Computational approaches and tools for modeling biomass pyrolysis

Blake R. Hough

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

Jim Pfaendtner, Chair

Daniel T. Schwartz

David Beck

Program Authorized to Offer Degree:

Chemical Engineering

University of Washington

Abstract

Computational approaches and tools for modeling biomass pyrolysis

Blake R. Hough

Chair of the Supervisory Committee:
Professor Jim Pfaendtner
Chemical Engineering

Pyrolysis of lignocellulosic biomass is a promising process capable of producing renewable fuels and chemicals that are currently derived from nonrenewable sources. However, industrial pyrolysis processes to make these products from biomass are not yet economically viable and require significant optimization before they can contribute to existing oil-based transportation and chemical systems. One means of optimization uses kinetic and transport models for predicting the products of biomass pyrolysis, which serve as the basis for designing pyrolysis reactors capable of producing the highest value products. The goal of this work is to improve upon current pyrolysis models using computational fluid dynamics, detailed kinetic schemes, and machine learning.

First, we develop a comprehensive two-dimensional particle model for wood pyrolysis that incorporates a multi-step semi-global reaction mechanism, prescribed particle shrinkage, and thermophysical properties that vary with temperature, composition, and orientation. This model is used to validate a new experimental technique that is capable of measuring the carbohydrate composition during fast pyrolysis under industrial conditions for the first time.

Motivated by the challenges and limitations we encountered while developing this particle model, we next present a detailed kinetic model capable of predicting the temporal evolution of molecules and functional groups during lignin pyrolysis. This kinetic model provides information beyond the lumped yields of common pyrolysis models without any fitting, allowing it to cover a wider range of feedstocks and reaction conditions than lumped models. We perform an exhaustive sensitivity analysis using over two million simulations to explore which reactions contribute most to variance in the model predictions.

Due to their computational expense, detailed kinetic models cannot be incorporated into comprehensive particle models that account for both kinetics and transport effects. To address this problem we demonstrate a machine learning approach to kinetic model reduction using neural networks that reduces the computational cost of our detailed kinetic model by four orders of magnitude. The trained neural networks generalize very well, predicting the outputs of the kinetic model with over 99.9% accuracy on new data.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I could not have completed this body of work without the guidance and support of many people who helped me stay focused, productive, and happy over the last five years. I often feel that my PhD experience has been better than any graduate student should reasonably expect, and attribute this to my advisors, Jim Pfaendtner and Daniel Schwartz. Together they have allowed me great intellectual and personal freedom, opportunities to travel and work on interdisciplinary projects, and constructive feedback and support when I need it. I would like to thank Dave Beck for his mentorship and the significant amount of time he has spent helping to accelerate my professional development over the last six months. I also need to thank the members of the Pfaendtner research group for their companionship and feedback over the years of group meetings, lunches, and office conversations – you brought levity to what could otherwise be a gloomy basement office.

I would like to thank all my friends who have helped keep my life balanced during this PhD by joining me so often for adventures – our time mountaineering, kayaking, remodeling my house, hiking, and taking trips recharged my soul and spending that time with all of you was important to my success as a graduate student. Much of the credit for my arrival at this milestone goes to my mom, who has always encouraged me to ask why, but let me find the answers on my own. Her unwavering love and support throughout my life is the foundation of my academic success.

Lastly, I thank my beautiful wife, Margo. Margo has been the linchpin of my PhD, supporting and providing for me emotionally, logistically, and financially – without her this would not have been possible. During stressful times her encouragement never falters, providing me the

motivation to keep working and an example I strive to emulate. Margo, thank you, I dedicate this dissertation to you.

Chapter 1

INTRODUCTION

The fast pyrolysis of woody biomass is a promising process capable of producing bio-oil to make renewable liquid transportation fuels and chemicals currently derived from petroleum.[1-4] Despite growing interest, biomass pyrolysis requires significant optimization before it can contribute to existing oil-based transportation and chemical systems. One method of optimization uses detailed kinetic and transport models that predict the products of biomass fast pyrolysis as the basis for designing reactors capable of producing the highest value products. Advances in this area continue to be hindered by a lack of detailed understanding of the underlying transport phenomena and reaction chemistry during thermal conversion of biomass.[5]

Progress toward the type of mechanistic understanding required to overcome these hurdles is limited by the complexity of the reaction environment.[6] Pyrolysis reactions occur in three distinct phases: solid virgin biopolymers, a short-lived (<100 ms) liquid intermediate, and gas-phase pyrolysis vapors. At the industrial scale, these phases exist within a multiscale system consisting of atomic-scale biopolymer/melt chemistry (10^{-10} to 10^{-8} m), cellular and particle heating and reaction (10^{-6} to 10^{-2} m), and reactor conversion (10^{-1} to 10^{1} m) (**Figure 1.1**). A complete kinetic description of pyrolysis systems will utilize a bottom-up approach, wherein biopolymer chemistry is integrated within reaction-transport particle models, which are ultimately included in complex fluidized bed reactor models. Significant challenges remain in modeling at each of these length scales, particularly at the atomic and particle-levels.

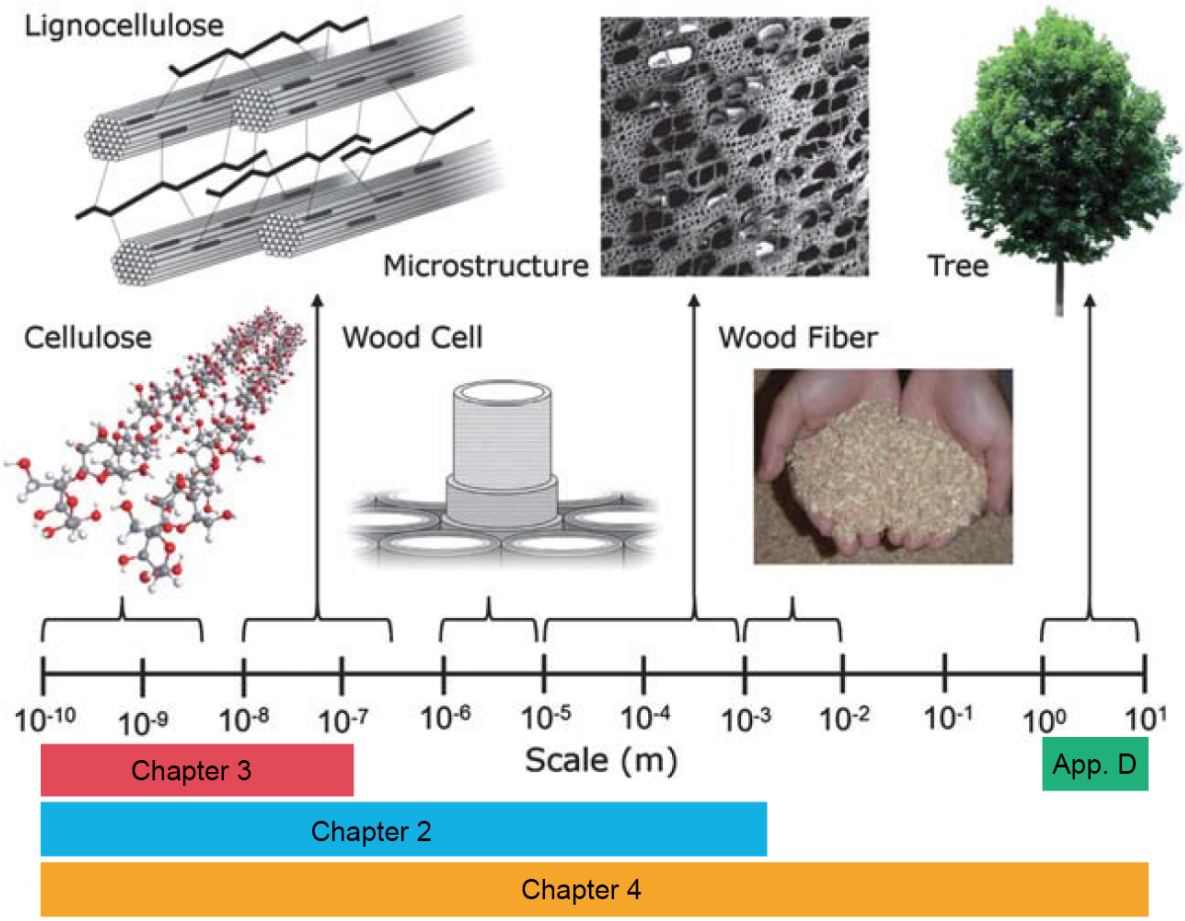


Figure 1.1: Multiscale nature of lignocellulosic feedstocks, with physical structures that span eleven orders of magnitude from the atomic level of biopolymers (10^{-10} to 10^{-8} m) to the microstructured cellular network (10^{-5} to 10^{-3} m) to the macrostructure of trees (1 to 10 m). Colored bars indicate the scale covered by the research described in each chapter of this Dissertation (the green macrostructure-scale project is not described as part of this dissertation but is included in Appendix D). Adapted from Mettler 2012.

Appreciating the challenges in modeling biomass pyrolysis requires some background knowledge of the structure and composition of biomass. A brief review is given here. For more details, please see references 7-9. Woody biomass is composed of three primary components: (1) cellulose is a long linear polymer of β-D-glucopyranose units (40-50 wt% of biomass); (2) hemicelluloses are short, branched heteropolymers made up of different sugars (15-35 wt%); (3) lignin is a highly cross-linked amorphous copolymer synthesized from the random

polymerization of three phenylpropane monomers (16-35 wt%). The structure and composition of hemicellulose and lignin differ significantly between hardwoods and softwoods, and even within a given species. Within plant cell walls long-chain cellulose molecules align to form microfibrils, which are bound together into macrofibrils by a matrix of lignin and hemicellulose. The majority of biomass pyrolysis research is done on purified cellulose because it makes up the bulk of biomass by weight and has a consistent, defined structure that can be characterized and approximated by well-defined models.

