactive at low temperatures, thus the importance of considering the dynamic periodic of the reactor.

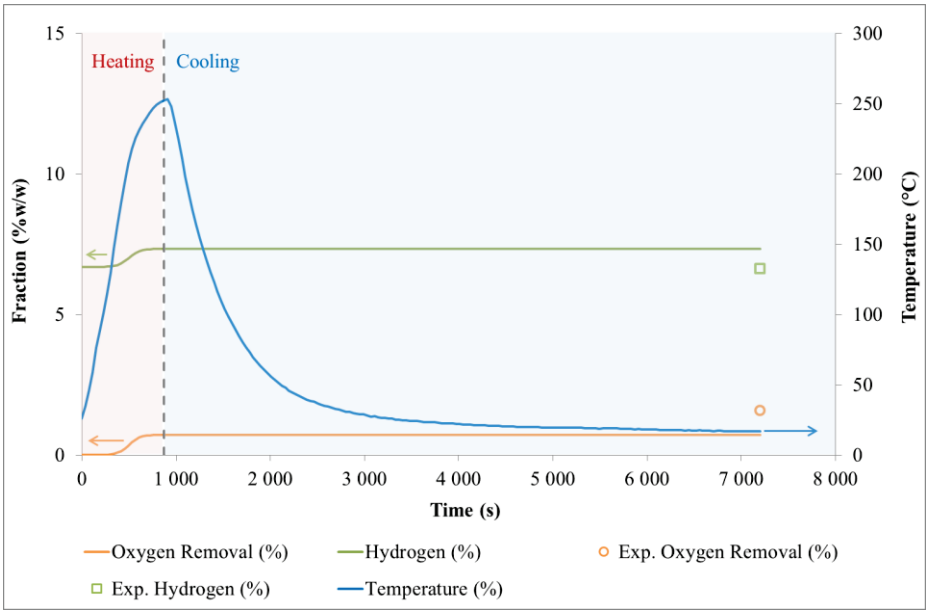


Figure XI-4 - Removal of oxygen, hydrogen content and temperature profile for the simulation of the blank test at 250 °C.

As also described by Figure XI-4, the removal of oxygen hits is maximum (around about 1 %) before the end of the heating process. Afterwards, no more oxygen is removed from the molecules. Regarding the hydrogen, it is possible to seen a light increase in the middle of the heating process. This indicates that hydrogenation reactions, as saturation, HDO or carbonyl hydrogenation, only happen at higher temperatures (≈ > 150 °C). Furthermore, and similar to the oxygen removal, the content of hydrogen stops increasing, meaning that either there are no more hydrogenation reactions, or these reactions are reversible (hydrogen content balance). However, the absence of reactions is the most likely. As seen before, CO↔OH reaction is the main pathway. This reaction does not require hydrogen, nor does it remove oxygen. For both reasons, the removal of oxygen and the content of hydrogen are constant through the rest of the test.

As shown in Figure XI-5, the behavior seen for the removal of oxygen is similar to all the simulations of the experimental tests. Occasionally, a reaction that removes oxygen may happen and increase the fraction of this heteroatom removal. This type of reactions is more predominant to occur at higher reaction times and temperatures, as reflected by the trends differences at 1 h – 250 °C, 3 h – 250 °C and 1 h – 300 °C, than at lower reaction times and temperatures. Be that at it may, the predicted evolution of the removal of oxygen does not matches the experimental HDO performance, as shown in Figure XI-5. To conclude, the time where each test undergoes the jump in the oxygen removal varies with the reactivity of each test (depends on the reaction time and temperature.)

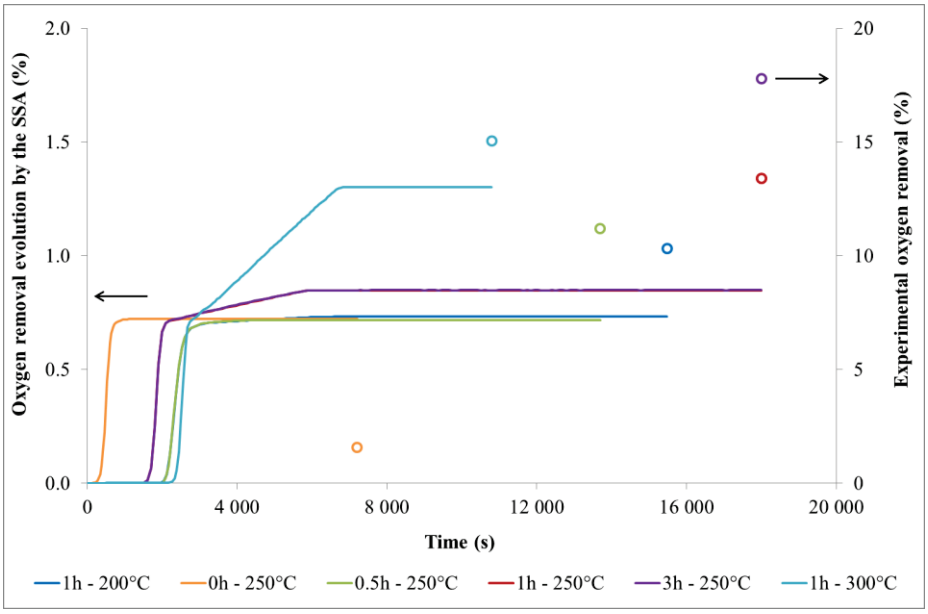


Figure XI-5 - Removal of oxygen in all the experimental tests: Full lines - Simulation, Circles - Experimental point.

Figure XI-6 presents the reactions which occur during the simulation apart the CO↔OH reaction. As represented, there is no demethylation, demethoxylation and transalkylation reactions. These are associated to the methoxyl groups, but, because the library of molecules used as bio-oil feed barely

contains any methoxyl (report to Figure X-18 in Chapter X), these reactions are not probable to happen and thus result in existent reaction pathways.

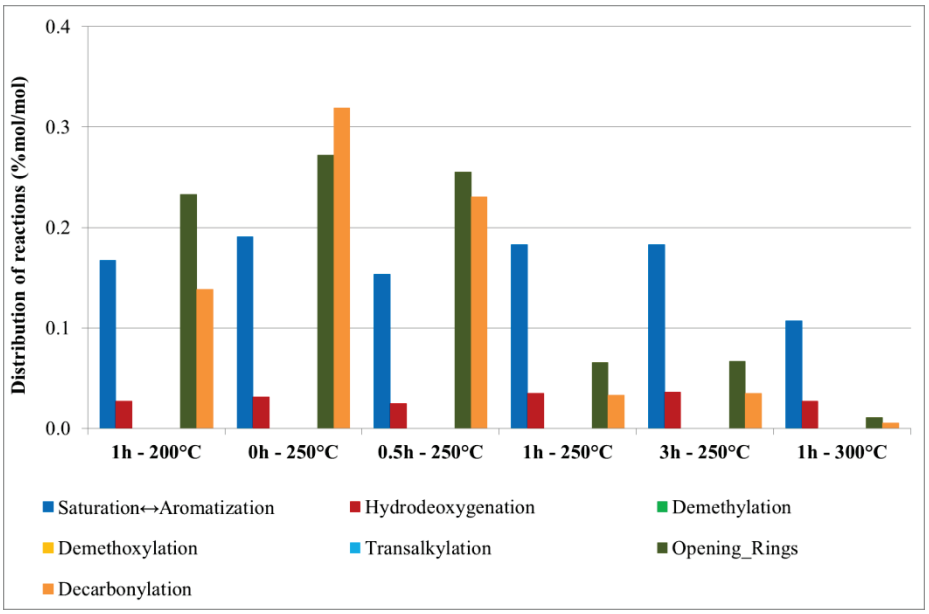


Figure XI-6 - Simulated reactions for all the experimental tests excluding the CO↔OH reaction.

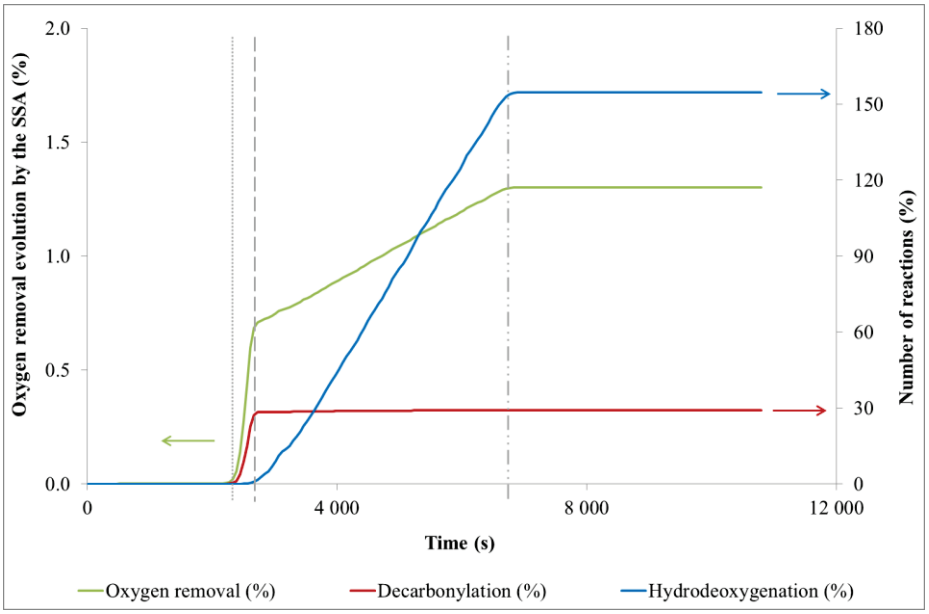


Figure XI-7 - Removal of oxygen and amount of decarbonylation and hydrodeoxygenation reactions at 1 h and 250°C.

Besides saturation↔aromatization and open rings reactions, Figure XI-6 also shows that the reactants also may undergo decarbonylation and hydrodeoxygenation during the simulations. These two last reactions are the responsible for the removal of oxygen from the mixtures as shown in Figure XI-7. As presented, at 250 °C and 1 h of reaction, decarbonylation starts removing oxygen in the form of carbon monoxide. Afterwards, this reaction is replaced by hydrodeoxygenation, which increases even more the

removal of oxygen. Eventually both reactions stop occurring due to the absence of aldehydes and also because of the higher reactivity of the CO↔OH reaction, which makes it the most probable to happen and therefore, the only reaction occurring in the system.

XI.5.3 Global properties of the effluent

Here the PIONA (Paraffins-Isoparaffins-Olefins-Naphthenes-Aromatics) and the oxygenated functional groups families will be analyzed. As illustrated by Figure XI-8, both the feedstock (bio-oil) and the hydrotreating effluents are mostly composed by linear molecules and aromatics[3]. In the experimental tests of the hydrotreating of bio-oil, there is a minor fraction of naphthenes. This increases with the temperature, as seen by the bars at 200 °C, 250 °C and 300 °C at 1h. Regarding the reaction time, for the blank and 30 min (0.5 h) test, the fractions of naphthenes was around about the same. However, after adding 30 min more to the reaction (total 1 h), the content of naphthenes triples. Such behavior is given by the kinetics behind the saturation↔aromatization equilibrium.

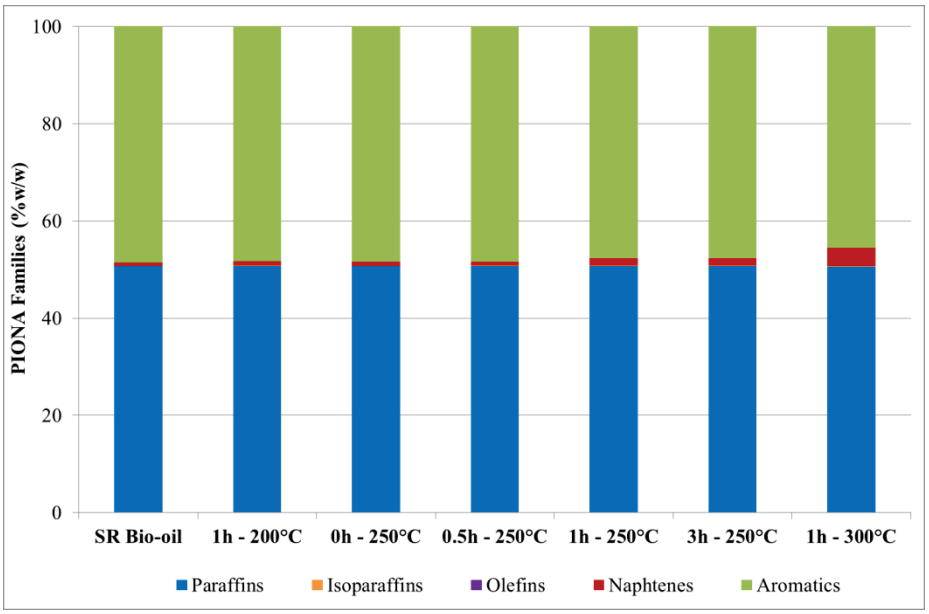


Figure XI-8 - PIONA Families for the reconstructed bio-oil and all the experimental tests.

In the simulation of the hydrotreating, no cracking reactions were considered with exception of the opening of furan and THF rings. However, since this reaction barely occurs in the system (report to Figure XI-3 and Figure XI-6), the number of linear molecules is almost constant. Both the absence of cracking reactions and the low impact of ring opening also justify the deficiency in isoparaffins. Regarding the olefins, these were not reconstructed by the MR, so the mixture does not content any

[3] Please note that the values are in %w/w, thus these do not match to those in Figure X-29, which are in %mol/mol.

unsaturated molecule apart the aromatics. Furthermore, the dehydration reaction, responsible for the production of unsaturated bonds, was not used in the simulations, justifying the absence of olefins.

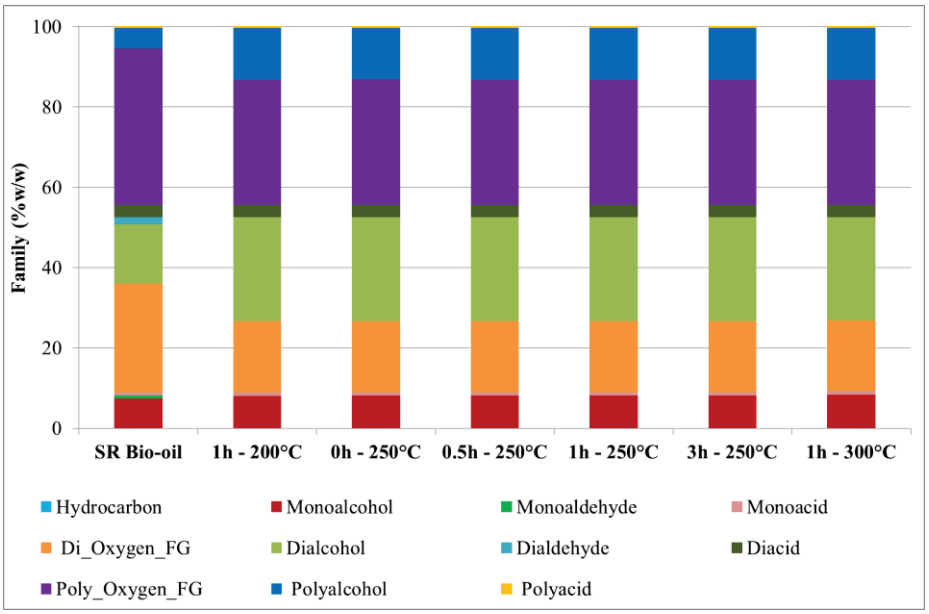


Figure XI-9 - Families of the oxygenated functional groups for the reconstructed bio-oil and all experimental tests.

Concerning the oxygenated functional group families, the feedstock and the hydrotreating effluents share a similar pattern. As illustrated in Figure XI-9, both are mostly composed by molecule with two (Di_Oxygen_FG) or more (Poly_Oxygen_FG) different functional groups, followed by molecules with one of more hydroxyl group and others with traces of carbonyls and carboxyls[4]. Be that as it may, the hydrotreating effluents do not have aldehydes (mono or di) and have a minor content of Di_Oxygen_FG and Poly_Oxygen_FG. These differences may be justified by the conversion of the aldehydes through decarbonylation into a carbon monoxide and by hydrodeoxygenation and CO↔OH into an alcohol.