1.1 KINETIC MODELS OF BIOMASS PYROLYSIS

Many kinetic and transport models for biomass pyrolysis have been presented in the literature, and the best reviews of the field are by Di Blasi,[10] Moghtaderi,[11] Babu,[4] and White et al.[12] Most kinetic models fall into one of three categories: one-step global reaction models, one-step multi-reaction models, and multistep semiglobal models. All of these models are built from lumped product and reactant components.

One-step global models involve the conversion of a homogeneous solid virgin fuel into a sum of lumped products, either volatiles plus char, or gases, tar, and char.

$$Virgin\ Fuel \xrightarrow{k} Gas + Tar + Char \tag{1.1}$$

The rate of reaction is expressed in Arrhenius form using **Equation 1.2** and is proportional to either the residue weight or the weight loss of the virgin fuel.

$$k = A \exp\left[\frac{-E}{RT}\right] \tag{1.2}$$

The kinetic parameters are generally fit from thermogravimetric analysis (TGA) experiments. These one-step global schemes assume a fixed mass ratio between products, and are thus unable

to predict product yields at process conditions different from the experimental data set used to fit the model. They can neither account for different components of the virgin fuel nor do they have any information on the composition of volatiles.

One-step multicomponent models were developed to address two of these limitations. Some models assume virgin fuel decomposes in a series of parallel independent reactions (**Equation 1.3**). These are most often applied when dealing with a homogenous fuel such as cellulose. In cases with inhomogenous fuels, like biomass, one-step reactions as in **Equation 1.1** are written for each pseudo-component of the virgin biomass (**Equation 1.4**).

$$Virgin\ Fuel \overset{k_i}{\rightarrow} Product_i \tag{1.3}$$

$$Virgin\ Fuel\ Component\ (i) \overset{k_i}{\rightarrow} Gas + Tar + Char \tag{1.4}$$

As with one-step global models, kinetic parameters are fit to experimental TGA data and thus these models are generally not valid across a range of conditions. A key limitation of these one-step multicomponent schemes is that they ignore secondary tar cracking reactions, which are known to be important at high temperatures and long residence times.[10]

Multistep semiglobal models consider both primary and secondary reactions. In most of these schemes the fuel decomposes to an "active" intermediate, often with a lower degree of polymerization or changes in physical properties. These intermediates then decompose to other products, which can themselves take part in secondary reactions. The most common multistep semiglobal mechanism is a modification of the well-known Broido-Shafizadeh[13] mechanism (**Figure 1.2**). This mechanism has also been extended to apply to biomass, rather than only cellulose, by adding analogous parallel reaction pathways for hemicellulose and lignin.[14]

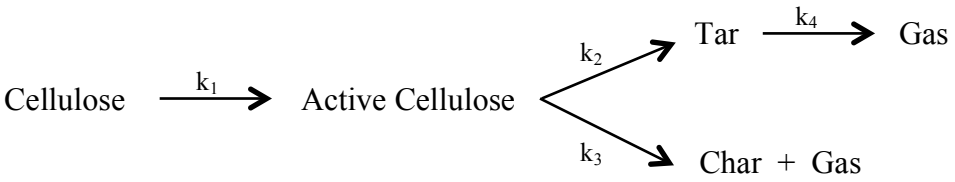


Figure 1.2: Extension of the Briodo-Shafizadeh cellulose pyrolysis mechanism to include tar cracking.

A variety of more complicated multistep mechanisms have been published[10] but are not widely used after publication, largely because multistep semiglobal mechanisms share the same limitation as the earlier models in that they are parameterized with experimental data for specific conditions and reactor configurations and are thus not broadly applicable.

The kinetic parameters for all of the models described thus far are parameterized to fit experimental data, usually from TGA experiments (for a discussion of TGA methods see section 2.2 of White et al.[12]). This dependence on TGA imposes noteworthy limitations on the models. First, since TGA measures only the decrease in sample mass as volatiles are released, it is useful only for the formulation of global or semiglobal mechanisms, and generally cannot provide mechanistic details, though extensions of TGA such as evolved gas analysis (EGA) can add information on the chemical composition of volatile products to generate more complex models. Second, to fit a kinetic model it is important that experimental conditions ensure reactions are kinetically controlled, requiring TGA to be performed at slow heating rates (1-100˚C min^{-1}) and with small sample masses. The use of slow heating rates means that kinetic parameters for these models are fit to slow pyrolysis experiments, and, since a different set of reaction pathways are thought to dominate under fast pyrolysis conditions, the global kinetic parameters for fast pyrolysis will have different values. Despite this understanding, TGA-derived parameters are

regularly applied to fast pyrolysis models due to experimental challenges in performing kinetically controlled fast pyrolysis experiments.

The many limitations of lumped kinetic models are discussed at great length in the reviews by Di Blasi[10] and White et al.[12] Their two key limitations are the fact that lumped models used today cannot predict the detailed product compositions necessary for optimizing pyrolysis reactors, nor are they accurate across ranges of reaction conditions or feedstocks. A mechanistic model built from elementary reactions would address all the lumped model limitations and allow prediction of product composition based on differing feed and process conditions. The first such mechanistic model was recently completed for pure cellulose fast pyrolysis.[15] This model was possible because cellulose has a consistent, defined structure and has been the subject of decades of intense experimental and computational work investigating the pathways, mechanisms, and kinetics of cellulose pyrolysis. Similar, purely mechanistic, models for lignin and hemicellulose are almost inconceivable at this time given their variable, and in the case of lignin, random, structures and compositions. Little effort has been put into moving toward a mechanistic model of hemicellulose pyrolysis to date, but two previously published detailed models of lignin pyrolysis incorporate mechanistic aspects. Klein and Virk published a model that combines mechanistic steps from model compound experiments with a statistical interpretation of lignin structure,[16] and Faravelli et al. published a semi-detailed model with mechanistic and lumped steps using elemental analysis to define initial pseudo-components from which to proceed mechanistically.[17] We build on the Faravelli et al. model in Chapter 3.

1.2 TRANSPORT MODELS FOR BIOMASS PYROLYSIS

Despite significant improvements in understanding molecular-level cellulose pyrolysis chemistry and refinements to kinetic model parameters, progress toward the modeling of transport phenomena within wood fibers or particles remains a challenge. Over the past few decades, it has been conclusively shown that fast pyrolysis of lignocellulosic particles produces higher yields of bio-oil than other thermochemical conversion strategies.[2] Under these conditions, initial heating of a particle conducts thermal energy through the cellular structure of lignocellulosic biomass. In turn, the particles exhibit a propagating thermal front, consistent with high Biot numbers (Bi >> 1), which drives pyrolysis chemistry in multiple zones,[18] as depicted in **Figure 1.3**. Leading this front is a drying zone where moisture evaporates,[19] followed by a pyrolysis zone where biopolymers are depolymerized to form intermediate liquids, driving microstructural collapse (shrinkage) of biomass[6, 20] and producing vapors and aerosols.[21] As pyrolysis goes to completion,

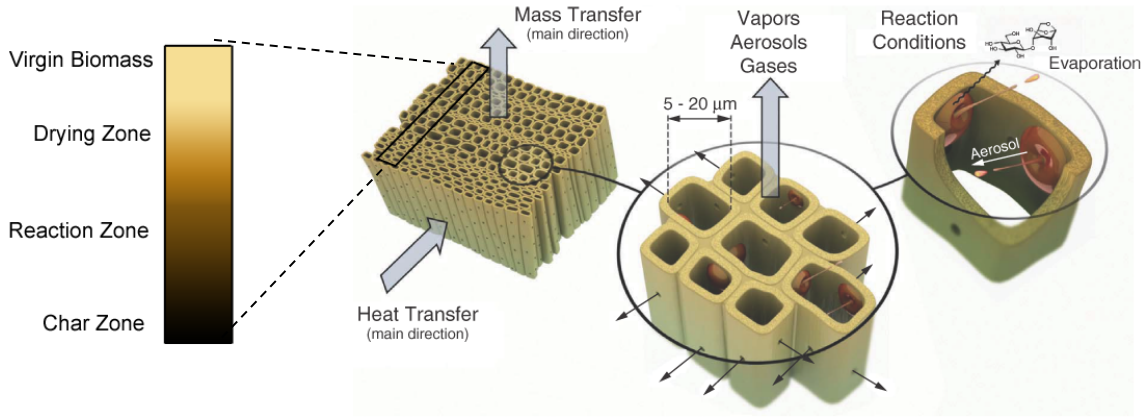


Figure 1.3: Schematic of lignocellulose chemistry and transport phenomena (shown in softwood), ranging from a single cell to a wood particle. The decomposition of particles occurs through thermal degradation of solid cell wall material to form an intermediate liquid. Subsequent decomposition of the liquid generates volatiles, which flow through pores and pits and exit the particle. A propagating thermal front results in multiple zones at the particle scale. Adapted from Kersten & Garcia-Perez 2013.

the resulting porous char zone (which will comprise the entire particle upon 100% conversion) conducts heat, transports volatiles, and traps aerosols produced earlier in pyrolysis.

Attempts to model the propagating pyrolysis reaction in wood particles have been confounded by the complexity of coupled reaction and transport phenomena models and the comparatively limited availability of reaction/transport parameters.[10] For example, reactions leading to local phase change between solid, liquid, and gas in the pyrolysis zone (**Figure 1.3**) complicate prediction of intraparticle heat transfer because model parameters, including anisotropic thermal conductivity, heat capacity, and latent heats associated with phase transition, will change significantly in this region (which also exhibits the largest temperature gradient).[6, 10] All of these parameters may be a strong function of composition and temperature and depend highly on the local cellular structure, which changes as the particles shrink.[22] Moreover, independent measurement of the physical properties of the different materials and phases is currently extremely difficult, resulting in a wide range of estimated model parameters in the literature.[10]

In summary, poor understanding of the fundamental transport phenomena and reaction chemistry during biomass pyrolysis imposes a number of limitations and challenges to the pyrolysis modeler. Existing kinetic models rely on lumped components that cannot provide valuable information on product composition. Their kinetic parameters are generally parameterized from TGA experiments using slow heating rates and specific reactor configurations, and lumped product distributions are often assigned as a fixed mass ratio from these experiments. For these reasons, such models are not valid outside the range of experimental conditions used to generate them. Furthermore, intraparticle heat transfer parameters are difficult to measure and are thought

to change significantly as particles react. Values for these parameters vary widely in the literature. These limitations have led to the large number of reaction-transport models of particle pyrolysis, which do not hold true outside of a narrow window of reaction conditions (initial composition, reaction temperature, particle size, reactor configuration, etc.). The objectives of the studies presented in this dissertation are designed to use computational tools to explore these limitations across multiple length scales (**Figure 1.1**) and propose a means around them.