XI.6 Conclusion of Chapter XI

The present chapter applied all the results obtained in this work into the simulation of the hydrotreating of bio-oil. From the model molecules, the kinetic parameters previously obtained and their reactions were implemented here, alongside an equimolar library of 10 000 molecules.

Concerning the results, the simulator predicted reasonably well the content of carbon and oxygen, but overestimated the content of hydrogen. The same happens for the saturated carbons, indicating a higher presence of hydrogenation reactions than expected. Furthermore, the saturated carbons were already

[4] Please note that the values are in %w/w, thus these do not match to those in Figure X-30, which are in %mol/mol.

overestimated by the SR. Considering that the SSA will promote even more saturated molecules, it is reasonable that the content of the saturated carbons is overestimated. The aromatic carbons were reasonably well predicted. However, these present slight deviations, mostly due to their correlation with the saturated carbons, via saturation↔aromatization reactions.

Regarding the aldehydes, ketones and carboxylic acids, these were underestimated by the simulator, while the hydroxyls and the methoxyls were overestimated. This behavior is connected to the reactions under stake and the potential calibration shortcomings of the ^{13}C NMR. In the present simulations, CO↔OH reaction is the main reactional pathway. The remaining reactions had a minor impact on the reactivity of the system. Due to this reactional pathway (and the smaller reactivity of the hydroxyl oxidation), the simulator converts almost all the aldehydes and ketones into hydroxyls.

Regarding the removal of oxygen, all experimental tests share the same pattern. The removal of oxygen starts at 150 °C in the heating period. However, it stabilizes around the 1 %, after the ending of the heating period. As it happens, decarbonylation and hydrodeoxygenation are the reactions responsible for the removal of the oxygen. These reactions occur to all the aldehydes in the mixture and eventually stop due to their low reactivity when compared to the CO↔OH reaction, which leads the reactional pathways for the conversion of molecules and does not removes oxygen.

Regarding the global properties, the PIONA families are always the same for all experimental tests and the feedstock. In fact, there is none cracking reactions used in the simulation, besides the ring opening, which barely occurs in the system. Therefore, the number of paraffins is basically the same and the number of isoparaffins as well. The aromatics follow a similar behavior. However, certain cycles are saturated during the hydrotreating of bio-oil, generating naphthenes, which do not exist in the feedstock analysis. Regarding the olefins, neither the SR nor the reactions which promote this type of olefins exist. Hence, there are no olefins in the mixture. Concerning the oxygenated functional group families, the feedstock and the hydrotreating effluents have similar trends. However, the upgraded mixtures do not present any form of aldehydes and also have less amount of Di_Oxygen_FG and Poly_Oxygen_FG. These differences may be justified by the conversion of the aldehydes via decarbonylation, hydrodeoxygenation and CO↔OH reactions.

To conclude, although the simulator can reasonable predict the elemental compounds and the carbon specification, it cannot correctly predict the atoms associated to oxygenated functions. Furthermore, the HDO performance is extremely low and does not match the experimental data. In the end, the simulation reveals that the system barely reacts and is trapped in reversibility between the carbonyls and the hydroxyls. The main reason behind this stagnation may be unsuitable kinetics, absence of reactions and also experimental deviations.

General Conclusions

Biomass is an organic feedstock with the potential to produce fuels and valuable chemical building blocks. In order to produce these, biomass must first be pre-treated and converted into a matter that is more easily manageable, such as liquids. One of such processes is fast pyrolysis, which generates a solid, a gas and a liquid, called bio-oil. This mixture is a complex oxygenated matrix, with a vast number of molecules, and consequently, a great diversity of characteristics. In bio-oil, it is possible to find compounds as alcohols, aldehydes, carbohydrates, carboxylic acids, furans, ketones and finally phenols and methoxy-phenols. A higher detail for the molecular characteristics of bio-oil is still quite complicated to acquire experimentally.

The high amount of undesired oxygen in bio-oil does not allow to mix it with conventional fuels composed of hydrocarbons. Hence, bio-oils need to be upgraded in order to serve as fuels. Hydrotreating is the most common process for the upgrading of bio-oils. However, this process is not simple, in particular due to the complex reactivity of the mixture. For this reason, the scientific community has started by analyzing the hydrotreating of model compounds of bio-oil. Briefly, bio-oil and its model molecules, can be converted under hydrotreating conditions into hydrocarbons and water by hydrogenation, hydrogenolysis, hydrodeoxygenation, decarboxylation, decarbonylation, cracking/hydrocracking and polymerization reactions (Wildschut 2009). It is important to stress that these reactions are extremely impacted by solvent, catalyst, temperature and structure of the molecule.

Even if these models are extremely useful to study, identify and retrieve intrinsic information about the bio-oils kinetics, the overall reaction network of bio-oil cannot be totally covered. Appropriate techniques for complex mixtures must be implemented in order to obtain a molecular description of the feed, the reaction network and its kinetics, and to allow for a correct simulation of the upgrading of complex mixture. For the present thesis, it was considered that coupling the stochastic reconstruction and the reconstruction by maximization of entropy would be a good approach for the molecular description of bio-oils, while the stochastic simulation algorithm would be the best method for the simulation of the hydrotreating process.

The aim of the present work was to understand the structure, the composition and the reactivity of bio-oil compounds through modeling of experimental data. In order to achieve this, the present thesis used a stochastic molecular reconstruction method and a stochastic simulation algorithm to, *in silico,* handle the complexity of the system, all the way from a simple model molecule up to a complex matrix, bio-oil.

In order to follow the strategy proposed for the present thesis, the hydrotreating process of two model molecules (guaiacol and furfural) was analyzed. The experimental tests were performed at 13 MPa over

a reduced $NiMo/Al_2O_3$ catalyst at a temperature range of 100 °C to 300 °C and a reaction time between 1 h to 5 h. Three blank tests were also performed for both model molecules (one with guaiacol and two with furfural). In the end, the global mass balance closed for all tests, with an average value of 101 %. The liquid and gas effluents of these tests were analyzed by GC-FID/MS and GC-FID/TCD. In summary, the experimental tests account for more than 80 % of the quantifiable carbon (60 % for the trial at 250 °C). For this reason, it was assumed that the results were suitable for the application of the strategy of this thesis.

For all the experimental tests, the consumption of hydrogen was almost constant. However, it increases with the reaction time, suggesting an increase in the hydrotreating reactions, and it decreases with the temperature, meaning that at high temperature hydrogen start to be produced. Regarding the conversion of guaiacol and furfural, both increase with the reaction time and temperature as expected. However, a baseline in the conversion of guaiacol can be seen between 1 h and 3 h of reaction. This may be explained by the competition between guaiacol and its derivatives for the catalytic sites (Wang et al. 2013). Concerning the products distributions, it is quite difficult to identify the pathways of each compound, as these may come from several different reactions and reactants. However, and in accordance with the retrieved products, it was considered that guaiacol converted mostly by demethylation, demethoxylation, hydrodeoxygenation, saturation and transalkylation, while furfural would preferably undergo hydrogenation of the carbonyl, hydrodeoxygenation, saturation, ring opening and decarbonylation. These reactions, and their associated compounds, were considered for the creation of a kinetic model for the hydrotreating of guaiacol and furfural.

A deterministic model was created for each model compound, in order to retrieve a suitable range for the kinetic parameters to be used later in the SSA. The ranges were based on experimental tests of both model compounds and on the mass balances for a batch reactor. These balances accounted for the dynamic periods of the reactor, the heating and the cooling. The reaction network of each model compound was created based on the reactant and on the compounds with a higher selectivity. The reaction network of guaiacol accounted for demethoxylation, demethylation, transalkylation, hydrodeoxygenation and saturation reactions, while the reaction network of furfural accounted not only for hydrodeoxygenation and saturation reactions, but also for decarbonylation, carbonyl hydrogenation and ring opening. It was assumed that all reactions were pseudo-first order and followed the Arrhenius law. For the reversible reactions, such as aromatization and hydroxyl oxidation, the equilibrium constants were extracted from the open literature and assumed to be equal for all reactions of the same family (all aromatization reaction have the same equilibrium constants and similarly for all oxidation reactions). For guaiacol, the model predicted reasonably well the concentration of certain compounds and experimental conditions, while others are overestimated or underestimated. For furfural, the model predicted reasonably well the concentration of all compounds for the lower temperatures (100 ºC and 150 ºC), but overestimated or underestimated at 250 ºC. The deviations found in both models may be

related to catalyst deactivation, unsuitable parameters, the absence of reactions and neglecting the impact of the molecule structure and the catalytic sites on the reactivity. Furthermore, the lack of experimental data also has an important impact on the estimation of parameters. Although both models present certain shortcomings, they were able to predict reasonably well most experimental tests and the trends of the products within 10 % error.

After the creation of the deterministic model, the estimated parameters were tested in the SSA. This approach is able to describe a reaction network, molecule by molecule, through a discrete evolution in time, while keeping track of the molecular detail, thereby avoiding the pre-definition of the reaction pathways. The SSA was adapted to the present work through the addition of the dynamic periods of the reactor operation and of the reactions undergone by guaiacol and furfural in hydrotreating conditions. In total, for the hydrotreating of guaiacol and furfural, twenty-six reactions were implemented in the algorithm: Methanol decomposition and steam reforming, methanation of monoxide and dioxide of carbon (reversible reaction), water-gas swift (reversible reaction), carbonyl hydrogenation (reversible reaction), decarbonylation, dehydration, demethoxylation, demethylation, hydrodeoxygenation, methylation, transalkylation, unsaturated hydrogenation, hydration and the opening of rings. For the simulations with the deterministic parameters, both model compounds predict similar trends for the majority compounds. The deviations found can originate not only from the equilibrium constant for aromatization and hydroxyl oxidation, but also from the presence of other molecules, which were not accounted for in the deterministic model. For the generation of the kinetic models via the SSA, the parameters from the deterministic model were used as a starting point. For guaiacol, the SSA can only predict the concentration of certain compounds, while the remaining molecules are overestimated or underestimated. Again, the deviations found in the model may be related to catalyst deactivation, unsuitable parameters, the absence of reactions and neglecting the impact of the molecule structure and the catalytic sites in the reactivity. For furfural, the SSA predicts reasonably well the concentrations of most of compounds at low temperatures. At higher temperatures, the model no longer correctly predicts the concentrations, probably due to the presence of reactions which were not accounted for neither in the deterministic nor in the stochastic model (for example hydration) (Ozagac 2016). For both model compounds, activation energy values similar to those of the deterministic models were used, while the pre-exponential constants were re-fitted via the SSA.

The kinetic parameters fitted via the SSA were applied to the simulation of the hydrotreating of bio-oil. However, the stochastic approach requires a library of molecules representative of bio-oil fractions. This library was obtained through a molecular reconstruction technique. This method consists in a first step, Stochastic Reconstruction or SR, which generates an equimolar library of molecules which is in accordance with the experimental data used for the reconstruction. The molar fractions of the molecules in the library are then adjusted to the properties of the mixture via a second step, Reconstruction by Entropy Maximization or REM. Before applying the method, new molecular structures, new ^{13}C NMR

families and a new correlation for the SEC analysis were implemented in the source code, in addition to the creation of several features related to the oxygenated functional groups. The reconstruction was applied to two bio-oil samples (A and B) using two different building diagrams. The generated compositional models were fitted based on the elemental analysis, the ^{13}C NMR analysis and the SEC analysis. For the first reconstruction approach (BD#1), the compositional model estimated reasonably well the elemental analysis, the ^{13}C NMR analysis and the initial fraction of the SEC analysis. The remaining experimental points of the latter analysis presented deviations, mostly associated to the lack of high molecular weight compounds in the libraries. The second reconstruction approach (BD#2) presented similar results to BD#1. Regarding validation of the models, both approaches predicted densities close to the expected values, but the molecules contained too few oxygen atoms per molecule. For this reason, the REM was not applied, as the method would overestimate the molar fractions of certain molecules and underestimate others, which is not desirable. To conclude, although both BDs present certain shortcomings, the compositional models here created can reconstruct libraries that are able to represent the known light fraction of bio-oil. Between both BDs, it was found that BD#2 had fewer deviations and could better estimate and predict the properties than BD#1. For this reason, the results from BD#2 will be applied for the simulation of the hydrotreating of bio-oil.

To consolidate the work, the synthetic mixture was used in the stochastic simulation of the hydrotreating of bio-oils, supported by the kinetics of the model compounds. The simulator could reasonably well predict the elemental compounds and the carbon type specification, except the carbon atoms associated to oxygenated functions. Furthermore, the SSA predicted an extremely low HDO performance, which does not match the experimental data. The main reason behind these deviations could be linked to unsuitable kinetics, absence of some reactions and also experimental deviations.

In summary, the present work has demonstrated that through a molecular reconstruction technique and a kinetic model based on the reactivity of oxygenated model compounds, it is possible to simulate the hydrotreating of bio-oil at the molecular level. However, to achieve a correct prediction of the behavior of bio-oil during the hydrotreating process, a suitable library of molecules and a correct kinetic model must be available. For both, support is required from the open literature and, in particular, a solid data base of experimental information. In the present case, both requirements were extremely constrained. Despite this, it was still possible to lay the foundations for a generic approach, to develop computer tools, and to generate meaningful results for the composition and reactivity of bio-oils and its model molecules.

Perspectives

Although the objectives of the present thesis are concluded, more can be done not only to improve the results presented here, but most importantly to understand and develop a hydrotreating process for compounds obtained from biomass:

Analytical data

As discussed in the State of the Art, the analysis of bio-oil is a complex process from which it is not always possible to extract vital information. Among the various analytical techniques, most of them only provide qualitative data, which is not sufficient for the creation of an appropriate description of the molecular composition of a bio-oil. For the generation of this type of composition models, quantitative data is required. However, this is difficult to obtain.

To achieve the quantification of the molecules, at least the molecular formula of each of the compounds is required. According to the scientific community, this can be obtained through a four-step strategy. First, the bio-oil sample is fractionated into several fractions. Afterwards, these are analyzed separately by GC, LC, RMN or combination of several analytical techniques. Each analytical technique would be associated to several detectors, as refractive index or ultra-violet detectors. Each detector provides different information concerning the fractionated samples. These would then be grouped and organized by chemometric techniques, which would later help to identify the more important features or molecules in bio-oil for a given property. In the end, a property of bio-oil could be obtained by grouping these features or molecules, generating quantitative data. These results could then be implemented in the molecular reconstruction in order to obtain a more detailed reconstruction of the bio-oil.

Molecular reconstruction

In the molecular reconstruction, several ideas may be implemented. Concerning the bio-oils, new building diagrams can be tested, alongside a better approach for representing the distribution of the molecule size. Regarding the algorithm, a system of identification of molecules can be implemented to ease the examination of the results.

As seen before, the building diagram and its PDFs are important in the reconstruction of the molecules. Therefore, new proposals of building diagrams and other PDFs besides the histogram, exponential and gamma distributions can be tested. This sensitivity analysis could then provide a better set of distributions and building diagrams, which would generate a correct representation of the entire bio-oil.

In addition, the definition of the molecule size should also be reshaped. For oil fractions, the molecular reconstruction techniques used the simulated distillation as information regarding the molecule size. However, this analysis is not available for bio-oil, forcing the user and the molecular reconstruction to base the distribution of the molecule size on the SEC analysis. This analytical technique is associated to

the uncertainties regarding the relation Mweq versus Mw. Furthermore, the implementation of the SEC analysis in the molecular reconstruction relies on a correlation, which may also contribute to larger deviations. Both reasons imply that SEC is not a reliable technique to obtain a quantitative distribution of molecule size (Hoekstra et al. 2011). However, it is currently the only analysis that is easily available. To solve this problem, either a new quantitative analysis of molecular size needs to be proposed, or a new correlation should be proposed. The first approach may rely on a detailed molecular analysis based on hyphenation between liquid chromatography and high-resolution mass spectrometry: such analytical set-up should provide exact molecular mass for each detected compound. However, the main issue will be to obtain quantitative data regarding the large chemical diversity of compounds present in bio-oils. On the other hand, the second approach may be more easily feasible, if further studies are focused on relating the results of SEC analysis to the actual molecular weight. Also, other types of correlations can be proposed, as for example a correlation between the viscosity and the molecular weight. However, as the prediction of viscosity is quite complex, the correlation of the SEC analysis currently still is the best approach for bio-oils and other complex oxygenated mixtures.

Regarding the molecule identification, the molecular reconstruction currently uses the MOL format of Oliveira (Oliveira 2013) for storing the molecule structure. However, besides being greedy in memory, this format does not allow the identification of the molecules. Other representations of molecule structures, such as SMILES (Weininger 1988) (Weininger et al. 1989) or InCHI (Heller et al. 2015) could be implemented in the algorithm. These, besides allowing the unambiguous identification of the compounds in the mixtures, consume less memory than the actual format.

Kinetics

As presented in the State of the Art, little information about the kinetics behind the reactivity of bio-oil and its model compounds can be found in the open literature. The existing information mostly focuses on reaction mechanisms and catalyst performance. Even so, when kinetic data is found, it is not always easy to apply, as it is mostly for the global conversion of the reactant and not for the individual reactions.

To face this obstacle, several approaches can be followed. Concerning the present thesis, more experimental tests with guaiacol, furfural and even other model molecules should be added to the experimental data base of the present thesis. Afterwards, these can be employed in the model to find better parameter values. In a more general perspective, several hydrotreating tests could be done to a group of model molecules with similar structure, as was done by Gevert et al. (Gevert et al. 1987) for phenol and methylphenols. These results could then be compared and thus used for the creation of Quantitative Structure Reactivity Correlations (QSRC) or similar techniques. These reactivity relations could then be implemented in the SSA.

Stochastic simulation

In the stochastic simulation, a system of identification of molecules can also be implemented to ease the analysis of the results. This proposal was already mentioned and explained in the section on the molecular reconstruction. Furthermore, the algorithm can be extended to bimolecular reactions. In the case of bio-oils, the extension of the SSA to bimolecular reactions would allow the implementation of undesirable side reactions, which could help understand the hydrotreating process of bio-oil, mainly the macromolecules and coke formation. Another improvement in the SSA could be the implementation of an optimization tool, for when the stochastic approach is used for tuning kinetic models.

As seen before, the SSA can currently only simulate monomolecular reactions. However, the method can be modified to account for bimolecular reactions. One approach could pass by the storage of all molecules responsible for bimolecular reactions. For instance, in bio-oil, molecules with carbonyls are the main responsible for the formation of macromolecules via bimolecular reactions (as described in the State of the Art). So, in the present thesis, the SSA would store, in addition to all molecules, a new feature indicating the presence of a carbonyl. Then, when a molecule would undergo a bimolecular reaction, one of the carbonyl molecules would react with the molecule under investigation, thereby creating a new and larger product.

Regarding the optimization tool, and as seen before, the SSA currently does not have one. In systems like the present thesis, where there are about ten parameters to estimate and where an entire simulation ($\approx$ 2 h) is required to have the results equivalent to a single experimental test, the lack of an optimization tool is a major barrier. For these reasons, an optimization tool should be added to the SSA. This tool could then be activated or deactivated according to the aim of the user. Since the SSA presents a stochastic nature, as the SR, a derivative-free optimizer, such as genetic algorithms, could be envisaged. To conclude, the implementation of the optimization tool would also require the input of the effluent properties.

References

Abdelnur PV, Vaz BG, Rocha JD, de Almeida, Marlon B. B., Teixeira MAG, Pereira RCL (2013) Characterization of Bio-oils from Different Pyrolysis Process Steps and Biomass Using High-Resolution Mass Spectrometry. Energy Fuels 27:6646–6654. doi: 10.1021/ef400788v

Achladas GE (1991) Analysis of biomass pyrolysis liquids: Separation and characterization of phenols. Journal of Chromatography A 542:263–275. doi: 10.1016/S0021-9673(01)88766-5

Allen DT (ed) (1991) Kinetic and Thermodynamic Lumping of Multicomponent Mixtures. Structural models of catalytic cracking chemistry. Elsevier

Allen DT, Liguras D (1991) Structural Models of Catalytic Cracking Chemistry: A Case Study of a Group Contribution Approach to Lumped Kinetic Modeling. Structural models of catalytic cracking chemistry. In: Sapre AV, Krambeck FJ (eds) Chemical Reactions in Complex Mixtures. Springer US, Boston, MA, pp 101–125

Alvarez-Majmutov A, Chen J (2017) Stochastic Modeling and Simulation Approach for Industrial Fixed-Bed Hydrocrackers. Ind. Eng. Chem. Res. 56:6926–6938. doi: 10.1021/acs.iecr.7b01743

Alvarez-Majmutov A, Chen J, Gieleciak R, Hager D, Heshka N, Salmon S (2014) Deriving the Molecular Composition of Middle Distillates by Integrating Statistical Modeling with Advanced Hydrocarbon Characterization. Energy Fuels 28:7385–7393. doi: 10.1021/ef5018169

Alvarez-Majmutov A, Gieleciak R, Chen J (2015) Deriving the Molecular Composition of Vacuum Distillates by Integrating Statistical Modeling and Detailed Hydrocarbon Characterization. Energy Fuels 29:7931–7940. doi: 10.1021/acs.energyfuels.5b02082

Alvarez-Majmutov A, Chen J, Gieleciak R (2016) Molecular-Level Modeling and Simulation of Vacuum Gas Oil Hydrocracking. Energy Fuels 30:138–148. doi: 10.1021/acs.energyfuels.5b02084

Andersson T, Hyötyläinen T, Riekkola M-L (2000) Analysis of phenols in pyrolysis oils by gel permeation chromatography and multidimensional liquid chromatography. Journal of Chromatography A 896:343–349. doi: 10.1016/S0021-9673(00)00678-6

Artok L, Erbatur O, Schobert HH (1996) Reaction of dinaphthyl and diphenyl ethers at liquefaction conditions. Fuel Processing Technology 47:153–176. doi: 10.1016/0378-3820(96)01013-2

Artok L, Su Y, Hirose Y, Hosokawa M, Murata S, Nomura M (1999) Structure and Reactivity of Petroleum-Derived Asphaltene †. Energy Fuels 13:287–296. doi: 10.1021/ef980216a

Bayerbach R, Meier D (2009) Characterization of the water-insoluble fraction from fast pyrolysis liquids (pyrolytic lignin). Part IV: Structure elucidation of oligomeric molecules. Journal of Analytical and Applied Pyrolysis 85:98–107. doi: 10.1016/j.jaap.2008.10.021

Bayerbach R, van Nguyen D, Schurr U, Meier D (2006) Characterization of the water-insoluble fraction from fast pyrolysis liquids (pyrolytic lignin). Journal of Analytical and Applied Pyrolysis 77:95–101. doi: 10.1016/j.jaap.2006.02.002

Bennett CA (2009) User-Controlled Network Generation with INGEN, Rutgers, The State University of New Jersey

Benson SW (1968) Thermochemical kinetics: Methods for the estimation of thermochemical data and rate parameters. Wiley, New York

Bertero M, La Puente G de, Sedran U (2012) Fuels from bio-oils: Bio-oil production from different residual sources, characterization and thermal conditioning. Fuel 95:263–271. doi: 10.1016/j.fuel.2011.08.041

Bie Y, Gutierrez A, Viljava TR, Kanervo JM, Lehtonen J (2013) Hydrodeoxygenation of Methyl Heptanoate over Noble Metal Catalysts: Catalyst Screening and Reaction Network. Ind. Eng. Chem. Res. 52:11544–11551. doi: 10.1021/ie4012485

Bie Y, Kanervo JM, Lehtonen J (2015) Hydrodeoxygenation of Methyl Heptanoate over Rh/ZrO 2 Catalyst as a Model Reaction for Biofuel Production: Kinetic Modeling Based On Reaction Mechanism. Ind. Eng. Chem. Res. 54:11986–11996. doi: 10.1021/acs.iecr.5b03232

Bindwal AB, Vaidya PD (2013) Kinetics of Aqueous-Phase Hydrogenation of Levoglucosan over Ru/C Catalyst. Ind. Eng. Chem. Res. 52:17781–17789. doi: 10.1021/ie401913j

Boduszynski MM (1987) Composition of heavy petroleums. 1. Molecular weight, hydrogen deficiency, and heteroatom concentration as a function of atmospheric equivalent boiling point up to 1400.degree.F (760.degree.C). Energy Fuels 1:2–11. doi: 10.1021/ef00001a001

Boduszynski MM (1988) Composition of heavy petroleums. 2. Molecular characterization. Energy Fuels 2:597–613. doi: 10.1021/ef00011a001

Boek ES, Yakovlev DS, Headen TF (2009) Quantitative Molecular Representation of Asphaltenes and Molecular Dynamics Simulation of Their Aggregation †. Energy Fuels 23:1209–1219. doi: 10.1021/ef800876b

Bondi A (1964) van der Waals Volumes and Radii. J. Phys. Chem. 68:441–451. doi: 10.1021/j100785a001

Bour H (2017) Confidential. Master Thesis

Branca C, Giudicianni P, Di Blasi C (2003) GC/MS Characterization of Liquids Generated from Low-Temperature Pyrolysis of Wood. Ind. Eng. Chem. Res. 42:3190–3202. doi: 10.1021/ie030066d

Bridgwater A (2003) Renewable fuels and chemicals by thermal processing of biomass. Chemical Engineering Journal 91:87–102. doi: 10.1016/S1385-8947(02)00142-0

Bridgwater T (2007) IEA Bioenergy 27th update. Biomass and Bioenergy 31:7–18. doi: 10.1016/S0961-9534(07)00047-5

Bunch AY, Wang X, Ozkan US (2007) Hydrodeoxygenation of benzofuran over sulfided and reduced Ni–Mo/γ-Al2O3 catalysts: Effect of H2S. Journal of Molecular Catalysis A: Chemical 270:264–272. doi: 10.1016/j.molcata.2007.02.006

Busetto L, Fabbri D, Mazzoni R, Salmi M, Torri C, Zanotti V (2011) Application of the Shvo catalyst in homogeneous hydrogenation of bio-oil obtained from pyrolysis of white poplar: New mild upgrading conditions. Fuel 90:1197–1207. doi: 10.1016/j.fuel.2010.10.036

Campbell DM, Bennett C, Hou Z, Klein MT (2009) Attribute-Based Modeling of Resid Structure and Reaction. Ind. Eng. Chem. Res. 48:1683–1693. doi: 10.1021/ie8012314

Centeno A, Laurent E, Delmon B (1995) Influence of the Support of CoMo Sulfide Catalysts and of the Addition of Potassium and Platinum on the Catalytic Performances for the Hydrodeoxygenation of Carbonyl, Carboxyl, and Guaiacol-Type Molecules. Journal of Catalysis 154:288–298. doi: 10.1006/jcat.1995.1170

Charon N, Ponthus J, Espinat D, Broust F, Volle G, Valette J, Meier D (2015) Multi-technique characterization of fast pyrolysis oils. Journal of Analytical and Applied Pyrolysis 116:18–26. doi: 10.1016/j.jaap.2015.10.012

Charon-Revellin N, Dulot H, López-García C, Jose J (2011) Kinetic Modeling of Vacuum Gas Oil Hydrotreatment using a Molecular Reconstruction Approach. Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles 66:479–490. doi: 10.2516/ogst/2010005

Cheng T, Han Y, Zhang Y, Xu C (2016) Molecular composition of oxygenated compounds in fast pyrolysis bio-oil and its supercritical fluid extracts. Fuel 172:49–57. doi: 10.1016/j.fuel.2015.12.075

Christensen ED, Chupka GM, Luecke J, Smurthwaite T, Alleman TL, Iisa K, Franz JA, Elliott DC, McCormick RL (2011) Analysis of Oxygenated Compounds in Hydrotreated Biomass Fast Pyrolysis Oil Distillate Fractions. Energy Fuels 25:5462–5471. doi: 10.1021/ef201357h

Crezee E (2003) Three-phase hydrogenation of D-glucose over a carbon supported ruthenium catalyst-mass transfer and kinetics. Applied Catalysis A: General 251:1–17. doi: 10.1016/S0926-860X(03)00587-8

Dahiya A (2015) Bioenergy: Biomass to biofuels. Pyrolysis of Lignocellulosic Biomass: Oil, Char and Gas

Djokic MR, Dijkmans T, Yildiz G, Prins W, Van Geem, Kevin M. (2012) Quantitative analysis of crude and stabilized bio-oils by comprehensive two-dimensional gas-chromatography. J Chromatogr A 1257:131–140. doi: 10.1016/j.chroma.2012.07.035

Elliott DC, Neuenschwander GG, Hart TR, Hu J, Solana AE, Cao C (eds) (2006) Hydrogenation of bio-oil for chemical and fuel production, Science in thermal and chemical biomass conversion. CPL Press, Speen

Elliott DC (2007) Historical Developments in Hydroprocessing Bio-oils. Energy Fuels 21:1792–1815. doi: 10.1021/ef070044u

Elliott DC, Hart TR (2009) Catalytic Hydroprocessing of Chemical Models for Bio-oil. Energy Fuels 23:631–637. doi: 10.1021/ef8007773

Evans RJ, Milne TA (1987) Molecular characterization of the pyrolysis of biomass. Energy Fuels 1:123–137. doi: 10.1021/ef00002a001

Faix O, Meier D, Fortmann I (1990) Thermal degradation products of wood: Gas chromatographic separation and mass spectrometric characterization of monomeric lignin derived products. Holz als Roh-und Werkstoff 48:281–285

Faix O, Fortmann I, Bremer J, Meier D (1991) Thermal degradation products of wood. Holz als Roh-und Werkstoff 49:213–219. doi: 10.1007/BF02613278

Faulon JL (1991) Prediction, elucidation et modelisation moleculaire : algorithmes et applications en geochimie, École Nationale Superieure des Mines

Faulon JL (1994) Stochastic Generator of Chemical Structure. 1. Application to the Structure Elucidation of Large Molecules. J. Chem. Inf. Model. 34:1204–1218. doi: 10.1021/ci00021a031

Faulon JL, Drappier JM, Romero M, Vandenbroucke M, Behar F (1990) Modélisation des structures chimiques des macromolécules sédimentaires: Le logiciel XMOL Software. Rev. Inst. Fr. Pét. 45:161–180. doi: 10.2516/ogst:1990014

Fisk CA, Morgan T, Ji Y, Crocker M, Crofcheck C, Lewis SA (2009) Bio-oil upgrading over platinum catalysts using in situ generated hydrogen. Applied Catalysis A: General 358:150–156. doi: 10.1016/j.apcata.2009.02.006

Freudenberg K, Neish AC (1968) Constitution and Biosynthesis of Lignin. Springer-Verlag, New York

Fullana A, Contreras JA, Striebich RC, Sidhu SS (2005) Multidimensional GC/MS analysis of pyrolytic oils. Journal of Analytical and Applied Pyrolysis 74:315–326. doi: 10.1016/j.jaap.2004.11.036

Furimsky E (2000) Catalytic hydrodeoxygenation. Applied Catalysis A: General 199:147–190. doi: 10.1016/S0926-860X(99)00555-4

Gevert SB, Eriksson M, Eriksson P, Massoth FE (1994) Direct hydrodeoxygenation and hydrogenation of 2,6- and 3,5-dimethylphenol over sulphided CoMo catalyst. Applied Catalysis A: General 117:151–162. doi: 10.1016/0926-860X(94)85095-X

Gevert BS, Otterstedt J-E, Massoth FE (1987) Kinetics of the HDO of methyl-substituted phenols. Applied Catalysis 31:119–131. doi: 10.1016/S0166-9834(00)80671-5

Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. Journal of Computational Physics 22:403–434. doi: 10.1016/0021-9991(76)90041-3

Gillespie DT (1992) A rigorous derivation of the chemical master equation. Physica A: Statistical Mechanics and its Applications 188:404–425. doi: 10.1016/0378-4371(92)90283-V