1.3 ORGANIZATION OF THIS DISSERTATION

This dissertation is organized as follows. In Chapter 2 we develop a comprehensive 2D shrinking particle model for biomass pyrolysis. Most experimental wood particle pyrolysis data provides information on conversion time, reactant weight loss, and lumped product yields. This integral information does not provide the spatiotemporally resolved compositional data needed for validation of complex multiscale models. For the first time, we introduce a robust wood particle reaction-transport model to confirm the presence of a visible carbohydrate reaction front observed during fast pyrolysis with a new experimental reflectance technique. Correlations between the model and experiment provide new insights into particle shrinkage during fast pyrolysis that suggest new approaches to incorporating predictive shrinkage in future models.

In Chapter 3 we apply a semi-detailed kinetic model for lignin pyrolysis. Models that rely on lumped reaction kinetics are unable to predict the detailed product compositions required for process optimization. A mechanistic model would address this issue, but such a model is not feasible at this time for an amorphous random polymer like lignin. We apply a kinetic model for lignin pyrolysis incorporating 406 mechanistic and lumped steps. This model shows promise for predicting the distribution of functional groups in tar across a wide range of conditions (slow to

fast pyrolysis) for the first time. Detailed sensitivity analyses highlight the impact of input parameters on various model predictions.

Chapter 4 presents a new approach to solving detailed kinetic models for biomass pyrolysis using machine learning (artificial neural networks and decision trees). The computational expense of solving detailed kinetic models, involving hundreds to thousands of ODEs, precludes their incorporation in comprehensive reaction-transport particle models. We demonstrate the ability of machine learning algorithms to learn from modeled or experimental data and speed-up the computation time for predicting these detailed models by up to four orders of magnitude. This approach can be generalized to other fields, or expanded to bypass physics-based models entirely when it is too expensive or experimentally challenging to gather the necessary data to formulate such models.

Chapter 5 wraps up by giving overall conclusions and recommendations for future work. As a compliment to this body of work, Appendix D contains an additional project that is tangential to my biomass pyrolysis research. In this project we perform a landscape-level assessment of the forest biomass resources available in managed forestlands of the Confederated Salish and Kootenai Tribes in Montana. Tribal foresters there are interested in exploring the potential supply of residual biomass, left over from ecologically-based active forest management, for use in bioenergy projects.

Chapter 2

FAST PYROLYSIS OF WOOD FOR BIOFUELS: SPATIOTEMPORALLY RESOLVED DIFFUSE REFLECTANCE IN SITU SPECTROSCOPY OF PARTICLES (STR-DRiSP)

Acknowledgement of coauthors: Alex Paulsen and I worked together closely on this project and are co-first authors. My contributions to this work are the computational modeling aspects of the research – development of a comprehensive 2D particle model of the pyrolyzing wood particle – and associated analysis. We collaborated on the writing of the content.

2.1 ABSTRACT

Fast pyrolysis of woody biomass is a promising process capable of producing renewable transportation fuels to replace gasoline, diesel, and chemicals currently derived from nonrenewable sources. However, biomass pyrolysis is not yet economically viable and requires significant optimization before it can contribute to the existing oil-based transportation system. One method of optimization uses detailed kinetic models for predicting the products of biomass fast pyrolysis, which serve as the basis for design of pyrolysis reactors capable of producing the

highest value products. The goal of this work is to improve upon current pyrolysis models, usually derived from experiments with low heating rates and temperatures, by developing models that account for both transport and pyrolysis decomposition kinetics at high heating rates and high temperatures (>400°C). In this work we propose a new experimental technique, Spatiotemporally-Resolved Diffuse Reflectance in situ Spectroscopy of Particles (STR-DRiSP), which is capable of measuring biomass composition during fast pyrolysis with high spatial (ten micron) and temporal (one millisecond) resolution. Compositional data were compared with a comprehensive two-dimensional single particle model, which incorporates a multi-step semi-global reaction mechanism, prescribed particle shrinkage, and thermophysical properties that vary with temperature, composition, and orientation. The STR-DRiSP technique can be used to determine the transport-limited kinetic parameters of biomass decomposition for a wide variety of biomass feedstocks.

2.2 INTRODUCTION

Lignocellulosic biomass, including fast-growing trees and grasses, has been identified as major source of sustainable carbon capable of generating renewable fuels and chemicals.[23, 24] With over one billion dry tons of lignocellulosic biomass available in the United States,[25, 26] the U.S. Department of Energy has made it a goal to replace 30% of all transportation fuels with biofuels.[23] Thermochemical technologies for biomass conversion such as 'fast pyrolysis' to produce bio-oil and biomass gasification to produce syngas exhibit high throughput, the capability to produce existing commodity scale products (e.g. plastics or chemicals), and the potential to generate liquid transportation fuels.[27-29] Pyrolysis processes operate at high temperatures (400 – 800°C) to break down biopolymers (20,000 to 400,000 a.m.u.) into smaller molecules (less than 200 a.m.u.) that are condensed at room temperature to form a liquid referred

to as 'bio-oil'.[2, 30-32] Currently, bio-oil is a viable replacement for heavy heating oils,[33] and it has the potential to replace traditional fossil fuel sources of gasoline and diesel via downstream catalytic upgrading.[1, 34-37] However, widespread commercialization of fast pyrolysis technologies will first require optimization of the process to maximize bio-oil yield and quality. Advances in this area continue to be hindered by a general lack of detailed understanding of the underlying transport phenomena and reaction chemistry.[38]

Progress toward mechanistic understanding of fast pyrolysis is limited by the complexity of the reaction environment.[6] Pyrolysis reactions occur in three distinct phases: solid virgin biopolymers,[39, 40] gas-phase pyrolysis vapors, and a short-lived (<100 ms) liquid intermediate.[21, 41] At the industrial scale, these phases exist within a multi-scale system consisting of atomic-scale biopolymer/melt chemistry (10^{-10} to 10^{-9} m), particle/cellular heating and reaction (10^{-6} to 10^{-3} m), and reactor conversion (10^{-1} to 10^{1} m).[6] A complete kinetic description of pyrolysis systems will utilize a bottom-up approach, whereby biopolymer chemistry is integrated within reaction/transport particle models, which are ultimately included in complex fluid bed reactor models.

Understanding of molecular-scale chemistry of pyrolysis has rapidly progressed from lumped-kinetic models of the past few decades. While initial lumped chemistries predicted the rate of generation of gases, vapors and char,[13, 42] new experimental and computational techniques are revealing the pathways, mechanisms and kinetics of cellulose and lignocellulose pyrolysis. For example, development of the technique, 'thin-film pyrolysis' (TFP), has characterized the first set of pyrolysis products produced by primary condensed-phase reactions absent heat and mass

transport limitations.[43] TFP has also led to the discovery of a chain-length effect in cellulose pyrolysis[44] and the stability of the five-membered furan ring within the liquid intermediate.[45] Additionally, secondary condensed-phase reactions of cellulose have been examined by another experimental technique called, 'co-pyrolysis,' which has revealed the condensed-phase reactions of levoglucosan to produce pyrans, anhydrosugars, and light oxygenates.[46] In parallel, the use of *ab initio* molecular dynamics and DFT has described the high temperature behavior of cellulose,[47, 48] the mechanisms associated with glycosidic cleavage,[43, 49] and the formation of pyrolysis products including furans[39, 43, 50] and light oxygenates.[43, 50] These discoveries are rapidly leading to the development of molecular-level kinetic models of cellulose[15] and eventually lignocellulose pyrolysis.

Despite significant improvement in understanding molecular-level cellulose pyrolysis chemistry, progress towards the modeling of lignocellulose pyrolysis chemistry and transport phenomena within wood fibers during pyrolysis remains a challenge. Over the past few decades, it has been conclusively shown that heating rates of lignocellulosic particles/fibers in the range of 0.1 – 1.0 MW m^{-2} produces higher yields of bio-oil.[13, 28, 42, 51, 52] At these conditions, initial heating of a particle conducts thermal energy through the cellular structure of lignocellulosic biomass. In turn, the particles exhibit a propagating thermal front, consistent with high Biot numbers (Bi>>1) which drives pyrolysis chemistry in multiple zones[18] as depicted by the cutaway in **Figure 2.1a**. Leading this front is a drying zone, wherein moisture evaporates.[19] This is followed by the pyrolysis zone wherein the lignocellulosic biopolymers are depolymerized to form intermediate liquids, driving microstructural collapse (shrinkage) of biomass[6, 20] and producing vapors and aerosols.[21] As pyrolysis goes to completion, the resulting porous char zone (which will comprise

the entire wood particle upon 100% conversion) conducts heat, transports volatiles, and traps aerosols produced earlier in pyrolysis.

Attempts to model the propagating pyrolysis reaction in wood particles have been confounded by the complexity of coupled reaction and transport phenomena models and the comparatively limited availability of reaction/transport parameters.[10] For example, reactions leading to local phase change between solid, liquid, and gas in the pyrolysis zone (**Figure 2.1a**) complicate prediction of intra-particle heat transfer, since model parameters including thermal conductivity, heat capacity, and latent heats associated with phase transition will change significantly in this region (which also exhibits the largest temperature gradient).[6, 10] All of these parameters may be a strong function of composition/temperature and depend highly on the local cellular structure, which changes as particles shrink.[22, 53] Moreover, independent measurement of physical properties of the different materials/phases is currently extremely difficult, resulting in a wide range of estimated model parameters in the literature.[10] These limitations have led to a large number of reaction-transport models of particle pyrolysis that do not hold true outside of a narrow window of reaction conditions (initial composition, reaction temperature, particle size, etc.).

The experimental challenge of acquiring a detailed description of particle pyrolysis arises from the small size of wood fibers (1-2 mm), the fast time-scales of the reaction (1-5 seconds), and the compositional complexity of lignocellulose, which is difficult to characterize. Existing experimental data sets of wood particle pyrolysis focus primarily on conversion time, reactant weight loss, and lumped-product yields,[10] which do not provide the spatiotemporally-resolved

compositional data needed for validation of complex multi-scale models. Additionally, many experimental data sets of pyrolyzing particles, such as thermogravimetric analysis (TGA), utilize low heating rates with only moderate rates of temperature change (1 – 150°C min^{-1}) not reflective of fast pyrolysis conditions (>10^3 °C min^{-1}). In contrast, spatially-resolved temperature profiles of wood particles have been collected (requiring particles as large as two-to-five centimeters), but they are unable to achieve sub-millimeter spatial resolutions necessary for tracking pyrolysis reaction zones.[54] A recent review of pyrolysis modeling by Di Blasi states that "significant effort, in both theoretical and experimental research activities, is still required to formulate and validate truly comprehensive models."[10] For this reason, overcoming the experimental challenges required to generate a data set of composition within reacting particles has recently been identified by us as one of the major fundamental challenges of biomass pyrolysis.[6]

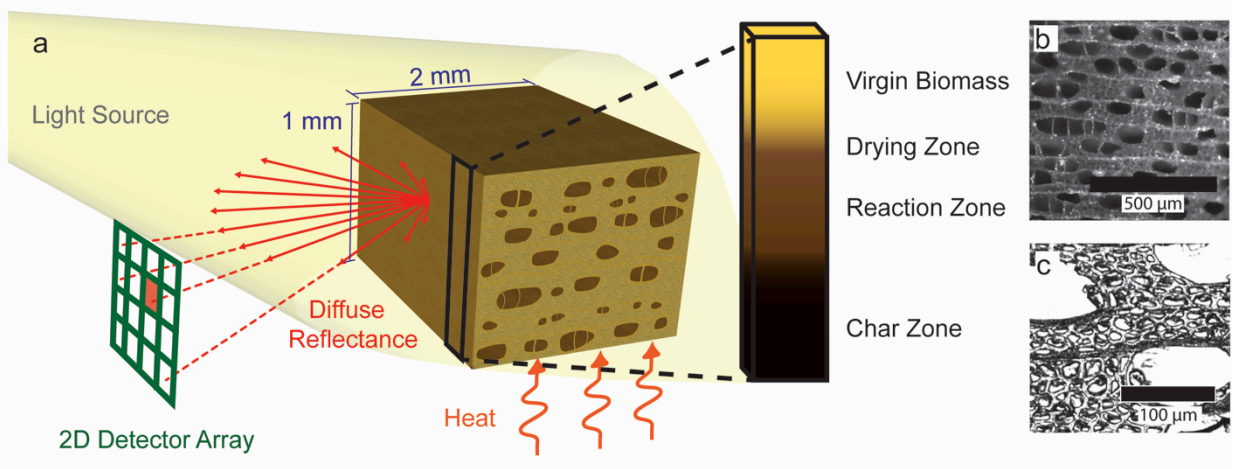


Figure 2.1: Diffuse Reflectance in situ Spectroscopy of Particles. (a) Fast pyrolysis experimental setup with a cutaway demonstrating the zones that exist during biomass pyrolysis at one moment in time, and bright field micrographs of yellow poplar structure (b) at 5x and (c) 50x magnification.

In this work, we introduce a robust wood fiber pyrolysis model based on data from a novel experimental reflectance technique capable of characterizing the carbohydrate fractions within a wood particle in both position and time at reaction conditions relevant to fast pyrolysis of biomass. The technique relies on experimental design (described in the Methods section) leading to one-dimensional heat transfer through yellow poplar (*Liriodendron tulipifera*) particles. In conventional fluidized bed reactors, wood particles are heated uniformly by gas convection such that the pyrolysis reaction front is not externally visible. However, the introduction of a heated surface for direct ablation of wood particles allows for significantly faster particle heating (relative to gas convection) which leads to one-dimensional heat transfer within the wood particle and a visible reaction front on the external particle surface as depicted in **Figure 2.1a**.

Direct observation of the pyrolysis reaction front on the external particle surface allows for compositional characterization by diffuse reflectance. Diffuse reflectance of visible and near infrared light (400 – 2500 nm) on particle surfaces has been used for characterization of woody biomass samples.[55-57] The characterization of wood chip composition, including initial lignin content and breakdown of sugars, has been demonstrated using only the 800 – 1100 nm spectrum.[58] Similarly, more general characterization of composition is available within the visible spectrum region, where lignin (highly absorptive) and carbohydrates (highly reflective) are easily distinguished.[59, 60] The distinct absorptive differences between lignin, carbohydrates (five- and six-carbon sugar-based biopolymers), and char over a broad range in the visible spectrum allows for characterization of moving or reacting wood fibers with limited spectral filtering.[56]

In the technique introduced here (**Figure 2.1**), Spatiotemporally-Resolved Diffuse Reflectance in situ Spectroscopy of Particles (STR-DRiSP), visible light (maximum intensity at 900 nm) is applied to the external side surface of a yellow poplar wood particle (1.0 mm by 2.0 mm by 4.0 mm), and diffusely reflected light is captured using a high-speed, monochrome camera. The overlap of detector absorption range and incident light allows for spectroscopic characterization in the range of 400 – 1100 nm, where significant differences in absorption between lignin/char and carbohydrates are maximized. Rapid response time of the camera (1000 Hz) combined with the capability for focusing on a two-dimensional surface (in focus on the particle surface) allows for the compositional characterization of carbohydrates within reacting particles in both position (ten micron resolution) and time (one millisecond temporal resolution).

The spatiotemporally resolved compositional data set is compared with a robust model for wood fiber pyrolysis at industrial conditions. Carbohydrate compositions within yellow poplar are measured for both position and time over a range of ablative surface temperatures (500 – 700°C). A reaction-transport model is developed by modifying the kinetic reaction model of Miller & Bellan[14] and combining it with a transport model developed for experiments described here.

2.3 METHODS

2.3.1 *WOOD PARTICLE PYROLYSIS EXPERIMENTAL DESIGN*

Experiments used yellow poplar wood from the same source and cut into blocks 1 mm x 2 mm x 4 mm (height x depth x length) prepared by Forestville Builders & Supply.[61] Wood samples consist of pores (~10 μm in diameter) that run lengthwise through the particle (**Figure 2.1c**). The micrograph at 5x was collected by boiling the wood in water for one hour before cutting blocks of the desired size (1 mm x 2 mm x 4 mm) and allowing them to dry. This process preserved the

structure of the pores at the freshly cut edge. Micrographs were then taken using an Olympus BX51 compound microscope and an Infinity 1 digital camera. Samples for the 50x image were prepared by incasing the wood sample in paraffin wax and using a microtome to cut ten micron thick slices of the wood. During ablative pyrolysis experiments, the wood particles were placed such that the pores were parallel to the heating apparatus and orthogonal to the imaging detector array. Yellow poplar was composed of cellulose (56.75%), hemicellulose (10.86%), and lignin (23.63%) with the remainder comprised primarily of water, ash, etc. (**Table 2.1**). Compositional characterization of the virgin biomass was conducted by V-Labs INC (Covington Louisiana). The density (573 kg m^{-3}) of the yellow poplar samples was determined by weighing the sample and measuring the volume using a caliper.

Table 2.1: Initial composition of the yellow poplar wood particle.

Compound	Weight Percent
Cellulose	56.75
Klason Lignin	18.38
Hemicellulose	10.86
Acid Soluble Lignin	5.25
Water	3.37
Nitrogen (as protein)	1.10
Ash	0.24
Other	4.05

Pyrolysis experiments were performed on a custom built ablative pyrolysis apparatus, which consisted of a cylindrical steel block (12.8 mm x 28.8 mm; height x diameter) that was heated by two cylindrical ceramic heating cartridges placed within the body of the steel cylinder (**Figure 2.2a**). To ensure pyrolysis (non-oxidative) conditions, a glass bell jar was suspended above the steel cylinder supplying a continuous flow of nitrogen (~13 L min^{-1}) over the steel cylinder and

wood particle. A high-speed camera (Phantom V7/Miro) with long working distance optics was placed to focus on the wood particle surface and record the evolution of the pyrolysis reaction. The source of applied light was a high intensity halogen lamp (250 W, 120 V, 3300 K tungsten filament) that was directed towards the wood particle such that prior to pyrolysis, the wood particle reflected the light, saturating the detectors (white). The orientation of the halogen lamp (light source) is shown in **Figure 2.2b** and **Figure 2.2c**. After pyrolysis, the wood sample was no longer reflective and resulted in a low sensor response (black).

Experiments were conducted by first heating the steel block to a constant reaction temperature (500, 600, 700°C) while nitrogen was directed over the steel cylinder. Once the steel block reached the desired reaction temperature (monitored using a thermocouple imbedded within the steel cylinder), a wood particle was placed on the heated surface and the reaction was allowed to proceed to completion (i.e. until no change in height or reflectance intensity was observed). This resulted in a heating rate of approximately 0.2 MW m^{-2}, as determined by balancing the heat flux at the interface between the wood particle and the heated steel surface. Additionally, total weight loss caused by pyrolysis was determined by weighing each wood particle prior to and after pyrolysis. In some cases, the wood particle moved on the surface from the convective flow of the nitrogen purge gas; these runs were discarded.