Glasser WG, Glasser HR, Morohoshi N (1981) Simulation of reactions with lignin by computer (SIMREL). 6. Interpretation of primary experimental analysis data ("analysis program"). Macromolecules 14:253–262. doi: 10.1021/ma50003a006

Gnansounou E, Dauriat A Le bioéthanol

Gomez-Prado J, Zhang N, Theodoropoulos C (2008) Characterisation of heavy petroleum fractions using modified molecular-type homologous series (MTHS) representation. Energy 33:974–987

González C, Marín P, Díez FV, Ordóñez S (2015) Hydrodeoxygenation of Acetophenone over Supported Precious Metal Catalysts at Mild Conditions: Process Optimization and Reaction Kinetics. Energy Fuels 29:8208–8215. doi: 10.1021/acs.energyfuels.5b02112

González C, Marín P, Díez FV, Ordóñez S (2016) Gas-Phase Hydrodeoxygenation of Benzaldehyde, Benzyl Alcohol, Phenyl Acetate, and Anisole over Precious Metal Catalysts. Ind. Eng. Chem. Res. 55:2319–2327. doi: 10.1021/acs.iecr.6b00036

Grilc M, Likozar B, Levec J (2014) Hydrotreatment of solvolytically liquefied lignocellulosic biomass over NiMo/Al2O3 catalyst: Reaction mechanism, hydrodeoxygenation kinetics and mass transfer model based on FTIR. Biomass and Bioenergy 63:300–312. doi: 10.1016/j.biombioe.2014.02.014

Gross JH (2010) Mass spectrometry: A textbook, 2nd ed. Springer, Berlin, London

Guo X, Wang S, Guo Z, Liu Q, Luo Z, Cen K (2010a) Pyrolysis characteristics of bio-oil fractions separated by molecular distillation. Applied Energy 87:2892–2898. doi: 10.1016/j.apenergy.2009.10.004

Guo Z, Wang S, Gu Y, Xu G, Li X, Luo Z (2010b) Separation characteristics of biomass pyrolysis oil in molecular distillation. Separation and Purification Technology 76:52–57. doi: 10.1016/j.seppur.2010.09.019

Hao N, Ben H, Yoo CG, Adhikari S, Ragauskas AJ (2016) Review of NMR Characterization of Pyrolysis Oils. Energy Fuels 30:6863–6880. doi: 10.1021/acs.energyfuels.6b01002

Harvey D (2000) Modern analytical chemistry. McGraw-Hill, Boston

He Z, Wang X (2012) Hydrodeoxygenation of model compounds and catalytic systems for pyrolysis bio-oils upgrading. Catalysis for Sustainable Energy 1. doi: 10.2478/cse-2012-0004

Heller SR, McNaught A, Pletnev I, Stein S, Tchekhovskoi D (2015) InChI, the IUPAC International Chemical Identifier. J Cheminform 7:23. doi: 10.1186/s13321-015-0068-4

Hilten RN, Das KC (2010) Comparison of three accelerated aging procedures to assess bio-oil stability. Fuel 89:2741–2749. doi: 10.1016/j.fuel.2010.03.033

Hirsch E, Altgelt KH (1970) Integrated structural analysis. Method for the determination of average structural parameters of petroleum heavy ends. Analytical Chemistry 42:1330–1339

Hoekstra E, Kersten SR, Tudos A, Meier D, Hogendoorn KJ (2011) Possibilities and pitfalls in analyzing (upgraded) pyrolysis oil by size exclusion chromatography (SEC). Journal of Analytical and Applied Pyrolysis 91:76–88. doi: 10.1016/j.jaap.2011.01.006

Hoffmann Ed, Stroobant V (2007) Mass spectrometry: Principles and applications, 3rd ed. J. Wiley, Chichester, West Sussex, England, Hoboken, NJ

Holland JH (1975) Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. University of Michigan Press, Ann Arbor

Horton SR, Hou Z, Moreno BM, Bennett CA, Klein MT (2013) Molecule-based modeling of heavy oil. Sci. China Chem. 56:840–847. doi: 10.1007/s11426-013-4895-8

Horton SR, Mohr RJ, Zhang Y, Petrocelli FP, Klein MT (2015a) Molecular-Level Kinetic Modeling of Biomass Gasification. Energy Fuels 30:1647–1661. doi: 10.1021/acs.energyfuels.5b01988

Horton SR, Zhang L, Hou Z, Bennett CA, Klein MT, Zhao S (2015b) Molecular-Level Kinetic Modeling of Resid Pyrolysis. Ind. Eng. Chem. Res. 54:4226–4235. doi: 10.1021/ie5041572

Horton SR, Woeckener J, Mohr R, Zhang Y, Petrocelli F, Klein MT (2015c) Molecular-Level Kinetic Modeling of the Gasification of Common Plastics. Energy Fuels 30:1662–1674. doi: 10.1021/acs.energyfuels.5b02047

Hou Z, Bennett C, Klein MT, Virk PS (2010) Approaches and Software Tools for Modeling Lignin Pyrolysis †. Energy Fuels 24:58–67. doi: 10.1021/ef900488k

Hu S, Towler G, Zhu XX (2002) Combine Molecular Modeling with Optimization to Stretch Refinery Operation. Ind. Eng. Chem. Res. 41:825–841. doi: 10.1021/ie0010215

Hu X, Wang Y, Mourant D, Gunawan R, Lievens C, Chaiwat W, Gholizadeh M, Wu L, Li X, Li C-Z (2013) Polymerization on heating up of bio-oil: A model compound study. AIChE J 59:888–900. doi: 10.1002/aic.13857

Hudebine D (2003) Reconstruction moléculaire de coupes pétrolières, École Normale Supérieure de Lyon

Hudebine D, Verstraete J (2004) Molecular reconstruction of LCO gasoils from overall petroleum analyses. Chemical Engineering Science 59:4755–4763. doi: 10.1016/j.ces.2004.09.019

Hudebine D, Verstraete JJ (2011) Reconstruction of Petroleum Feedstocks by Entropy Maximization. Application to FCC Gasolines. Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles 66:437–460. doi: 10.2516/ogst/2011110

Hudebine D, Verstraete J, Chapus T (2011) Statistical Reconstruction of Gas Oil Cuts. Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles 66:461–477. doi: 10.2516/ogst/2009047

IEA. https://www.iea.org/topics/renewables/subtopics/bioenergy/

Iwase M, Sugiyama S, Liang Y, Masuda Y, Morimoto M, Matsuoka T, Boek ES, Ueda R, Nakagawa K (2018) Development of Digital Oil for Heavy Crude Oil: Molecular Model and Molecular Dynamics Simulations. Energy Fuels 32:2781–2792. doi: 10.1021/acs.energyfuels.7b02881

J. P. Diebold (2000) A Review of the Chemical and Physical Mechanisms of the Sorage Stability of Fast Pyrolysis Bio-oils

Jaffe SB, Freund H, Olmstead WN (2005) Extension of Structure-Oriented Lumping to Vacuum Residua. Ind. Eng. Chem. Res. 44:9840–9852. doi: 10.1021/ie058048e

Jarvis JM, McKenna AM, Hilten RN, Das KC, Rodgers RP, Marshall AG (2012) Characterization of Pine Pellet and Peanut Hull Pyrolysis Bio-oils by Negative-Ion Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Energy Fuels 26:3810–3815. doi: 10.1021/ef300385f

Joback KG, Reid RC (2007) Estimation of pure-component properties from group-contributions. Chemical Engineering Communications 57:233–243

Jong W, Ommen JR (2015) Biomass as a sustainable energy source for the future: Fundamentals of conversion processes. Biomass Composition, Properties and Characterization

Joseph J, Baker C, Mukkamala S, Beis SH, Wheeler MC, DeSisto WJ, Jensen BL, Frederick BG (2010) Chemical Shifts and Lifetimes for Nuclear Magnetic Resonance (NMR) Analysis of Biofuels. Energy Fuels 24:5153–5162. doi: 10.1021/ef100504d

Joshi N, Lawal A (2012) Hydrodeoxygenation of acetic acid in a microreactor. Chemical Engineering Science 84:761–771. doi: 10.1016/j.ces.2012.09.018

Kantarelis E, Yang W, Blasiak W (2013) Production of Liquid Feedstock from Biomass via Steam Pyrolysis in a Fluidized Bed Reactor. Energy Fuels 27:4748–4759. doi: 10.1021/ef400580x

Kirby SR, Song C, Schobert HH (1996) Hydrodeoxygenation of O-containing polycyclic model compounds using a novel organometallic catalyst precursor. Catalysis Today 31:121–135. doi: 10.1016/0920-5861(96)00083-1

Kirilin A, Wärnå J, Tokarev A, Murzin DY (2014) Kinetic Modeling of Sorbitol Aqueous-Phase Reforming over Pt/Al 2 O 3. Ind. Eng. Chem. Res. 53:4580–4588. doi: 10.1021/ie403813y

Klein MT, Hou G, Ralph BJ, Broadbelt LJ, Kumar A (2006) Molecular modeling in heavy hydrocarbon conversions. Chemical industries, vol 109. CRC/Taylor & Francis, Boca Raton

Korre SC (1994) Quantitative structure/reactivity correlations as a reation engineering tol: Applications to hydrocracking of polynuclear aromatics, University of Delaware

Kurtz TG (1972) The Relationship between Stochastic and Deterministic Models for Chemical Reactions. J. Chem. Phys. 57:2976. doi: 10.1063/1.1678692

Laurent E, Delmon B (1994) Study of the hydrodeoxygenation of carbonyl, car□ylic and guaiacyl groups over sulfided CoMo/γ-Al2O3 and NiMo/γ-Al2O3 catalysts. Applied Catalysis A: General 109:77–96. doi: 10.1016/0926-860X(94)85004-6

Li N, Huber GW (2010) Aqueous-phase hydrodeoxygenation of sorbitol with Pt/SiO2–Al2O3: Identification of reaction intermediates. Journal of Catalysis 270:48–59. doi: 10.1016/j.jcat.2009.12.006

Libaniti C (1992) Monte Carlo Simulation of complex reactive macromolecular systems, University of Delaware

Lievens C, Mourant D, He M, Gunawan R, Li X, Li C-Z (2011) An FT-IR spectroscopic study of carbonyl functionalities in bio-oils. Fuel 90:3417–3423. doi: 10.1016/j.fuel.2011.06.001

Liu Y, Shi Q, Zhang Y, He Y, Chung KH, Zhao S, Xu C (2012) Characterization of Red Pine Pyrolysis Bio-oil by Gas Chromatography–Mass Spectrometry and Negative-Ion Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Energy Fuels 26:4532–4539. doi: 10.1021/ef300501t

Lopez Abelairas M, Oliveira LP de, Verstraete JJ (2016) Application of Monte Carlo techniques to LCO gas oil hydrotreating: Molecular reconstruction and kinetic modelling. Catalysis Today 271:188–198. doi: 10.1016/j.cattod.2016.02.041

López García C, Hudebine D, Schweitzer J-M, Verstraete JJ, Ferré D (2010) In-depth modeling of gas oil hydrotreating: From feedstock reconstruction to reactor stability analysis. Catalysis Today 150:279–299. doi: 10.1016/j.cattod.2009.08.002

Lu Q, Yang X-l, Zhu X-f (2008) Analysis on chemical and physical properties of bio-oil pyrolyzed from rice husk. Journal of Analytical and Applied Pyrolysis 82:191–198. doi: 10.1016/j.jaap.2008.03.003

Lu R, Sheng G-P, Hu Y-Y, Zheng P, Jiang H, Tang Y, Yu H-Q (2011) Fractional characterization of a bio-oil derived from rice husk. Biomass and Bioenergy 35:671–678. doi: 10.1016/j.biombioe.2010.10.017

Mahinpey N, Murugan P, Mani T, Raina R (2009) Analysis of Bio-Oil, Biogas, and Biochar from Pressurized Pyrolysis of Wheat Straw Using a Tubular Reactor. Energy Fuels 23:2736–2742. doi: 10.1021/ef8010959

Marsman JH, Wildschut J, Mahfud F, Heeres HJ (2007) Identification of components in fast pyrolysis oil and upgraded products by comprehensive two-dimensional gas chromatography and flame ionisation detection. J Chromatogr A 1150:21–27. doi: 10.1016/j.chroma.2006.11.047

Marsman JH, Wildschut J, Evers P, Koning S de, Heeres HJ (2008) Identification and classification of components in flash pyrolysis oil and hydrodeoxygenated oils by two-dimensional gas chromatography and time-of-flight mass spectrometry. J Chromatogr A 1188:17–25. doi: 10.1016/j.chroma.2008.02.034

Massoth FE, Politzer P, Concha MC, Murray JS, Jakowski J, Simons J (2006) Catalytic hydrodeoxygenation of methyl-substituted phenols: correlations of kinetic parameters with molecular properties. J Phys Chem B 110:14283–14291. doi: 10.1021/jp057332g

McDermott JB, Libanati C, LaMarca C, Klein MT (1990) Quantitative use of model compound information: Monte Carlo simulation of the reactions of complex macromolecules. Ind. Eng. Chem. Res. 29:22–29. doi: 10.1021/ie00097a004

Medeiros, D. C. M. (2013) Chemical Representation of Complex Mixtures. Dissertation, FEUP

Melero JA, García A, Iglesias J (2011) Biomass catalysis in conventional refineries. In: Advances in Clean Hydrocarbon Fuel Processing. Elsevier, pp 199–240

Metaxas K, Thybaut JW, Morra G, Farrusseng D, Mirodatos C, Marin GB (2010) A Microkinetic Vision on High-Throughput Catalyst Formulation and Optimization: Development of an Appropriate Software Tool. Top Catal 53:64–76. doi: 10.1007/s11244-009-9432-9

Miettinen I, Kuittinen S, Paasikallio V, Mäkinen M, Pappinen A, Jänis J (2017) Characterization of fast pyrolysis oil from short-rotation willow by high-resolution Fourier transform ion cyclotron resonance mass spectrometry. Fuel 207:189–197. doi: 10.1016/j.fuel.2017.06.053

Migliorini MV, Moraes MSA, Machado ME, Caramão EB (2013) Caracterizaçãodefenóisnobio-óleodapirólisedecaroço de pêssego por GC/MS e GC×GC/TOFMS. SC 5:47–65. doi: 10.4322/sc.2013.006

Milne T, Agblevor F, Davis M, Deutch S, Johnson D (1997) Developments in Thermochemical Biomass Conversion. A Review of the Chemical Composition of Fast-Pyrolysis Oils from Biomass. Springer Netherlands, Dordrecht

Moreno BM (2014) Thermochemical conversion of biomass: Models and modeling approaches, University of Delaware

Moreno BM, Li N, Lee J, Huber GW, Klein MT (2013) Modeling aqueous-phase hydrodeoxygenation of sorbitol over Pt/SiO2–Al2O3. RSC Adv. 3:23769. doi: 10.1039/c3ra45179h

Mullen CA, Strahan GD, Boateng AA (2009) Characterization of Various Fast-Pyrolysis Bio-Oils by NMR Spectroscopy †. Energy Fuels 23:2707–2718. doi: 10.1021/ef801048b

Murwanashyaka JN, Pakdel H, Roy C (2001) Step-wise and one-step vacuum pyrolysis of birch-derived biomass to monitor the evolution of phenols. Journal of Analytical and Applied Pyrolysis 60:219–231. doi: 10.1016/S0165-2370(00)00206-0

Negahdar L, Gonzalez-Quiroga A, Otyuskaya D, Toraman HE, Liu L, Jastrzebski, Johann T B H, Van Geem, Kevin M., Marin GB, Thybaut JW, Weckhuysen BM (2016) Characterization and Comparison of Fast Pyrolysis Bio-oils from Pinewood, Rapeseed Cake, and Wheat Straw Using 13C NMR and Comprehensive GC × GC. ACS Sustain Chem Eng 4:4974–4985. doi: 10.1021/acssuschemeng.6b01329

Neurock M (1992) A computational chemical reaction engineering analysis of complex heavy hydrocarbon reaction systems, University of Delaware