STR-DRiSP data sets were analyzed using the National Institute of Health (NIH) program ImageJ[62] by measuring the intensity of the light reflected from the face of the wood particle. The intensity of the reflected light ranged from 0 to 255 (arbitrary units) where 0 intensity

corresponds to a completely black surface and 255 to white. Data were collected at one millisecond intervals and then analyzed at 200 millisecond increments until the wood particle

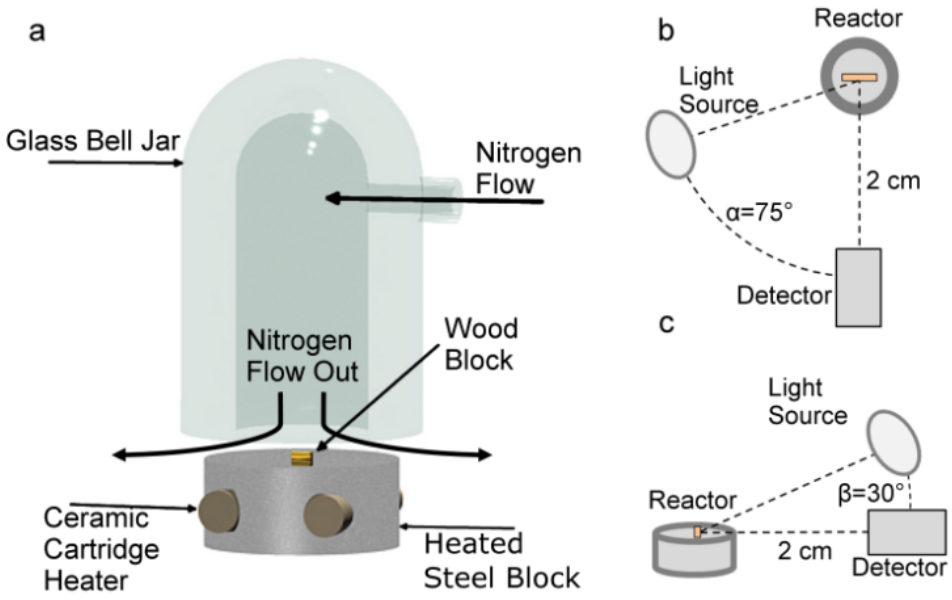


Figure 2.2: Fast pyrolysis experimental reactor design. (a) Front face view of the reactor setup. A steel block was heated using ceramic cartridge heaters, while nitrogen passed over the steel block inside a glass bell jar. Wood particles were placed on the heated steel surface at the start of the experiment, and the subsequent pyrolysis behavior was captured with spectroscopy. (b) Overhead view of the source/sample/detector arrangement. (c) Side view of the source/sample/detector arrangement.

reaction went to completion. Ten runs were performed for each considered temperature (500, 600, 700°C). The resulting data were averaged by taking a selected vertical position relative to the surface on the wood particle and averaging the corresponding spectral intensity data for each of the ten experimental runs. The resulting spectral intensity data were then converted to percentage of unreacted carbohydrate using the methods described in the next section.

2.3.2 *Diffuse Reflectance Measurements*

Experimental conditions were designed to capture the compositional transition from carbohydrate-rich biomass feedstock (cellulose and hemicellulose) to carbon rich char. The

technique employed here utilizes diffuse reflectance due to its relative ease in distinguishing between the key components in the visible and near infrared regions. Due to the Lambertian nature of wood fibers (surface roughness on the order of the wavelengths of light) it is appropriate to neglect the presence of specular reflectance and assume that all observed light is from diffuse reflectance.[63]

Figure 2.3a shows the percent of incident light that is reflected by the three dominant constituents of biomass (cellulose, hemicellulose, and lignin) and the primary solid pyrolysis product (char). Cellulose was purchased from Alfa Aesar (Part Number: A17730), xylose was purchased from Sigma Aldrich (Part Number: X1500), lignin was obtained from the Kraft process, and char was produced by pyrolyzing yellow poplar at 500°C as described previously. All diffuse reflectance measurements were conducted using a Shimadzu UV-3600 (UV-Vis-NIR spectrometer) with a Harrick Scientific Praying Mantis assembly for capturing diffusely reflected light. The baseline (100% reflectance) was measured using a spectrolon disk and the reflectance spectrum for each sample was obtained by lightly covering the spectrolon disk with the desired powder sample. All diffuse reflectance measurements were collected over the spectral range of 350 nm to 3200 nm and were conducted at room temperature in air.

Figure 2.3a presents the spectral range of the high-speed camera detector (detector specific data provided by Vison Research, an AMETEK company), demonstrating the selective detector response to excitement from incident light in the range of 9,000 - 25,000 cm^{-1} (400 – 1100 nm). The intensity of the 3300 K tungsten light source (a black body radiator) was calculated using Planck's Law and is represented in **Figure 2.3a**. From the overlap of these two curves, it is clear

that the dominant detection region is between 9,000 - 25,000 cm^{-1} where the light source emits and the detector is sensitive to light. **Figure 2.3a** shows cellulose and xylose (an approximation of hemicellulose) as highly reflective within this dominant detection range, while lignin and char are highly absorptive, indicating the ability to use diffuse reflectance to differentiate the two classes of compounds.

This contrast between the carbohydrates and the char/lignin species is made even starker when examining the predicted spectral response of the camera to the different species. When light reflected from the wood particle reaches the detector, an electrical response is triggered and the relative response is sensitive to the wavelength of light, which is demonstrated by the spectral response curve (**Figure 2.3a**). The predicted detector response, R^{λ}, is the product of the spectral response of the detector, $R^{\lambda}_{spectral}$, (**Figure 2.3b**) and the incident power observed by the detector at a particular wavelength, P^{λ}_{in}, to give, $R^{\lambda} = R^{\lambda}_{spectral} * P^{\lambda}_{in}$. If it is assumed that the incoming power is proportional to the total incident light and the diffuse reflectance of a particular sample, r^{λ}, then the electrical response of the detector can be described by Equation 1 where λ is the wavelength of light, h is Planck's constant, c is the speed of light, and k_B is Boltzmann's constant,

$$R^{\lambda} \propto \frac{R^{\lambda}_{spectral} r^{\lambda}}{\lambda^5 \exp\left(\frac{hc}{\lambda k_B T_{3300}}\right) - 1} \qquad \textbf{(2.1)}$$

The detector response to cellulose, xylose, lignin and char are predicted using Equation 1 and shown graphically in **Figure 2.3b**. This data again shows that the predicted detector response by

the STR-DRiSP technique due to carbohydrates is overwhelmingly dominant relative to the response of the highly absorptive lignin and char in the visible and near-IR regions.

The diffusely reflected light from STR-DRiSP can be directly converted into a carbohydrate composition by assuming that the observed reflected light is a linear combination of the individual components where r_i^{λ} is the diffuse reflectance and X_i is the mass fraction of species i,

$$r^{\lambda} = r_{cell}^{\lambda}X_{cell} + r_{hemi}^{\lambda}X_{hemi} + r_{lignin}^{\lambda}X_{lignin} + r_{char}^{\lambda}X_{char} \quad \textbf{(2.2)}$$

This assumption is verified by calculating the predicted detector response of yellow poplar (**Equation 2.1**) from diffuse reflectance spectroscopic data obtained for ball milled yellow poplar ($r^{\lambda} = r_{wood}^{\lambda}$). This is directly compared against the linear approximation for the experimental system here by substituting **Equation 2.2** into the detector response relationship (**Equation 2.1**), to obtain the theoretical predicted detector response for the sum of the components. Using the measured diffuse reflectance of cellulose, xylose, and lignin (**Figure 2.3a**) with the known mass fractions of each of the three components in the yellow poplar sample (**Table 2.1**), the actual and theoretical detector response for yellow poplar is shown in **Figure 2.3c** as a parity plot; this comparison demonstrates good agreement by accurately predicting the reflectance of a virgin wood sample. The linear combination of the component reflections are shown to approximate the predicted detector response for the yellow poplar, validating the ability of the linear combination model to accurately describe the experimental system.

Substituting **Equation 2.2** into **Equation 2.1** and integrating over all wavelengths for each pure component species data set ($X_{j=i} = 1, X_{j\neq i} = 0$) allows for the determination of a response factor

(α_i) for each individual component *i* (i.e. cellulose, xylose, lignin, and char) to be calculated for the STR-DRiSP system,

$$\alpha_i = \int_0^{\infty} R^{\lambda}\, d\lambda \qquad \textbf{(2.3)}$$

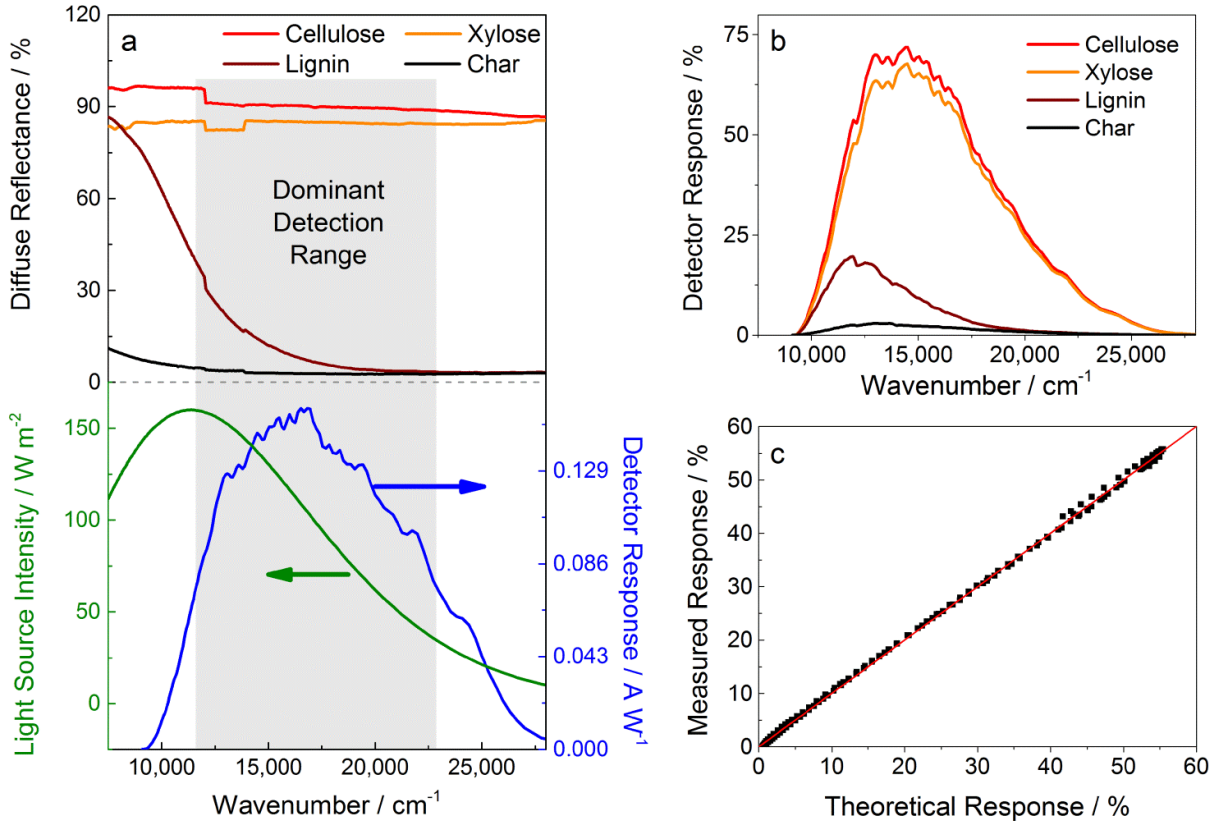


Figure 2.3: Diffuse reflectance spectroscopy. (a). Diffuse reflectivity of the components of wood shows a distinct difference in reflectivity within the range of the light source intensity, and the detector response. (b) Electrical response of the camera detector to different wood components and pyrolysis products shows a strong response to cellulose and xylose. (c) Parity plot demonstrating the effectiveness of modeling the reflectance of wood using a simple linear combination model of the components of wood.

Equation 2.3 can then be reduced to the total observed response at each pixel, R_t, (corresponding to the grayscale experimental pyrolysis data) by,

$$R_t = \alpha_{cell}X_{cell} + \alpha_{hemi}X_{hemi} + \alpha_{lignin}X_{lignin} + \alpha_{char}X_{char} \quad \textbf{(2.4)}$$

Utilizing this analysis, we are able to convert one-dimensional spatially resolved experimental data to carbohydrate mass fraction by using the following equation where I_{carb} is the Intensity measured at a specific position and time and I_{max} and I_{min} are the maximum and minimum observed intensities.

$$X_{carb} = \frac{I_{carb} - I_{min}}{I_{max} - I_{min}} \quad \textbf{(2.5)}$$

2.3.3 *COMPUTATIONAL MODEL*

The pyrolysis experiments presented here were simulated using the computational fluid dynamics (CFD) software COMSOL Multiphysics™, version 4.2. The major simplifying assumptions made in modeling the system are: (a) the wood particle is treated as a non-porous solid;[11] (b) wood is in good thermal contact with the surrounding gas;[11] (c) there is instantaneous outflow of volatile and tar products;[11] (d) crack formation in the wood particle is not considered;[11] (e) the thermophysical properties of the wood particle vary linearly between virgin wood and char as the virgin wood is consumed and char is formed;[11] (f) reaction intermediates have the same properties as virgin wood;[14] (g) secondary tar cracking reactions are neglected due to high nitrogen flow rate.[14]

2.3.3.1 Transport Model

The experiment was modeled as a two-dimensional system consisting of three distinct domains: (1) a glass bell jar, (2) a nitrogen atmosphere, and (3) a shrinking wood particle with anisotropic material properties (**Supporting Figure A.1**). The system geometry, dimensions, and nitrogen flow rate correspond to measurements of the experimental apparatus. Thermophysical properties

for the glass bell jar and nitrogen atmosphere were assigned from COMSOL's built-in material database as silica glass and nitrogen, respectively, and the properties of virgin wood and char are defined in **Table 2.2**. For consistency, the thermophysical properties for wood and char were gathered from Mehrabian et al.,[64] except for the density of yellow poplar (measured in this experiment), and the heat capacity of wood.[65] An expression different from that of Mehrabian et al. was assigned for the heat capacity of wood (**Table 2.2**), because Mehrabian et al. neglect water content; and their value is based on a specific type of wood (spruce) with dissimilar properties to yellow poplar.

The transport model is defined by the following governing equations for energy, continuity, and motion (motion applies only in the nitrogen domain). The energy conservation equation is given as:

$$\rho C_p \frac{\partial T}{\partial t} + \rho C_p \bar{v} \cdot \nabla T = \nabla \cdot (k \nabla T) + Q \qquad \textbf{(2.6)}$$

The first term represents the accumulation of energy; the second term is convective heat transfer (applies only in the nitrogen domain); the third term represents conductive heat transfer; and Q is a heat source term accounting for the heats of reaction and the heat of vaporization of water. For the glass and nitrogen domains Q = 0 as there is no reaction within these domains. Within the wood domain, Q is defined as:

$$\begin{aligned} Q = \frac{\rho}{m_{block}} \big[& \Delta H_2 \left(-m_{Acell} k_2^C - m_{Ahemi} k_2^H - m_{Alig} k_2^L \right) \\ & + \Delta H_3 \left(-m_{Acell} k_3^C - m_{Ahemi} k_3^H - m_{Alig} k_3^L \right) \\ & + \Delta H_{H_2O} \left(-m_{H_2O} k_{H_2O} \right) \big] \end{aligned} \qquad \textbf{(2.7)}$$

where the meaning of each kinetic and mass fraction parameter is described in Section 2.3.3.3. The equation of continuity is:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0 \quad \textbf{(2.8)}$$

The second term, representing the rate mass is added by convection, only applies to the nitrogen domain. The equation of motion, which also only applies in the nitrogen domain, is:

$$\rho \frac{\partial \bar{v}}{\partial t} + \rho(\bar{v} \cdot \nabla)\,\bar{v} = \nabla \cdot \left[-p\bar{I} + \mu(\nabla \bar{v} + (\nabla \bar{v})^{T}) - \frac{2}{3}\mu(\nabla \cdot \bar{v})\bar{I}\right] \quad \textbf{(2.9)}$$

The first term represents the rate of increase in momentum; the second term is the rate momentum is added by convection; and the right hand side of the equation accounts for the rate momentum is added by molecular transport.

2.3.3.2 Boundary Conditions

The boundary conditions for Equations 2.6 through 2.9 are:

No slip on all solid surfaces:

$$\bar{v} = 0 \quad \textbf{(2.10)}$$

Constant velocity at N_2 inlet:

$$\nabla_t \cdot \bar{v} = 0 \quad \textbf{(2.11)}$$

Laminar flow at N_2 inlet:

$$L_{entr} \nabla \cdot [-p_{entr}\bar{I} + \mu(\nabla \bar{v} + (\nabla \bar{v})^{T}] = p_{entr}\bar{n} \quad \textbf{(2.12)}$$

Outlet condition:

$$p = p_0 \quad \textbf{(2.13)}$$

Thermal insulation of gas inlet:

$$-\bar{n} \cdot (-k\nabla T) = 0 \quad \textbf{(2.14)}$$

Convective cooling of the bell jar in air:

$$-\bar{n} \cdot (-k\nabla T) = h_{air}(T_{ext} - T) \quad \textbf{(2.15)}$$

To account for imperfect thermal contact at the interface between wood and the heated steel block the following heat flux boundary condition was applied along that interface:

$$-\bar{n} \cdot \left(-\bar{k}\nabla T\right) = h_{block}(T_{block} - T) \quad \textbf{(2.16)}$$

Di Blasi's review of pyrolysis modeling literature[10] provided a range of values for heat transfer coefficients used in modeling fast pyrolysis systems. To choose the heat transfer coefficient for our model, we examined the results of simulations with a steel block temperature of 600˚C and heat transfer coefficients within the range reported by Di Blasi (between 80 and 1000 W m^{-2} K^{-1}). A heat transfer coefficient of 400 W m^{-2} K^{-1} resulted in the best fit to experimental data and was used in all further simulations. Along all of the surfaces of the heated steel block exposed to nitrogen gas the boundary condition is:

$$T = T_{block} \quad \textbf{(2.17)}$$

The bell jar rests on two steel supports, which are in contact with the heated steel block, and they heat the bottom surface of the bell jar. The temperature at this surface was estimated to be 90% of the steel block temperature due to convective heat losses along the length of the supports.

$$T_{\text{bell jar bottom}} = 0.9T_{block} \quad \textbf{(2.18)}$$

These initial conditions apply to all domains in the system:

$$p_0 = 1 \text{ atm} \quad \textbf{(2.19)}$$

$$T_0 = 293.15 \text{ K} \quad \textbf{(2.20)}$$

The experimental technique used in this study allowed us to observe the shrinkage of the wood particle with time at each steel block temperature (**Figure 2.4a**). While many pyrolysis models in the literature ignore particle shrinkage, we chose to prescribe the experimentally observed shrinkage into the model geometry. This was accomplished by defining a moving mesh, which adjusts the z-position of the top boundary of the wood domain with time according to an

equation derived by fitting the experimentally observed shrinkage for each different steel block temperature.

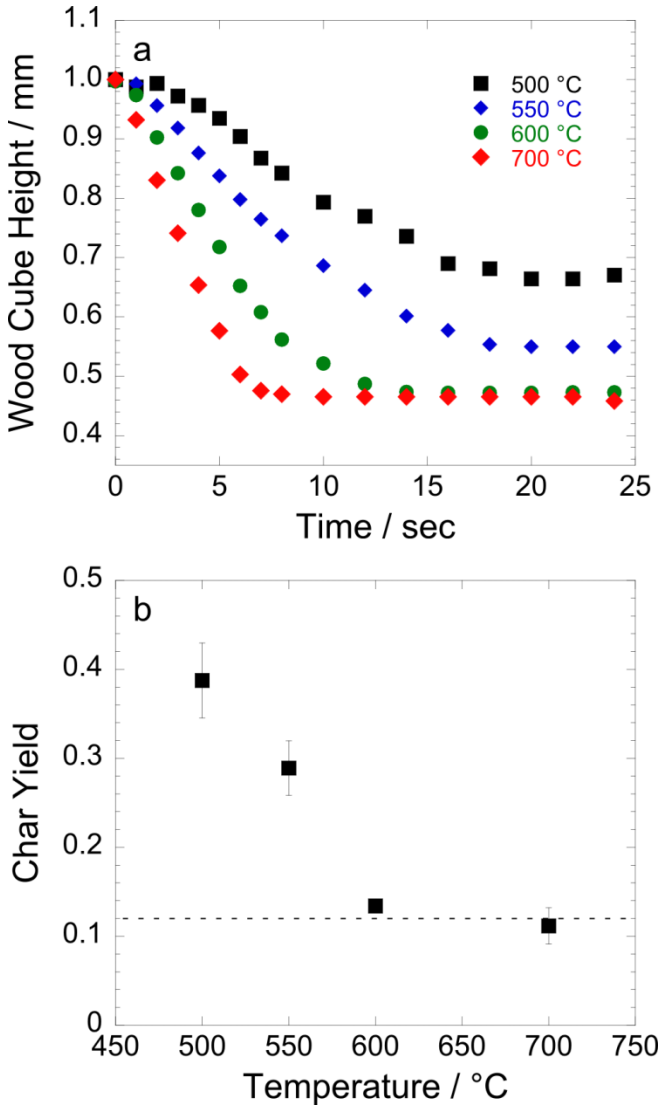


Figure 2.4: Particle shrinkage and mass loss. (a) The average yellow poplar particle height for varying reaction temperatures, and (b) the average fraction of the initial wood particle mass which remains as char and ash after completion of the pyrolysis reaction; this value reaches a plateau at approximately 0.12 (dashed line).