Neurock M, Nigam A, Trauth D, Klein MT (1994) Molecular representation of complex hydrocarbon feedstocks through efficient characterization and stochastic algorithms. Chemical Engineering Science 49:4153–4177. doi: 10.1016/S0009-2509(05)80013-2

Nimmanwudipong T, Runnebaum RC, Block DE, Gates BC (2011) Catalytic Conversion of Guaiacol Catalyzed by Platinum Supported on Alumina: Reaction Network Including Hydrodeoxygenation Reactions. Energy Fuels 25:3417–3427. doi: 10.1021/ef200803d

Oasmaa A, Meier D (2002) Fast Pyrolysis of Biomass: A Handbook. Analysis, Characterization and Test Methods of Fast Pyrolysis Liquids

Oasmaa A, Kuoppala E, Solantausta Y (2003) Fast Pyrolysis of Forestry Residue. 2. Physicochemical Composition of Product Liquid. Energy Fuels 17:433–443. doi: 10.1021/ef020206g

Olcese R, Carré V, Aubriet F, Dufour A (2013) Selectivity of Bio-oils Catalytic Hydrotreatment Assessed by Petroleomic and GC*GC/MS-FID Analysis. Energy Fuels 27:2135–2145. doi: 10.1021/ef302145g

Oliveira LP (2013) Développement d'une méthodologie de modélisation cinétique de procédés de raffinage traitant des charges lourdes, École Normale Supérieure de Lyon

Oliveira LP, Verstraete JJ, Kolb M (2012a) A Monte Carlo modeling methodology for the simulation of hydrotreating processes. Chemical Engineering Journal 207-208:94–102. doi: 10.1016/j.cej.2012.05.039

Oliveira LP, Verstraete JJ, KoIb M (2012b) Molecular representation of petroleum fractions and molecule-based kinetic modeling by Monte Carlo methods. Prepr. Pap.-Am. Chem. Soc., Div. Energy Fuels Chem 57(2), 982

Oliveira LP, Verstraete JJ, KoIb M (2013a) Development of a General Modelling Methodology for Vacuum Residue Hydroconversion. Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles 68:1027–1038. doi: 10.2516/ogst/2013135

Oliveira LP, Verstraete JJ, Kolb M (2013b) Molecule-based kinetic modeling by Monte Carlo methods for heavy petroleum conversion. Sci. China Chem. 56:1608–1622. doi: 10.1007/s11426-013-4989-3

Oliveira LP, Verstraete JJ, Kolb M (2014) Simulating vacuum residue hydroconversion by means of Monte-Carlo techniques. Catalysis Today 220-222:208–220. doi: 10.1016/j.cattod.2013.08.011

Oliveira LP de, Hudebine D, Guillaume D, Verstraete JJ, Joly JF (2016) A Review of Kinetic Modeling Methodologies for Complex Processes. Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles 71:45. doi: 10.2516/ogst/2016011

Oliveira LP de, Vazquez AT, Verstraete JJ, Kolb M (2013c) Molecular Reconstruction of Petroleum Fractions: Application to Vacuum Residues from Different Origins. Energy Fuels 27:3622–3641. doi: 10.1021/ef300768u

Omais B, Crepier J, Charon N, Courtiade M, Quignard A, Thiebaut D (2013) Oxygen speciation in upgraded fast pyrolysis bio-oils by comprehensive two-dimensional gas chromatography. Analyst 138:2258–2268. doi: 10.1039/c2an35597c

Onyestyák G, Harnos S, Kaszonyi A, Štolcová M, Kalló D (2012) Acetic acid hydroconversion to ethanol over novel InNi/Al2O3 catalysts. Catalysis Communications 27:159–163. doi: 10.1016/j.catcom.2012.07.021

Oppenheim I (1969) Stochastic and Deterministic Formulation of Chemical Rate Equations. J. Chem. Phys. 50:460. doi: 10.1063/1.1670820

Orata F (2012) Derivatization Reactions and Reagents for Gas Chromatography Analysis. In: Ali Mohd M (ed) Advanced Gas Chromatography - Progress in Agricultural, Biomedical and Industrial Applications. InTech

Otyuskaya D, Thybaut JW, Lødeng R, Marin GB (2017) Anisole Hydrotreatment Kinetics on CoMo Catalyst in the Absence of Sulfur: Experimental Investigation and Model Construction. Energy Fuels 31:7082–7092. doi: 10.1021/acs.energyfuels.7b00519

Ourique J, Silva Telles A (1997) Estimation of Properties of Pure Organic Substances with Group and Pair Contributions. Braz. J. Chem. Eng. 14. doi: 10.1590/S0104-66321997000200005

Ozagac M (2016) Etude mécanistique de l'hydroconversion catalytique de bio-huiles de pyrolyse. PhD, Université Claude Bernard - Lyon 1

Ozagac M, Bertino-Ghera C, Uzio D, Rivallan M, Laurenti D, Geantet C (2016a) Impact of guaiacol on the formation of undesired macromolecules during catalytic hydroconversion of bio-oil: A model compounds study. Biomass and Bioenergy 95:194–205. doi: 10.1016/j.biombioe.2016.09.022

Ozagac M, Bertino-Ghera C, Uzio D, Rivallan M, Laurenti D, Geantet C (2016b) Understanding macromolecules formation from the catalytic hydroconversion of pyrolysis bio-oil model compounds. Biomass and Bioenergy 95:182–193. doi: 10.1016/j.biombioe.2016.10.007

Özbay N, Apaydın-Varol E, Burcu Uzun B, Eren Pütün A (2008) Characterization of bio-oil obtained from fruit pulp pyrolysis. Energy 33:1233–1240. doi: 10.1016/j.energy.2008.04.006

Özçimen D, Karaosmanoğlu F (2004) Production and characterization of bio-oil and biochar from rapeseed cake. Renewable Energy 29:779–787. doi: 10.1016/j.renene.2003.09.006

Parker JR (2002) Genetic Algorithms for Continuous Problems. In: Goos G, Hartmanis J, van Leeuwen J, Cohen R, Spencer B (eds) Advances in Artificial Intelligence, vol 2338. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 176–184

Patwardhan PR (2010) Understanting the product distribution from biomass fast pyrolysis, Iowa State University

Peng B (1999) Molecular modelling of petroleum processes, University of Manchester

Peng B, Zhao C, Mejía-Centeno I, Fuentes GA, Jentys A, Lercher JA (2012) Comparison of kinetics and reaction pathways for hydrodeoxygenation of C3 alcohols on Pt/Al2O3. Catalysis Today 183:3–9. doi: 10.1016/j.cattod.2011.10.022

Poole CF (2003) New trends in solid-phase extraction. TrAC Trends in Analytical Chemistry 22:362–373. doi: 10.1016/S0165-9936(03)00605-8

Procházková D, Zámostný P, Bejblová M, Červený L, Čejka J (2007) Hydrodeoxygenation of aldehydes catalyzed by supported palladium catalysts. Applied Catalysis A: General 332:56–64. doi: 10.1016/j.apcata.2007.08.009

Prosen H, Zupančič-Kralj L (1999) Solid-phase microextraction. TrAC Trends in Analytical Chemistry 18:272–282. doi: 10.1016/S0165-9936(98)00109-5

Pütün A, Özcan A, Pütün E (1999) Pyrolysis of hazelnut shells in a fixed-bed tubular reactor: Yields and structural analysis of bio-oil. Journal of Analytical and Applied Pyrolysis 52:33–49. doi: 10.1016/S0165-2370(99)00044-3

Pyl SP, Hou Z, Van Geem, Kevin M., Reyniers M-F, Marin GB, Klein MT (2011) Modeling the Composition of Crude Oil Fractions Using Constrained Homologous Series. Ind. Eng. Chem. Res. 50:10850–10858. doi: 10.1021/ie200583t

Qiang L, Wen-zhi L, Dong Z, Xi-feng Z (2009) Analytical pyrolysis–gas chromatography/mass spectrometry (Py–GC/MS) of sawdust with Al/SBA-15 catalysts. Journal of Analytical and Applied Pyrolysis 84:131–138. doi: 10.1016/j.jaap.2009.01.002

Quann RJ, Jaffe SB (1996) Building useful models of complex reaction systems in petroleum refining. Chemical Engineering Science 51:1615–1635. doi: 10.1016/0009-2509(96)00023-1

Rangarajan S, Bhan A, Daoutidis P (2010) Rule-Based Generation of Thermochemical Routes to Biomass Conversion. Ind. Eng. Chem. Res. 49:10459–10470. doi: 10.1021/ie100546t

Rangarajan S, Bhan A, Daoutidis P (2012) Language-oriented rule-based reaction network generation and analysis: Description of RING. Computers & Chemical Engineering 45:114–123. doi: 10.1016/j.compchemeng.2012.06.008

Rangarajan S, Bhan A, Daoutidis P (2014) Identification and analysis of synthesis routes in complex catalytic reaction networks for biomass upgrading. Applied Catalysis B: Environmental 145:149–160. doi: 10.1016/j.apcatb.2013.01.030

Ribeiro R (2018) Molecular Reconstruction of Fast-pyrolysis Bio-oil. Master thesis, IST Lisbon

Rita, A. I. B. (2014) Chemical Representation of Various Biomass Compounds. Dissertation, Instituto Superior Técnico

Rout PK, Naik MK, Naik SN, Goud VV, Das LM, Dalai AK (2009) Supercritical CO 2 Fractionation of Bio-oil Produced from Mixed Biomass of Wheat and Wood Sawdust. Energy Fuels 23:6181–6188. doi: 10.1021/ef900663a

Runnebaum RC, Lobo-Lapidus RJ, Nimmanwudipong T, Block DE, Gates BC (2011) Conversion of Anisole Catalyzed by Platinum Supported on Alumina: The Reaction Network. Energy Fuels 25:4776–4785. doi: 10.1021/ef2010699

Sarkanen KV, Ludwig CH (1971) Lignins: Occurrence, formation, structure and reactions. Wiley-Interscience, New York

Sastri S, Rao KK (1992) A new group contribution method for predicting viscosity of organic liquids. The Chemical Engineering Journal 50:9–25. doi: 10.1016/0300-9467(92)80002-R

Sato S (1997) The Development of Support Program for the Analysis of Average Molecular Structures by Personal Computer. Sekiyu Gakkaishi 40:46–51. doi: 10.1627/jpi1958.40.46

Scanlon JT, Willis DE (1985) Calculation of Flame Ionization Detector Relative Response Factors Using the Effective Carbon Number Concept. Journal of Chromatographic Science 23:333–340. doi: 10.1093/chromsci/23.8.333

Schnongs P (2005) Recosntruction moléculaire de coupes pétrolières lourdes, Université Catholique de Louvain

Scholze B, Meier D (2001) Characterization of the water-insoluble fraction from pyrolysis oil (pyrolytic lignin). Part I. PY–GC/MS, FTIR, and functional groups. Journal of Analytical and Applied Pyrolysis 60:41–54. doi: 10.1016/S0165-2370(00)00110-8

Şenol O, Viljava T-R, Krause A (2005) Hydrodeoxygenation of methyl esters on sulphided NiMo/γ-Al2O3 and CoMo/γ-Al2O3 catalysts. Catalysis Today 100:331–335. doi: 10.1016/j.cattod.2004.10.021

Şenol O, Viljava T-R, Krause A (2007) Effect of sulphiding agents on the hydrodeoxygenation of aliphatic esters on sulphided catalysts. Applied Catalysis A: General 326:236–244. doi: 10.1016/j.apcata.2007.04.022

Sfetsas T, Michailof C, Lappas A, Li Q, Kneale B (2011) Qualitative and quantitative analysis of pyrolysis oil by gas chromatography with flame ionization detection and comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry. J Chromatogr A 1218:3317–3325. doi: 10.1016/j.chroma.2010.10.034

Shahrouzi JR, Guillaume D, Rouchon P, Da Costa P (2008) Stochastic Simulation and Single Events Kinetic Modeling: Application to Olefin Oligomerization. Ind. Eng. Chem. Res. 47:4308–4316. doi: 10.1021/ie071215l

Shannon CE (1948) A Mathematical Theory of Comunication. The Bell System Technical Journal 27:379–423

Sheremata JM, Gray MR, Dettman HD, McCaffrey WC (2004) Quantitative Molecular Representation and Sequential Optimization of Athabasca Asphaltenes. Energy Fuels 18:1377–1384. doi: 10.1021/ef049936%2B

Sipilä K, Kuoppala E, Fagernäs L, Oasmaa A (1998) Characterization of biomass-based flash pyrolysis oils. Biomass and Bioenergy 14:103–113. doi: 10.1016/S0961-9534(97)10024-1

Sitthisa S, Sooknoi T, Ma Y, Balbuena PB, Resasco DE (2011) Kinetics and mechanism of hydrogenation of furfural on Cu/SiO2 catalysts. Journal of Catalysis 277:1–13. doi: 10.1016/j.jcat.2010.10.005

Smith EA, Lee YJ (2010) Petroleomic Analysis of Bio-oils from the Fast Pyrolysis of Biomass: Laser Desorption Ionization−Linear Ion Trap−Orbitrap Mass Spectrometry Approach. Energy Fuels 24:5190–5198. doi: 10.1021/ef100629a

Smith EA, Park S, Klein AT, Lee YJ (2012) Bio-oil Analysis Using Negative Electrospray Ionization: Comparative Study of High-Resolution Mass Spectrometers and Phenolic versus Sugaric Components. Energy Fuels 26:3796–3802. doi: 10.1021/ef3003558

Solomons TW, Fryhle CB (2004) Organic chemistry, 8. ed. Wiley, New York

Speight J (1970) A structural investigation of the constituents of Athabasca bitumen by proton magnetic resonance spectroscopy. Fuel 49:76–90. doi: 10.1016/0016-2361(70)90010-4

Staš M, Kubička D, Chudoba J, Pospíšil M (2014) Overview of Analytical Methods Used for Chemical Characterization of Pyrolysis Bio-oil. Energy Fuels 28:385–402. doi: 10.1021/ef402047y

Staš M, Chudoba J, Kubička D, Blažek J, Pospíšil M (2017) Petroleomic Characterization of Pyrolysis Bio-oils: A Review. Energy Fuels 31:10283–10299. doi: 10.1021/acs.energyfuels.7b00826

Strahan GD, Mullen CA, Boateng AA (2011) Characterizing Biomass Fast Pyrolysis Oils by 13 C NMR and Chemometric Analysis. Energy Fuels 25:5452–5461. doi: 10.1021/ef2013166

Surla K, Guillaume D, Verstraete JJ, Galtier P (2011) Kinetic Modeling using the Single-Event Methodology: Application to the Isomerization of Light Paraffins. Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles 66:343–365. doi: 10.2516/ogst/2011119

Suzuki T, Itoh M, Takegami Y, Watanabe Y (1982) Chemical structure of tar-sand bitumens by 13C and 1H n.m.r. spectroscopic methods. Fuel 61:402–410. doi: 10.1016/0016-2361(82)90062-X

Takegami Y, Watanabe Y, Suzuki T, Mitsudo T, Itoh M (1980) Structural investigation on column-chomatographed vaccum residues of various petroleum crudes by 13C nuclear magnetic resonance spectroscopy. Fuel 59:253–259

Tessarolo NS, dos Santos, Luciana R M, Silva RSF, Azevedo DA (2013) Chemical characterization of bio-oils using comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry. J Chromatogr A 1279:68–75. doi: 10.1016/j.chroma.2012.12.052

Tessini C, Muller N, Mardones C, Meier D, Berg A, Baer D von (2012) Chromatographic approaches for determination of low-molecular mass aldehydes in bio-oil. J Chromatogr A 1219:154–160. doi: 10.1016/j.chroma.2011.10.093

Torri C, Fabbri D (2009) Application of off-line pyrolysis with dynamic solid-phase microextraction to the GC–MS analysis of biomass pyrolysis products. Microchemical Journal 93:133–139. doi: 10.1016/j.microc.2009.05.010

Trauth DM (1993) Structure of complex mixtures through characterization, reaction and modeling, University of Delaware

Trauth DM, Stark SM, Petti TF, Neurock M, Klein MT (1994) Representation of the Molecular Structure of Petroleum Resid through Characterization and Monte Carlo Modeling. Energy Fuels 8:576–580. doi: 10.1021/ef00045a010

Tripathi MM, Hassan EBM, Yueh F-Y, Singh JP, Steele PH, Ingram LL (2009) Reflection–absorption-based near infrared spectroscopy for predicting water content in bio-oil. Sensors and Actuators B: Chemical 136:20–25. doi: 10.1016/j.snb.2008.10.055

Ugi I, Bauer J, Bley K, Dengler A, Dietz A, Fontain E, Gruber B, Herges R, Knauer M, Reitsam K, Stein N (1993) Computer-Assisted Solution of Chemical Problems—The Historical Development and the Present State of the Art of a New Discipline of Chemistry. Angew. Chem. Int. Ed. Engl. 32:201–227. doi: 10.1002/anie.199302011

Vaidya PD, Mahajani VV (2003) Kinetics of Liquid-Phase Hydrogenation of Furfuraldehyde to Furfuryl Alcohol over a Pt/C Catalyst. Ind. Eng. Chem. Res. 42:3881–3885. doi: 10.1021/ie030055k

Van Geem, K. M., Hudebine D, Reyniers MF, Wahl F, Verstraete JJ, Marin GB (2007) Molecular reconstruction of naphtha steam cracking feedstocks based on commercial indices. Computers & Chemical Engineering 31:1020–1034. doi: 10.1016/j.compchemeng.2006.09.001

Vandewiele NM, Van Geem, Kevin M., Reyniers M-F, Marin GB (2012) Genesys: Kinetic model construction using chemo-informatics. Chemical Engineering Journal 207-208:526–538. doi: 10.1016/j.cej.2012.07.014

Venderbosch RH, Ardiyanti AR, Wildschut J, Oasmaa A, Heeres HJ (2010) Stabilization of biomass-derived pyrolysis oils. J. Chem. Technol. Biotechnol. 85:674–686. doi: 10.1002/jctb.2354

Verstraete JJ, Revellin N, Dulot H, Hudebine D (2004) Molecular Reconstruction of Vacuum Gas Oils. Abstr. Pap. Am. Chem. Soc.