2.3.3.1 Pyrolysis Kinetics Model

In choosing a set of published kinetic parameters for the model, we required that the selected reaction kinetics allow calculation of the mass fractions of each solid pseudo-component (cellulose, hemicellulose, and lignin) and char. Modeling the fractions of individual solid components allowed us to tie the model directly to experimental observations of the changing composition of solid carbohydrate. Unfortunately, most kinetic parameters reported in the literature for biomass pyrolysis are derived from TG or DTG curves, based on measurements of total sample mass, and therefore kinetic parameters can only be reported for the volatile fraction of each pseudo-component. These "devolatilization mechanisms" do not provide an accurate means of tracking the behavior of solid pseudo-components, although relations are usually provided for calculating the final yield of char.

Miller & Bellan published the most complete reaction scheme that allows calculation of the mass fractions of each solid pseudo-component.[14] Their scheme applies the skeleton of the well-known Broido-Shafizadeh model for cellulose pyrolysis to all three pseudo-components (cellulose, hemicellulose, and lignin). In this scheme, virgin pseudo-components react to form "active intermediates" (Reaction 1), which decompose into tar (Reaction 2) or char plus non-condensable gases (Reaction 3). Two additional kinetic schemes[66, 67] were modeled but failed to reproduce the decomposition rates and times observed in our experimental results due to their lower rates of cellulose decomposition, so all simulation results presented are based on the scheme of Miller & Bellan. Reaction rate constants are determined using the Arrhenius equation with the values in **Table 2.2**.

$$k_i = A_i \exp\left[-\frac{E_i}{RT}\right] \qquad \textbf{(2.21)}$$

The consumption of pseudo-components and formation of "active intermediates" is described by the following equations:

$$\frac{\partial m_i}{\partial t} = -k_1^s m_i \quad \textbf{(2.22)}$$

$$\frac{\partial m_j}{\partial t} = k_1^s m_i - k_2^s m_j - k_3^s m_j \quad \textbf{(2.23)}$$

where i = (cellulose, hemicellulose, lignin),
j = ($A_{cellulose}$, $A_{hemicellulose}$, A_{lignin}) in which A means "active"
s = (cellulose, hemicellulose, lignin) corresponding to the rate constants in **Table 2.2**

All extractives are included in the mass of hemicellulose.[14, 68] Mass conservation equations for the three lumped products are:

$$\frac{\partial m_{char}}{\partial t} = 0.35k_3^C m_{Acell} + 0.6k_3^H m_{Ahemi} + 0.75k_3^L m_{Alig} \quad \textbf{(2.24)}$$

$$\frac{\partial m_{gas}}{\partial t} = 0.65k_3^C m_{Acell} + 0.4k_3^H m_{Ahemi} + 0.25k_3^L m_{Alig} \quad \textbf{(2.25)}$$

$$\frac{\partial m_{tar}}{\partial t} = k_2^C m_{Acell} + k_2^H m_{Ahemi} + k_2^L m_{Alig} \quad \textbf{(2.26)}$$

The importance of accounting for the evaporation of water during particle drying in comprehensive pyrolysis models has been demonstrated repeatedly,[10, 11, 64, 68] therefore we added water as another reactant in our reaction scheme following the treatment of Bryden, Ragland, & Rutland.[69]

$$\frac{\partial m_{H_2O}}{\partial t} = k_{H_2O} m_{H_2O} \quad \textbf{(2.27)}$$

Table 2.2: Physical parameters used in the CFD simulation

Density			
Yellow poplar	573	[kg m^{-3}]	Measured
Char	200	[kg m^{-3}]	[64]
Thermal conductivity			
Wood (z-axis)	$(0.129 - 0.049X_{H_2O})(1 + T[℃](0.986 + 2.695X_{H_2O})(2.05 + 4X_{H_2O})\times10^{-3})$	[W m^{-1} K^{-1}]	[64]
Wood (y-axis)	$2.5k_z$	[W m^{-1} K^{-1}]	[64]
Char	0.071	[W m^{-1} K^{-1}]	[64]
Heat capacity			
Wood	$1000\left[(0.1031 + 0.003867T + 4.18X_{H_2O})(1 + X_{H_2O})^{-1} + X_{H_2O}(6.191 + 0.0236T - 0.0133X_{H_2O})\right]$	[J kg^{-1} K^{-1}]	[65]
Char	$420 + 2.09T + 6.85\times10^{-4}T^2$	[J kg^{-1} K^{-1}]	[64]
Heat transfer coefficient (wood-steel interface)			
h	400	[W m^{-2} K^{-1}]	Fit
Heat of reaction			
ΔH_1	0	[kJ kg^{-1}]	[14]
ΔH_2	255	[kJ kg^{-1}]	[14]
ΔH_3	-20	[kJ kg^{-1}]	[14]
ΔH_{H2O}	2440	[kJ kg^{-1}]	[69]

Reaction rate constants	A [s^{-1}]	E [kJ mol^{-1}]	
k_1^C	2.8×10^{19}	242.4	[14]
k_1^H	2.1×10^{16}	186.7	[14]
k_1^L	9.6×10^{8}	107.6	[14]
k_2^C	3.28×10^{14}	196.5	[14]
k_2^H	8.75×10^{15}	202.4	[14]
k_2^L	1.5×10^{9}	143.8	[14]
k_3^C	1.3×10^{10}	150.5	[14]
k_3^H	2.6×10^{11}	145.7	[14]
k_3^L	7.7×10^{6}	111.4	[14]
k_{H_2O}	5.13×10^{10}	88	[69]

2.3.3.2 Solution Strategy

COMSOL uses the finite element method to solve the system of model equations based on a user-defined mesh. The final mesh was determined by shrinking the mesh element size in each domain until the solution at specific coordinates in the wood particle did not change significantly with further mesh refinement. This optimization was done for a steel block temperature of 700°C to account for the simulation with the steepest temperature and concentration gradients. Details regarding the settings for tolerances, mesh, and the solver used in these simulations are included in Appendix A, along with **Supporting Figure A.2** showing the final mesh and model setup. Complete simulations, covering 24 seconds of observed experimental time, took approximately six days each to run on a Linux workstation with two four core 2.27 GHz Intel® Xeon® E5607 processors and 5.7 GB of memory. Mass fractions along the two external faces of the simulated wood particle were averaged and compared with experimental data.

Additionally, a sensitivity analysis for the model was performed by varying the value of 20 different model parameters to explore which parameters had the greatest control on the results of the simulation. Rate constants were varied by one order of magnitude up (+) or down (-), and both the heats of reaction and char formation mass ratios were varied by 20%. Thermophysical parameters of wood and char span the range of published literature values (**Supporting Table A.1**). Parameter values were varied independently and used to run three seconds of the 600˚C steel block simulation, which was then compared to the "base case" at 600˚C. Sensitivity analysis simulations were limited to three seconds to avoid unnecessarily long computation times while still allowing the entire simulated wood particle to begin reacting. All parameters were varied except: (1) the heat transfer coefficient between wood and the steel block, because it was

used to fit the experimental data as described earlier; (2) the initial composition of the wood, because this was measured analytically; (3) the heat of vaporization of water, because its value is well established; and (4) the dimensions of the wood particle and experimental apparatus.

2.4 RESULTS

Figure 2.4a shows the evolution of particle shrinkage for different reaction temperatures. Higher temperatures (700°C) result in a greater amount of shrinkage (i.e. the total height decrease of the particle) (50%) than lower temperatures (500°C, 32% shrinkage). Higher temperatures also result in a faster rate of shrinkage (~70 μm s^{-1} at 700°C compared to ~20 μm s^{-1} at 500°C). Despite an increase in extent of shrinkage at higher temperatures, the time required to finish shrinking is much shorter at high temperatures. Total shrinkage of the particle corresponds to a decrease in non-volatilized mass, i.e. char yield (**Figure 2.4b**). While the 700°C reaction showed a faster shrinkage rate compared to the 600°C sample, both total shrinkage and the char yield is nearly identical for the two temperatures.

Figure 2.5 shows the development of the pyrolysis reaction front and the shrinkage of the wood particle. Initially the wood particle is bright white, corresponding to virgin biomass (i.e. carbohydrate rich). As the reaction progresses, the particle darkens starting at the base of the particle and proceeding up the particle face. After a few seconds the lower half of the particle is completely black, corresponding to the presence of char (a solid pyrolysis product) and absence of carbohydrates. The transition from white to dark on the particle face corresponds to the pyrolysis zone (i.e. where the pyrolysis reaction is in progress), which moves up the face of the wood particle. Additionally, the particle shrinks in size as the reaction progresses and continues to shrink after the entire face of the particle has become completely dark.