Verstraete JJ, Schnongs P, Dulot H, Hudebine D (2010) Molecular reconstruction of heavy petroleum residue fractions. Chemical Engineering Science 65:304–312. doi: 10.1016/j.ces.2009.08.033

Vinu R, Broadbelt LJ (2012) Unraveling reaction pathways and specifying reaction kinetics for complex systems. Annu Rev Chem Biomol Eng 3:29–54. doi: 10.1146/annurev-chembioeng-062011-081108

Vispute TP, Huber GW (2009) Production of hydrogen, alkanes and polyols by aqueous phase processing of wood-derived pyrolysis oils. Green Chem. 11:1433. doi: 10.1039/b912522c

Wan H, Chaudhari RV, Subramaniam B (2013) Aqueous Phase Hydrogenation of Acetic Acid and Its Promotional Effect on p -Cresol Hydrodeoxygenation. Energy Fuels 27:487–493. doi: 10.1021/ef301400c

Wang J, Cui H, Wei S, Zhuo S, Wang L, Li Z, Yi W (2010) Separation of Biomass Pyrolysis Oil by Supercritical CO2 Extraction. SGRE 01:98–107. doi: 10.4236/sgre.2010.12015

Wang S (2013) High-Efficiency Separation of Bio-Oil. In: Matovic MD (ed) Biomass Now - Sustainable Growth and Use. InTech

Wang S, Gu Y, Liu Q, Yao Y, Guo Z, Luo Z, Cen K (2009) Separation of bio-oil by molecular distillation. Fuel Processing Technology 90:738–745. doi: 10.1016/j.fuproc.2009.02.005

Wang H, Male J, Wang Y (2013) Recent Advances in Hydrotreating of Pyrolysis Bio-Oil and Its Oxygen-Containing Model Compounds. ACS Catal. 3:1047–1070. doi: 10.1021/cs400069z

Wei W, Bennett CA, Tanaka R, Hou G, Klein MT (2008) Computer aided kinetic modeling with KMT and KME. Fuel Processing Technology 89:350–363. doi: 10.1016/j.fuproc.2007.11.015

Weininger D (1988) SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J. Chem. Inf. Model. 28:31–36. doi: 10.1021/ci00057a005

Weininger D, Weininger A, Weininger JL (1989) SMILES. 2. Algorithm for generation of unique SMILES notation. J. Chem. Inf. Model. 29:97–101. doi: 10.1021/ci00062a008

Wildschut J, Arentz J, Rasrendra CB, Venderbosch RH, Heeres HJ (2009) Catalytic hydrotreatment of fast pyrolysis oil: Model studies on reaction pathways for the carbohydrate fraction. Environ. Prog. Sustainable Energy 28:450–460. doi: 10.1002/ep.10390

Wildschut J (2009) Pyrolysis oil upgrading to transportation fuels by catalytic hydrotreatment. s.n.]; University Library Groningen], [S.l., [Groningen

Xu J, Jiang J, Lv W, Dai W, Sun Y (2010) Rice husk bio-oil upgrading by means of phase separation and the production of esters from the water phase, and novolac resins from the insoluble phase. Biomass and Bioenergy 34:1059–1063. doi: 10.1016/j.biombioe.2010.01.040

Xu Q, Zhang Z, Zhang S, Wang F, Yan Y (2014) Molecular Structure Models of Asphaltene in Crude and Upgraded Bio-Oil. Chem. Eng. Technol. 37:1198–1204. doi: 10.1002/ceat.201300158

Yakovlev VA, Khromova SA, Sherstyuk OV, Dundich VO, Ermakov D, Novopashina VM, Lebedev M, Bulavchenko O, Parmon VN (2009) Development of new catalytic systems for upgraded bio-fuels production from bio-crude-oil and biodiesel. Catalysis Today 144:362–366. doi: 10.1016/j.cattod.2009.03.002

Zhang Y (1999) A Molecular Approach for Characterization and Property Predictions of Petroleum Mixtures with Applications to Refinery Modelling, University of Manchester

Zhang S, Yan Y, Ren Z, Li T (2003) Study of Hydrodeoxygenation of Bio-Oil from the Fast Pyrolysis of Biomass. Energy Sources 25:57–65. doi: 10.1080/00908310290142118

Zhang S, Yan Y, Li T, Ren Z (2009) Lumping Kinetic Model for Hydrotreating of Bio-oil from the Fast Pyrolysis of Biomass. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 31:639–645. doi: 10.1080/15567030701750523

Zhang L, Hou Z, Horton SR, Klein MT, Shi Q, Zhao S, Xu C (2014) Molecular Representation of Petroleum Vacuum Resid. Energy Fuels 28:1736–1749. doi: 10.1021/ef402081x

Glossary

µKE	Microkinetic Engine
A	Acetone
AC	Acetone Catalyst
ACE	Acetol
AL	Acetone Liquid
APCI	Atmospheric Pressure Chemical Ionization
Ar	Aromatic
ASTM	ASTM International
B	Benzenediol
B	Butanol
BA	Butanal
BD	Building Diagram
BO	Bio-Oil
BtL	Biomass To Liquid
C	Catalyst
C	Cyclohexane
CAT	Catechol
CC	Clean Catalyst
CDF	Cumulative Distribution Function
CDM	Carbon Dioxide Methanation
CHONS	Carbon, Hydrogen, Oxygen, Nitrogen And Sulfur Analysis
CL	Centrifuge Liquid
CL	Cyclohexanol
CME	Composition Model Editor / Chemical Master Equation
CMM	Carbon Monoxide Methanation
CO	Carbon Monoxide
LC	Liquid Chromatography
LR	Liquid Residues
M	Methanol
M4	Methane
MBL	Methylbenzenediol
MBL	Methoxybenzene
MC	Monte Carlo
MET	Methylation
METD	Methanol Decomposition
METSR	Methanol Steam Reforming
MF / 2MF	Methylfuran
MLC	Multidimensional Liquid Chromatography
MLC	Methylphenol
MR	Molecular Reconstruction
MS	Mass Spectroscopy
MS	Methyl Shift
MTHS	Method Type And Homologous Series
MX	Methoxycyclohexane
MY	Methylcyclohexanediol
NHV	Net Heating Value
NMR	Nuclear Magnetic Ressonance
ODE	Ordinary Differential Equation
OF	Objective Function
P	Phenol
PA	Pentanal
PCP	Protonated Cyclo Propane

CO	Decarbonylation
CO2	Decarboxylation
COH	Carbonyl Hydrogenation
CPU	Central Process Unit
Cr	Cracking
CY	Cyclohexanediol
DEM	Demethylation
DEMO	Demethoxylation
DH	Dehydrogenation
DH	Hydrogenolysis
DHYD	Dehydration
EA	Elemental Analysis
ECN	Effective Carbon Number
EM	Entropy Maximization
ESI	Electrospray Ionization
F	Furfural
FA	Furfuryl Alcohol
FID	Flame Ionization Detector
FL	Filtration Liquid
FN	Furan
FT-ICR	Fourier-Transform Ion Cyclotron Resonance
G	Guaiacol
GA	Genetic Algorithm
GC	Gas Chromatography
GPC	Gel Permeation Chromatography
GRG	Generalized Reduced Gradient
HDO	Hydrodeoxygenation
HDT	Hydrotreating / Hydrogenation
HMF	Hydroxymethylfuran
PDF	Probablility Distribution Function
PIONA	Paraffins, Isoparaffins, Olefins, Naphthenes, Aromatics
PL	Pentanol
PO	Pentanone
PR	Protonation
PY	Pentanediol
PY-IR	Pyrolysis-Infrared Detection
QMR	Quantitative Molecular Representation
QSRC	Quantitative Structure / Reactivity Correlation
REM	Reconstruction By Entropy Maximization
RI	Refraction Index
RING	Rule Input Network Generator
RL	Rotavap Liquid
RN	Random Number
RO	Ring Opening
S	Solids
SARA	Saturated, Aromatic, Resins, Asphalthenes
SAT	Saturation
SEC	Size Exclusion Chromatography
SL	Soiled Liquid
SMR	Stochastic Molecular Reconstruction
SOL	Structure Oriented Lumping
SR	Stochactic Reconstruction
SSA	Stochastic Simulation Algorithm
T	Tetrahydrofuran
TA	Tetrahydrofuran Alcohol
TAN	Total Acid Number
TCD	Thermal Conductivity Detector
THF	Tetrahydrofuran

HPLC	High-Performance Liquid Chromatography	THP	Tetrahydropyran
HRMS	High Resolution Mass Spectrometry	TOF	Time Of Flight
HS	Hydride Shift	TRANS	Transalkylation
HYD	Hydration	UC	Unrefined Catalyst
IEA	International Energy Association	UHDT	Unsaturated Hydrogenation
InGen	Interactive Generator	UL	Unrefined Liquid
kMC	Kinetic Of Monte Carlo	WGS	Water Gas Swift

APPENDIX

Appendix A – Experimental results

A.1. Guaiacol

Table 4 - Experimental results of the hydrotreating tests of guaiacol.

Temperature (°C)	**150**			**200**	**250**		**300**			
Reaction Time (h)	**1**	**3**	**5**	**3**	**1**	**3**	**0**	**1**	**3**	**5**
Fractions	**Carbon balance (%)**									
Liquid (%)	95.11	96.74	93.22	86.73	82.95	88.46	91.58	85.91	85.97	84.55
Gas (%)	0.00	0.00	0.00	0.03	0.02	0.03	0.03	0.03	0.14	0.22
Non-Detected (%)	4.89	3.26	6.78	13.24	17.03	11.51	8.39	14.05	13.89	15.23
Reactants	**Quantity (g)**									
Guaiacol initial	45.06	45.01	45.16	45.05	45.03	44.97	45.05	44.99	45.00	45.01
Guaiacol final	42.81	43.53	42.04	37.26	35.03	35.46	39.91	33.82	30.33	26.88
Consumption	2.25	1.48	3.12	7.79	10.00	9.51	5.14	11.17	14.67	18.13
Guaiacol conversion (%)	5.00	3.29	6.91	17.30	22.20	21.14	11.42	24.83	32.60	40.27
Hydrogen initial	3.93	4.07	4.07	3.25	2.29	2.45	1.01	1.14	1.17	1.25
Hydrogen final	2.90	2.92	2.87	2.69	1.92	1.91	1.01	1.08	1.01	1.01
Hydrogen consumption	1.03	1.15	1.20	0.56	0.37	0.53	0.00	0.07	0.16	0.23
Compounds	**Selectivity (%)**									
Methylcylcohexanediol	0.00	0.00	0.00	53.15	40.31	36.69	21.79	8.04	4.33	5.47
Methanol	25.66	80.76	32.73	15.06	18.67	18.81	23.62	27.46	25.93	24.94
5,7-Methylcylcopentane	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Cyclohexane	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.09	0.24
Methoxycyclohexane	2.97	8.72	7.09	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Cyclohexanol	0.00	0.00	0.00	8.34	7.87	7.35	2.74	2.33	2.52	3.93
Cyclohexanone	0.00	0.00	0.00	0.12	0.39	0.24	1.15	2.28	4.36	5.84
Cyclopentanemethanol	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.30	0.51	0.87
Methoxybenzene	0.00	0.00	0.00	0.97	0.70	0.48	0.14	0.11	0.11	0.14
2-Methylcyclohexanone	0.00	0.00	0.00	0.13	0.00	0.19	0.00	0.07	0.15	0.31
Phenol	3.17	5.26	3.81	0.69	5.95	3.82	5.38	5.00	8.00	10.58
2-Hydroxycyclohexanone	0.00	0.00	0.00	6.96	0.00	0.07	0.74	0.98	0.57	0.66
2-Methoxycyclohexanone	0.00	0.00	0.00	0.64	1.51	0.76	6.10	1.12	0.32	0.13
2-Methylphenol	56.26	5.26	46.84	3.10	2.26	1.13	4.87	1.71	1.03	0.26
Cyclohexanediol	0.00	0.00	0.00	9.93	17.78	27.71	10.49	11.40	8.02	8.58
Benzendiol	0.00	0.00	0.00	0.65	4.08	1.93	22.06	37.16	41.35	34.93
3-Methyl-1,2-benzendiol	0.00	0.00	0.00	0.00	0.07	0.05	0.38	1.19	1.79	2.05
Dimethylcyclopentanone	11.95	0.00	9.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Methane	0.00	0.00	0.00	0.26	0.42	0.55	0.53	0.86	0.91	1.04

A.2. Furfural

Table 5 - Experimental results of the hydrotreating tests of furfural

Temperature (°C)	**100**		**150**	**250**
Reaction Time (h)	**0**	**1**	**1**	**0**
Fractions	**Carbon balance (%)**			
Liquid (%)	83.56	82.48	76.93	61.27
Gas (%)	0.00	0.00	0.00	1.61
Non Detected (%)	16.44	17.52	23.07	37.11
Reactants	**Quantity (g)**			
Furfural initial	19.55	19.60	19.60	20.00
Furfural final	14.64	10.07	0.55	9.41
Consumption	4.91	9.53	19.05	10.59
Furfural conversion (%)	25.12	48.61	97.17	52.96
Hydrogen initial	4.57	4.64	4.45	0.97
Hydrogen final	3.24	3.16	2.76	0.67
Hydrogen consumption	1.32	1.47	1.69	0.30
Compounds	**Selectivity (%)**			
Furan	0.77	0.29	0.19	8.71
Pent-1-ene	0.00	0.00	0.00	0.43
2-Methylfuran	0.20	0.37	2.15	5.52
THF	0.00	0.00	0.00	0.04
Butanol	0.01	0.00	0.02	0.27
Cyclopentanol	0.00	0.00	0.00	36.89
Butan-2-one	0.00	0.00	0.00	8.10
Furfuryl alcohol	95.37	97.35	92.54	9.47
Cyclohexanol	0.00	0.00	0.00	0.36
THF alcohol	0.93	1.16	4.55	0.49
Butanediol	0.00	0.00	0.00	1.31
Pentane-1,5-diol	0.00	0.00	0.00	4.91
Methylcyclohexanediol	0.00	0.00	0.00	0.08
2-Methoxyphenol	0.66	0.05	0.04	4.59
1-Cyclohexane-1-acetaldehyde-2-dimethyl	0.00	0.00	0.00	1.27
Bicyclopentyl-2-one	0.00	0.00	0.00	0.61
Bicyclopentyl	0.00	0.00	0.00	0.36
Methylfurfural	0.20	0.04	0.02	0.00
Methylphenol	0.01	0.08	0.01	0.00
Acetylfuran	0.95	0.26	0.12	0.00
2,4-Difurfurylfuran	0.13	0.13	0.13	0.00
2-Furanmethanol, 5-(2-furanylmethyl)-	0.00	0.00	0.00	0.00
Difurfurylether	0.08	0.05	0.13	0.00
Methane	0.00	0.00	0.00	0.13
Carbon monoxide	0.00	0.00	0.00	2.26
Carbon dioxide	0.00	0.00	0.00	14.19

Pentanone	0.40	0.12	0.04	0.00
Pentanal	0.09	0.03	0.01	0.00
Butanal	0.18	0.05	0.00	0.00
Pentanol	0.03	0.01	0.04	0.00

Appendix B – Deterministic results

B.1. Guaiacol

Table 6 - Results of the deterministic model for the HDT of guaiacol.

Temperature (°C)	150			200	250		300			
Reaction time (h)	1	3	5	3	1	3	0	1	3	5
Compounds	Simulated concentration (mol/L)									
Guaiacol	7.2E-01	7.0E-01	6.9E-01	5.5E-01	4.1E-01	4.1E-01	5.7E-01	5.2E-02	3.0E-04	2.0E-06
Benzenediol	3.8E-04	9.7E-04	1.5E-03	8.0E-03	1.7E-02	1.7E-02	1.1E-02	4.6E-02	4.8E-02	4.7E-02
Cyclohexane	1.3E-09	2.2E-08	9.3E-08	9.1E-06	4.3E-05	4.3E-05	1.1E-06	2.9E-05	8.0E-05	1.3E-04
Cyclohexanediol	4.8E-06	3.1E-05	8.2E-05	1.6E-03	3.6E-03	3.6E-03	4.3E-04	2.1E-03	2.2E-03	2.1E-03
Cyclohexanol	3.1E-05	2.0E-04	5.2E-04	8.1E-03	1.6E-02	1.6E-02	1.7E-03	8.0E-03	8.8E-03	8.8E-03
Methane	3.9E-04	1.0E-03	1.6E-03	9.6E-03	2.1E-02	2.1E-02	1.1E-02	5.2E-02	6.4E-02	7.1E-02
Methanol	2.5E-03	6.4E-03	1.0E-02	4.9E-02	8.9E-02	8.9E-02	4.4E-02	1.8E-01	1.9E-01	1.8E-01
Methylbenzenediol	5.7E-03	1.4E-02	2.3E-02	9.5E-02	1.7E-01	1.7E-01	9.7E-02	4.1E-01	4.4E-01	4.2E-01
Methylcyclohexanediol	7.2E-05	4.6E-04	1.2E-03	1.9E-02	3.6E-02	3.6E-02	3.9E-03	1.8E-02	1.9E-02	1.9E-02
Methylphenol	3.6E-07	2.4E-06	6.2E-06	1.7E-04	7.2E-04	7.2E-04	1.7E-04	8.9E-03	2.8E-02	4.4E-02
Methoxybenzene	4.5E-05	1.1E-04	1.8E-04	8.1E-04	1.5E-03	1.5E-03	1.0E-03	1.5E-03	6.3E-05	2.0E-06
Methoxycyclohexane	5.6E-07	3.7E-06	9.5E-06	1.6E-04	3.1E-04	3.1E-04	4.0E-05	6.6E-05	2.9E-06	8.9E-08
Phenol	2.5E-03	6.2E-03	9.8E-03	4.1E-02	7.3E-02	7.4E-02	4.2E-02	1.8E-01	2.0E-01	2.0E-01

B.2. Furfural

Table 7 - Results of the deterministic model for the HDT of furfural.

Temperature (°C)	100		150	250
Reaction time (h)	0	1	1	0
Compounds	Simulated concentration (mol/L)			
Furfural	3.3E-01	2.0E-01	3.7E-03	3.2E-03
Butanal	5.4E-07	5.4E-06	3.4E-05	2.7E-05
Butanol	3.9E-08	1.5E-06	1.7E-04	5.8E-04
Carbon Monoxide	3.5E-05	1.1E-04	3.6E-04	7.0E-04
Furan	3.5E-05	1.0E-04	1.5E-04	7.8E-05
Furfuryl alcohol	7.7E-02	2.1E-01	3.6E-01	2.5E-01
Methylfuran	3.6E-05	4.2E-04	1.2E-02	3.1E-02
Pentanal	4.2E-07	1.5E-05	2.0E-03	7.6E-03
Pentanediol	1.2E-06	4.7E-05	9.1E-03	3.7E-02
Pentanol	2.4E-08	3.1E-06	4.6E-03	6.1E-02
Pentanone	4.5E-07	1.9E-05	6.4E-03	6.4E-02
Tetrahydrofuran alcohol	1.1E-04	1.1E-03	1.8E-02	2.1E-02

Tetrahydrofuran	4.9E-08	5.5E-07	1.0E-05	1.0E-05

Appendix C – Stochastic results

It is important to note that, the SSA there is not a molecules identifier. The identification of a molecules requires a direct coding in the source code. Therefore, only the molecules used for the estimation of parameters were implemented in the source code of the SSA. Thus, only these have a value of concentration.

C.1. Guaiacol

The temperature profiles will not be presented in here as they are too large to be summarized in the present work in the form of a table.

C.1.1. Deterministic kinetic model

Table 8 - Results of the deterministic model for the stochastic simulation of the HDT of guaiacol.

Temperature (°C)	**150**			**200**	**250**		**300**			
Reaction time (h)	**1**	**3**	**5**	**3**	**1**	**3**	**0**	**1**	**3**	**5**
Compounds	**Simulated concentration (mol/L)**									
Guaiacol	6.8E-01	6.2E-01	5.7E-01	2.9E-01	2.2E-01	2.2E-01	4.2E-01	1.1E-01	1.1E-01	1.1E-01
Benzenediol	3.2E-03	7.4E-03	1.1E-02	3.4E-02	4.5E-02	4.5E-02	3.9E-02	5.3E-02	5.3E-02	5.2E-02
Cyclohexanol	1.3E-04	7.1E-04	1.6E-03	2.4E-02	4.1E-02	4.1E-02	1.2E-02	8.4E-02	8.6E-02	8.7E-02
Methane	3.4E-03	8.5E-03	1.3E-02	6.9E-02	1.1E-01	1.1E-01	5.7E-02	1.8E-01	1.9E-01	1.9E-01
Methanol	3.0E-03	7.2E-03	1.1E-02	6.0E-02	9.3E-02	9.3E-02	4.9E-02	1.6E-01	1.6E-01	1.6E-01
Methylcyclohexanediol	3.8E-02	9.1E-02	1.4E-01	3.0E-01	3.0E-01	3.0E-01	2.0E-01	2.7E-01	2.6E-01	2.6E-01

C.1.2. Stochastic kinetic model

Table 9 - Results of the stochastic model for the HDT of guaiacol.

Temperature (°C)	**150**			**200**	**250**		**300**			
Reaction time (h)	**1**	**3**	**5**	**3**	**1**	**3**	**0**	**1**	**3**	**5**
Compounds	**Simulated concentration (mol/L)**									
Guaiacol	7.2E-01	7.1E-01	7.0E-01	5.9E-01	5.0E-01	5.0E-01	6.2E-01	2.2E-01	1.0E-01	7.6E-02
Benzenediol	1.1E-03	2.6E-03	4.2E-03	2.5E-02	4.0E-02	4.0E-02	2.0E-02	7.3E-02	5.1E-02	3.8E-02
Cyclohexanol	7.3E-06	3.0E-05	8.3E-05	3.3E-03	8.6E-03	8.6E-03	1.7E-03	6.4E-02	1.4E-01	1.6E-01
Methane	1.1E-03	2.6E-03	4.2E-03	2.6E-02	4.4E-02	4.4E-02	2.1E-02	1.1E-01	1.3E-01	1.3E-01
Methanol	3.1E-03	8.0E-03	1.3E-02	7.8E-02	1.3E-01	1.3E-01	6.3E-02	3.1E-01	3.6E-01	3.5E-01
Methylcyclohexanediol	1.7E-03	4.2E-03	6.8E-03	3.4E-02	4.9E-02	4.9E-02	2.3E-02	8.9E-02	1.3E-01	1.5E-01

C.2. Furfural

The temperature profiles will not be presented in here as they are too large to be summarized in the present work in the form of a table.

C.2.1. Deterministic kinetic model

Table 10 - Results of the deterministic model for the stochastic simulation of the HDT of furfural.

Temperature (°C)	**100**		**150**	**250**
Reaction time (h)	**0**	**1**	**1**	**0**
Compounds	**Simulated concentration (mol/L)**			
Furfural	3.0E-01	1.4E-01	1.1E-04	0.0E+00
Butanol	0.0E+00	0.0E+00	0.0E+00	0.0E+00
Furan	3.3E-05	1.0E-04	3.3E-05	#DIV/0!
Furfuryl alcohol	7.7E-02	1.7E-01	3.6E-02	#DIV/0!
Methylfuran	5.0E-05	3.8E-04	2.2E-03	#DIV/0!
Tetrahydrofuran alcohol	7.7E-03	4.9E-02	2.7E-02	#DIV/0!

C.2.2. Stochastic kinetic model

Table 11 - Results of the stochastic model for the HDT of furfural.

Temperature (°C)	**100**		**150**	**250**
Reaction time (h)	**0**	**1**	**1**	**0**
Compounds	**Simulated concentration (mol/L)**			
Furfural	3.2E-01	1.8E-01	9.7E-04	1.6E-06
Butanol	0.0E+00	0.0E+00	0.0E+00	0.0E+00
Furan	2.4E-04	6.5E-04	1.7E-03	2.0E-03
Furfuryl alcohol	8.3E-02	2.1E-01	3.1E-01	2.0E-01
Methylfuran	1.5E-05	1.7E-04	5.6E-03	2.2E-02
Tetrahydrofuran alcohol	8.3E-04	5.7E-03	1.6E-02	1.5E-02

Appendix D – Stochastic reconstruction results

D.1. Building diagram #1

D.1.1. Estimated data

Table 12 - Results of the stochastic reconstruction with building diagram #1 for sample A and B.

Analysis		Sample A					Sample B				
		Exp.	Pred.	Abs. Error	Rel. Error (%)	Glo. Error (%)	Exp.	Pred.	Abs. Error	Rel. Error (%)	Glo. Error (%)
Elemental analysis	% Carbon	53.79	49.83	3.96	7.35	8	54.12	49.36	4.76	8.80	8
	% Hydrogen	6.48	6.09	0.39	6.06		6.67	6.54	0.13	1.89	
	% Oxygen	39.73	44.08	4.35	10.94		39.21	44.10	4.89	12.47	
13C NMR analysis	Aromatic carbons	30.20	36.11	5.91	19.57	16	31.20	31.33	0.13	0.42	5
	Saturated carbons	17.90	22.78	4.88	27.26		25.50	27.80	2.30	9.02	
	Aldehydes & Ketones	6.20	6.34	0.14	2.26		5.90	5.80	0.10	1.69	
	Carboxyls	4.10	3.96	0.14	3.41		6.70	6.96	0.26	3.88	
	Methoxyls & Hydroxyls	41.60	30.82	10.78	25.91		30.70	28.11	2.59	8.44	
SEC analysis	10%	57.2	97.6	40.4	70.6	44	60.1	97.6	37.50	62.4	45
	20%	97.6	97.6	0.0	0.0		96.9	97.6	0.70	0.7	
	30%	136.5	124.4	12.1	8.9		143.6	107.4	36.20	25.2	
	40%	180.7	136.8	43.9	24.3		171.0	122.2	48.80	28.5	
	50%	235.0	146.6	88.4	37.6		220.2	135.2	85.00	38.6	
	60%	293.4	155.1	138.3	47.1		278.9	148.3	130.60	46.8	
	70%	364.2	172.2	192.0	52.7		341.3	164.0	177.30	51.9	
	80%	470.3	198.4	271.9	57.8		430.6	179.7	250.90	58.3	
	90%	652.2	229.2	423.0	64.9		597.2	211.9	385.30	64.5	
	95%	1019.0	264.0	755.0	74.1		1091.1	248.6	842.50	77.2	
	100%										

Exp. – Experimental; Pred. – Predicted; Abs. – Absolute; Rel. – Relative; Glo. – Global.

$$Abs.Error = \left|x_i^{exp} - x_i^{cal}\right| \qquad Rel.Error = \left(Abs.Error/x_i^{exp}\right)\cdot 100 \qquad Glo.Error = \sum_{i}^{A} Rel.Error \Big/ A$$

x_i^{exp} experimental point i of a given analysis, x_i^{exp} calculated point i of a given analysis, A number of points in a given analysis.

D.1.2. Validation data

Table 13 - [Continuation] Results of the stochastic reconstruction with building diagram #1 for sample A and B.

		Sample A					Sample B				
		Exp.	Pred.	Abs. Error	Rel. Error (%)	Glo. Error (%)	Exp.	Pred.	Abs. Error	Rel. Error (%)	Glo. Error (%)
Density (g/cm^3)		1.24	1.3058	0.07	5.31	5	1.19	1.2296	0.04	3.33	3
Molecular Weight (g/mol)			103.4					90.4			
Oxygenated families (%) (mol/mol)	Hydrocarbon		0.0					0.1			
	Monoalcohol		21.3					23.3			
	Monoaldehyde		0.1					0.1			
	Monoacid		0.0					0.3			
	Di_Oxygen_FG		25.4					34.0			
	Dialcohol		26.3					23.8			
	Dialdehyde		1.5					1.7			
	Diacid		0.7					2.4			
	Poly_Oxygen_FG		18.2					10.9			
	Polyalcohol		6.4					3.4			
	Polyaldehyde		0.0					0.0			
	Polyketone		0.0					0.0			
	Polyacid		0.0					0.1			
$O_xN_yS_z$ Families - FT-ICR-MS (%)(mol/mol)	O0	0.0	0.0	0.02		3063	0.00	0.1	0.1		775
	O1	0.0	21.3	21.2	43644.9		0.36	23.4	23.0	6436.3	
	O2	1.0	41.5	40.5	4062.2		1.60	39.4	37.8	2366.0	
	O3	3.5	14.3	10.9	313.3		4.26	22.4	18.1	425.7	
	O4	8.1	5.6	2.5	31.0		8.48	6.0	2.5	29.0	
	O5	12.4	5.6	6.7	54.4		13.00	3.3	9.7	74.8	
	O6	14.1	4.7	9.4	66.8		17.49	2.2	15.3	87.2	
	O7	15.3	3.1	12.2	79.9		20.82	1.5	19.3	92.7	
	O8	12.8	1.5	11.2	88.1		15.97	0.8	15.2	95.1	
	O9	10.4	0.9	9.5	91.7		9.14	0.4	8.7	95.6	
	O10	8.3	0.7	7.6	91.1		4.81	0.3	4.5	92.9	
	O11	5.9	0.4	5.4	92.8		2.44	0.1	2.4	96.7	
	O12	3.8	0.2	3.7	95.8		1.04	0.1	1.0	92.3	
	O13	2.2	0.1	2.0	93.5		0.18	0.0	0.2	88.7	
	O14	1.3	0.1	1.2	92.2		0.00		0.0		
	O15	0.6	0.0	0.6	100.0		0.00		0.0		
	O16	0.0	0.0	0.0	7.9		0.00		0.0		
	O17	0.0	0.0	0			0.00		0.0		

Exp. – Experimental; Pred. – Predicted; Abs. – Absolute; Rel. – Relative; Glo. – Global; FG – Functional group; O – Oxygen; N – Nitrogen; S – Sulfur; Mono – one FG; Di – two FG; Poly – three or more FG.

D.2. Building diagram #2

D.2.1. Estimated data

Table 14 - Results of the stochastic reconstruction with building diagram #2 for sample A and B.

Analysis		Sample A					Sample B				
		Exp.	Pred.	Abs. Error	Rel. Error (%)	Glo. Error (%)	Exp.	Pred.	Abs. Error	Rel. Error (%)	Glo. Error (%)
Elemental analysis	% Carbon	53.79	51.15	2.64	4.91	6	54.12	52.91	1.21	2.23	2
	% Hydrogen	6.48	6.19	0.29	4.43		6.67	6.69	0.02	0.27	
	% Oxygen	39.73	42.66	2.93	7.37		39.21	40.40	1.19	3.04	
13C NMR analysis	Aromatic carbons	30.20	37.85	7.65	25.33	17	31.20	31.70	0.50	1.60	11
	Saturated carbons	17.90	22.17	4.27	23.85		25.50	31.82	6.32	24.78	
	Aldehydes & Ketones	6.20	6.46	0.26	4.19		5.90	5.84	0.06	1.02	
	Carboxyls	4.10	4.07	0.03	0.73		6.70	6.94	0.24	3.58	
	Methoxyls & Hydroxyls	41.60	29.45	12.15	29.21		30.70	23.70	7.00	22.80	
SEC analysis	10%	57.2	97.6	40.4	70.6	46	60.1	97.6	37.50	62.4	40
	20%	97.6	102.1	4.5	4.6		96.9	97.6	0.70	0.7	
	30%	136.5	124.4	12.1	8.9		143.6	112.2	31.40	21.9	
	40%	180.7	135.2	45.5	25.2		171.0	127.5	43.50	25.4	
	50%	235.0	141.6	93.4	39.7		220.2	143.3	76.90	34.9	
	60%	293.4	148.3	145.1	49.5		278.9	163.1	115.80	41.5	
	70%	364.2	157.8	206.4	56.7		341.3	185.5	155.80	45.6	
	80%	470.3	176.9	293.4	62.4		430.6	219.4	211.20	49.0	
	90%	652.2	206.6	445.6	68.3		597.2	285.0	312.20	52.3	
	95%	1019.0	243.9	775.1	76.1		1091.1	339.8	751.30	68.9	
	100%										

Exp. – Experimental; Pred. – Predicted; Abs. – Absolute; Rel. – Relative; Glo. – Global.

D.2.2. Validation data

Table 15 - [Continuation] Results of the stochastic reconstruction with building diagram #2 for sample A and B.

		Sample A					Sample B				
		Exp.	Pred.	Abs. Error	Rel. Error (%)	Glo. Error (%)	Exp.	Pred.	Abs. Error	Rel. Error (%)	Glo. Error (%)
Density (g/cm^3)		1.24	1.2808	0.04	3.29	3	1.19	1.2424	0.05	4.40	4
Molecular Weight (g/mol)			99.8					106.5			
Oxygenated families (%)(mol/mol)	Hydrocarbon		0.1					0.1			
	Monoalcohol		21.1					21.9			
	Monoaldehyde		0.3					0.8			
	Monoacid		0.2					0.9			
	Di_Oxygen_FG		25.6					33.5			
	Dialcohol		25.2					21.5			
	Dialdehyde		1.8					2.2			
	Diacid		0.7					2.9			
	Poly_Oxygen_FG		18.2					13.4			
	Polyalcohol		6.8					2.8			
	Polyaldehyde		0.0					0.0			
	Polyketone		0.0					0.1			
	Polyacid		0.1					0.1			
$O_xN_yS_z$ Families - FT-ICR-MS (%)(mol/mol)	O0	0.0	0.1	0.1		3091	0.00	0.1	0.1		743
	O1	0.0	21.4	21.4	43932.9		0.36	22.6	22.3	6234.8	
	O2	1.0	41.3	40.3	4042.1		1.60	37.6	36.0	2253.3	
	O3	3.5	16.6	13.2	380.3		4.26	22.7	18.4	431.8	
	O4	8.1	7.3	0.8	10.0		8.48	5.7	2.8	32.5	
	O5	12.4	5.6	6.8	54.9		13.00	3.5	9.5	72.9	
	O6	14.1	3.3	10.8	76.6		17.49	2.3	15.2	86.9	
	O7	15.3	2.2	13.1	85.4		20.82	1.9	18.9	90.7	
	O8	12.8	1.0	11.7	91.8		15.97	1.2	14.8	92.5	
	O9	10.4	0.6	9.8	94.6		9.14	1.0	8.2	89.3	
	O10	8.3	0.2	8.1	97.4		4.81	0.5	4.3	90.0	
	O11	5.9	0.2	5.7	97.3		2.44	0.3	2.2	88.5	
	O12	3.8	0.1	3.7	96.4		1.04	0.2	0.8	80.7	
	O13	2.2	0.0	2.1	99.1		0.18	0.1	0.0	20.8	
	O14	1.3	0.0	1.3	98.4		0.00	0.1	0.1		
	O15	0.6	0.0	0.6	100.0		0.00	0.0	0.0		
	O16	0.0	0.0	0.0	100.0		0.00	0.1	0.1		
	O17	0.0	0.0	0.0			0.00	0.0	0.0		

Exp. – Experimental; Pred. – Predicted; Abs. – Absolute; Rel. – Relative; Glo. – Global; FG – Functional group; O – Oxygen; N – Nitrogen; S – Sulfur; Mono – one FG; Di – two FG; Poly – three or more FG.

Appendix E – Library of molecules

E.1. Building diagram #1

Table 16 - Identification of the molecules reconstructed by BD#1 for sample A and B.

Molecule	Name	
	Sample A	**Sample B**
a)	Methanol	Methanol
b)	Ethane-1,2-diol	Acetaldehyde
c)	Formic acid	Methanediol
d)	Hydroxyacetaldehyde	Hydroxyacetaldehyde
e)	3-Hydroxypropanal	3-Hydroxypropanal
f)	Propane-1,3-diol	Propanedial
g)	4-Hydroxybutanal	4-Oxobutanoic acid
h)	Catechol	Furan-2,3-diol
i)	(2R,4S,5S,6R)-2,4,5,6-tetrahydroxytetrahydro-2H-pyran-2-carbaldehyde	2,4-dihydroxybenzaldehyde
j)	6-hydroxy-1-benzofuran-3-carboxylic acid	5'-formyl-2',4'-dihydroxybiphenyl-2-carboxylic acid
k)	2,3,4,5,8-pentahydroxyanthracene-1-carbaldehyde	1-benzofuran-4-ol
l)	1/C39H36O12/c40-18-22-10-6-9-21-17-30(43)33-24(12-4-5-13-26-34(44)28(19-41)36(46)37(47)35(26)45)27(39(50)51)16-20(32(33)31(21)22)8-2-1-3-11-23-25(38(48)49)14-7-15-29(23)42/h6-7,9-10,14-19,42-47H,1-5,8,11-13H2,(H,48,49)(H,50,51) (Structure name not available in ChemSketch, replaced by InChI)	1/C28H18O10/c29-9-17-15-5-11(6-20(33)23(15)24(28(37)38)14-2-1-3-19(32)22(14)17)4-12-7-13(31)8-16-21(12)27(36)26(35)18(10-30)25(16)34/h1-3,5-10,31-36H,4H2,(H,37,38) (Structure name not available in ChemSketch, replaced by InChI)

E.2. Building diagram #2

Table 17 - Identification of the molecules reconstructed by BD#2 for sample A and B.

Molecule	Name	
	Sample A	**Sample B**
a)	Methanol	Methanol
b)	Acetaldehyde	Formic acid
c)	Hydroxylacetaldehyde	Ethanedial
d)	Ethane-1,2-diol	Methanediol
e)	3-Hydroxypropanal	3-Hydroxypropanal
f)	3-Hydroxypropanoic acid	Propanoic acid
g)	4-Hydroxybutanoic acid	4-Oxobutanoic acid
h)	Benzene-1,2,4-triol	3,5-dihydroxyfuran-2-carbaldehyde
i)	(2R,4R,5R,6S)-5,6-dihydroxytetrahydro-2H-pyran-2,4-dicarbaldehyde	4-hydroxy-6-(2-hydroxyethyl)-5-methylbenzene-1,3-dicarboxylic acid
j)	2',3,3',4,4',6-hexahydroxybiphenyl-2-carboxylic acid	1-benzofuran-3,6,7-triol
k)	2,3,5,6-tetrahydroxy-1-benzofuran-4-carbaldehyde	
l)	1/C35H30O13/c36-22-10-4-8-17(24(22)34(45)46)11-12-19-26(35(47)48)28(38)18-9-3-7-16(23(18)27(19)37)6-2-1-5-15-13-20-25(21(14-15)33(43)44)30(40)32(42)31(41)29(20)39/h3-4,7-10,13-14,36-42H,1-2,5-6,11-12H2,(H,43,44)(H,45,46)(H,47,48) (Structure name not available in ChemSketch, replaced by InChI)	1/C44H46O19/c1-15-18(11-19-26-27(39(56)40(57)31(37(26)54)42(58)59)30-33(50)17(8-6-10-46)44(61)63-41(30)35(19)52)34(51)29-25(32(15)49)16(7-4-2-3-5-9-45)23-20(12-47)22-14-62-43(60)21(13-48)24(22)36(53)28(23)38(29)55/h12-13,17,21,33,43-46,49-57,60-61H,2-11,14H2,1H3,(H,58,59)/t17-,21+,33+,43-,44+/m1/s1 (Structure name not available in ChemSketch, replaced by InChI)

Appendix F – Stochastic Simulation of bio-oil

The results simulations of the hydrotreating tests of bio-oil which are given by the graphics with the time of the experimental test, as for instance the oxygen removal, will not be here presented, as the time and temperature profiles are too large to be summarized in the present work in the form of a table.

Table 18 - Results of the stochastic simulation of the hydrotreating tests of bio-oil.

Temperature (°C)		**200**	**250**				**300**
Reaction time (h)		**1**	**0**	**0.5**	**1**	**3**	**1**
13C NMR (%w/w)	**Saturated carbons**	35.8	35.7	35.7	36.7	36.7	39.6
	Aromatic carbons	29.7	29.8	29.8	28.8	28.9	26.0
	Aldehydes and Ketones	0.0	0.0	0.0	0.0	0.0	0.0
	Carboxylic acids	6.6	6.6	6.6	6.5	6.5	6.5
	Hydroxyls and Methoxyls	27.9	27.9	27.9	27.9	27.9	27.9
Elemental analysis (%w/w)	**Carbon**	52.6	52.6	52.6	52.5	52.5	52.5
	Hydrogen	7.3	7.3	7.3	7.4	7.4	7.6
	Oxygen	40.1	40.1	40.1	40.1	40.1	39.9
	Oxygen initial experimental	39.2	39.2	39.2	39.2	39.2	39.2
	Oxygen final experimental	35.2	38.6	34.8	34.0	32.2	33.3
Removal of oxygen (%w/w) experimental		10.3	1.6	11.2	13.4	17.8	15.0
Removal of oxygen (%w/w) simulation		0.7	0.7	0.7	0.8	0.8	1.3

Table 19 - Distributions of the resultant reactions of the stochastic simulation per experimental test.

	1h - 200°C	**0h - 250°C**	**0.5h - 250°C**	**1h - 250°C**	**3h - 250°C**	**1h - 300°C**
Temperature (°C)	**200**	**250**				**300**
Reaction time (h)	**1**	**0**	**0.5**	**1**	**3**	**1**
Reactions	**Fractions (%mol/mol)**					
Saturation↔Aromatization	0.2	0.2	0.2	0.2	0.2	0.1
Hydrodeoxygenation	0.0	0.0	0.0	0.0	0.0	0.0
Demethylation	0.0	0.0	0.0	0.0	0.0	0.0
Demethoxylation	0.0	0.0	0.0	0.0	0.0	0.0
Carbonyl↔Hydroxyl	99.4	99.2	99.3	99.7	99.7	99.9
Transalkylation	0.0	0.0	0.0	0.0	0.0	0.0
Rings Opening	0.2	0.3	0.3	0.1	0.1	0.0
Decarbonylation	0.1	0.3	0.2	0.0	0.0	0.0
Decarboxylation	0.0	0.0	0.0	0.0	0.0	0.0

Table 20 - PIONA families resultant of the stochastic simulation per experimental test.

		1h - 200°C	0h - 250°C	0.5h - 250°C	1h - 250°C	3h - 250°C	1h - 300°C
Temperature (°C)	**Reconstructed Bio-oil**	**200**	**250**				**300**
Reaction time (h)		**1**	**0**	**0.5**	**1**	**3**	**1**
PIONA	**Fractions (%w/w)**						
Paraffins	50.8	50.8	50.8	50.8	50.8	50.8	50.7
Isoparaffins	0.0	0.0	0.0	0.0	0.1	0.1	0.1
Olefins	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Naphthenes	0.8	0.9	0.9	0.9	1.5	1.5	3.7
Aromatics	48.5	48.2	48.3	48.3	47.6	47.6	45.5

Table 21 - Oxygenated functional groups families resultant of the stochastic simulation per experimental test.

		1h - 200°C	0h - 250°C	0.5h - 250°C	1h - 250°C	3h - 250°C	1h - 300°C
Temperature (°C)	**Reconstructed Bio-oil**	**200**	**250**				**300**
Reaction time (h)		**1**	**0**	**0.5**	**1**	**3**	**1**
Families Oxygenated	**Fractions (%w/w)**						
Hydrocarbon	0.0	0.0	0.0	0.0	0.1	0.1	0.1
Monoalcohol	7.5	8.1	8.1	8.1	8.2	8.1	8.3
Monoaldehyde	0.6	0.0	0.0	0.0	0.0	0.0	0.0
Monoacid	0.9	0.8	0.9	0.9	0.8	0.8	0.9
Di_Oxygen_FG	27.1	17.8	17.8	17.8	17.8	17.8	17.6
Dialcohol	14.8	25.9	25.9	25.9	25.8	25.8	25.7
Dialdehyde	1.7	0.0	0.0	0.0	0.0	0.0	0.0
Diacid	3.1	3.0	3.0	3.0	3.0	3.0	3.0
Poly_Oxygen_FG	39.0	31.2	31.2	31.2	31.2	31.2	31.1
Polyalcohol	5.2	12.9	12.9	12.9	12.9	12.9	13.0
Polyacid	0.2	0.2	0.2	0.2	0.2	0.2	0.2

Appendix G – Scientific Contribution

Here is presented is the articles and oral communications produced in the past three years (2015-2018) in the present work. The second article is still under processing thus it does not have a volume number.

Articles

A.R. Costa da Cruz, J. Verstraete, N. Charon, J.-F. Joly (2018), Adaptation of a Monte Carlo method to the hydrotreating of bio-oil model compounds, Computer Aided Chemical Engineering, Volume 43 (1559-1564), 2018. doi: 10.1016/B978-0-444-64235-6.50272-2.

A.R. Costa da Cruz, J.J. Verstraete, N. Charon, J.-F. Joly (2018), A Monte Carlo method for the simulating hydrotreating of bio-oil model compounds, Chemical Engineering Journal, Volume XX (XX - XX), 2018, doi: 10.1016/j.cej.2018.10.081.

Oral Communications

A.R. Costa da Cruz, J. Verstraete, N. Charon, J.-F. Joly (2018), Adaptation of a Monte Carlo method to the hydrotreating of bio-oil model compounds, ESCAPE-28, June 10-13, Graz, Austria.

A.R. Costa da Cruz, J.J. Verstraete, N. Charon, J.-F. Joly (2018), A Monte Carlo method for the hydrotreating of bio-oil model compounds, ISCRE-25, May 20-23, Florence, Italy (Presented by J. Verstraete).

A.R. Costa da Cruz, J.J. Verstraete, N. Charon, J.-F. Joly (2017), Monte Carlo simulation of the hydrotreating reactions of pyrolysis bio-oil model compounds,16 ème Congrès SFGP, July 11-13, Nancy, France.