Figure 2.6 shows the percentage of unreacted carbohydrate as a function of time and position for three different experimental reaction temperatures (500, 600, and 700°C). **Figure 2.6** also includes computational results for comparison overlaid with the experimental data. The comparison shows good agreement between the experimental and computational results, especially at early times. However, in a few cases (typically as the reaction of carbohydrates nears completion) experimental and computational agreement suffers. As expected, higher temperatures result in faster reaction times and generally poorer agreement between experimental and computational results.

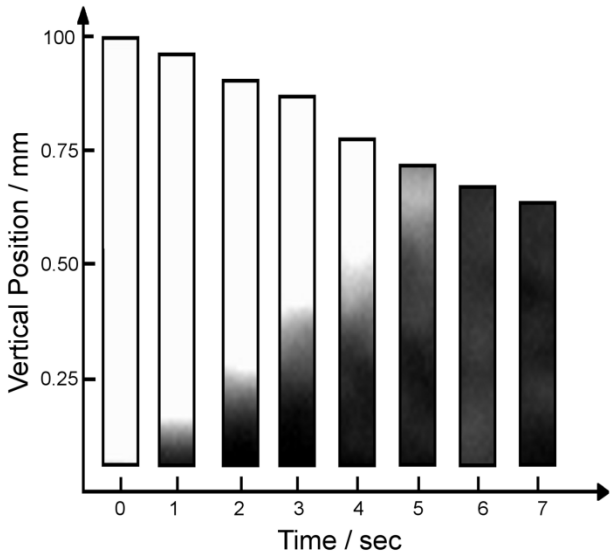


Figure 2.5: Spectral intensity of a reacting yellow poplar wood particle. Raw spectral intensity data collected for a cross section of a yellow poplar wood particle is shown during pyrolysis with a surface temperature of 600°C at different times. The unreacted wood particle with complete cellulose/hemicellulose content exhibits high spectral intensity (0 seconds, white). A visible reaction front proceeds from the bottom to the top of the particle as it shrinks. After seven seconds, the particle is completely dark consistent with low carbohydrate content.

Figure 2.7 shows the results of a sensitivity analysis undertaken to determine which model parameters, if changed, will have the greatest impact on the results of the simulation. Of the 40 simulations performed where one parameter was either increased (+) or decreased (-), only nine resulted in deviations (defined as [carbohydrate %] – [carbohydrate %]$_{base\ case}$) with absolute values greater than 10% at any z-position or time step. The remaining 31 sensitivity analysis simulations all showed non-zero deviations less than 10%. All simulations were run at a reaction temperature of 600°C and for a total of three seconds of reaction time.

2.5 DISCUSSION

Based on the observations noted above, the STR-DRiSP technique and associated computational pyrolysis particle model lead to three salient points relevant to future studies in biomass fast pyrolysis. First, we have developed a new approach for characterization of the composition of a pyrolyzing wood particle both spatially (ten micron resolution) and temporally (one millisecond resolution) using diffuse reflectance spectroscopy. Second, we have developed a coupled transport-kinetic model using the STR-DRiSP data that is capable of studying fast pyrolysis of wood at industrial conditions. Third, the observation and associated model analysis of the role of wood particle shrinkage in fast pyrolysis holds critical importance for mechanistic understanding of pyrolysis and guidance for future modeling efforts. These aspects are further discussed below.

2.5.1 *COMPARISON OF EXPERIMENT AND MODEL*

Initial reflectance of the wood particle is caused by the presence of the carbohydrates cellulose and hemicellulose. **Figure 2.3a** shows the highly reflective nature of both cellulose and xylose (a good approximation of hemicellulose), while also demonstrating the highly absorptive nature of lignin and char within the visible and near IR spectrum. The basis for this technique was this

difference in reflectivity, which can be used to decouple carbohydrate content from lignin/char and be used to monitor the concentration of carbohydrates both spatially and temporally. **Figure 2.3c** substantiates this method and shows that the reflectance spectrum of a yellow poplar wood sample can be predicted accurately by a simple linear combination of the reflectance of the three constituents (cellulose, hemicellulose, and lignin) in their pure form.

We apply the above technique to the fast pyrolysis of yellow poplar in order to track the carbohydrate content of the wood particle (**Figure 2.6**) in both position and time. **Figure 2.6** shows the results of applying this technique to the experimental results and compares it to the results of a computational model based on kinetics (gathered from TGA experiments) taken from Miller & Bellan.[14] The experimental and computational results show qualitative and quantitative agreement at short times (1-3 sec) and lower temperatures (500°C and 600°C). In all cases the basic trends are observed in the model, with minor discrepancies in the slope and shape of the model curves. However, at a few intermediate time steps, the model and experimental agreement is no longer as accurate. The pyrolysis zone is considerably broader (approximately 50%) in the experimental results when compared to the simulation results, especially at 700°C. This difference may derive from the reaction kinetics used in the computational model, which were developed from experiments with a slow rate of temperature change (5-80°C min^{-1}) or from inaccuracies in parameter estimation. This narrowing of the pyrolysis zone by simulation also occurs to a lesser extent for all of the data sets (i.e. the simulations react to completion more quickly than the experimental results). This discrepancy may be an indication that the low temperature kinetics over-predict the speed of pyrolysis reactions at high temperatures.

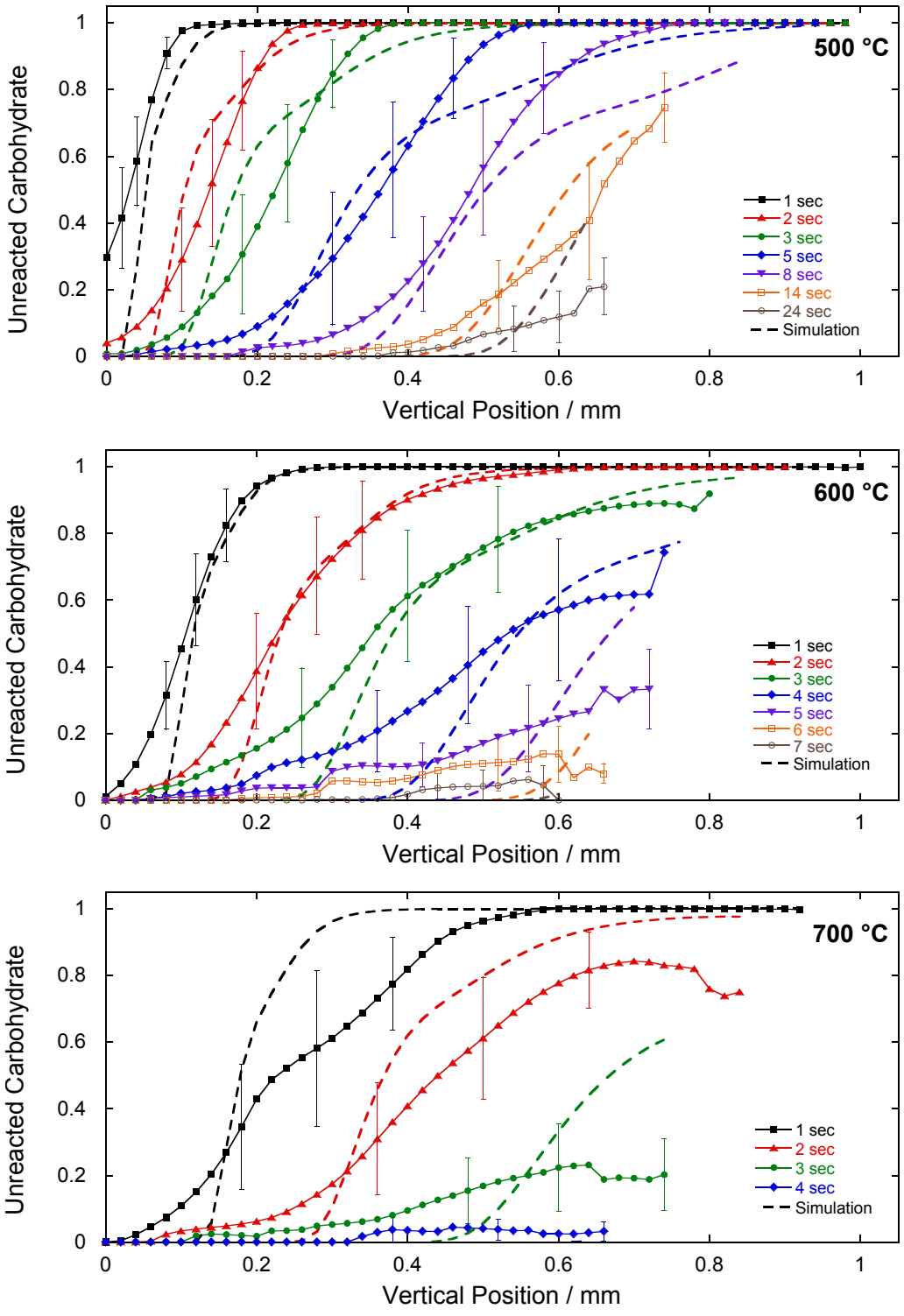


Figure 2.6: Spatiotemporally resolved carbohydrate composition profiles of pyrolyzing yellow poplar. Modeling and experimental results of the mass fraction of unreacted carbohydrate content as a function of time and position. Plots depict pyrolysis at 500°C, 600°C, and 700°C.

2.5.1 *PARTICLE SHRINKAGE*

Figure 2.5 shows the evolution of a visible reaction front during the pyrolysis of a wood particle. Wood is completely reflective prior to the start of the reaction, but upon contact with the heated surface the base of the wood particle begins to darken (i.e. is no longer reflective). As the reaction progresses, the reaction front (defined as the transition from reflective to absorptive) travels from the base of the particle to the top. In addition to the loss of reflectivity, the particle also shrinks in size over the course of the reaction. However, the particle continues to shrink after the majority of the carbohydrates in the particle have reacted (true for all reaction temperatures), which implies that particle shrinkage is not driven completely by carbohydrate pyrolysis. Davidsson et al. has shown that at low temperatures (<500°C) shrinkage does not begin until after 60% of the biomass has been converted at which point much of the cellulose has been consumed.[70] Davidsson et al. goes on to postulate that the decomposition of the rigid structures formed around the microfibrils by lignin and hemicellulose char account for the shrinkage of biomass particles during pyrolysis.[70] In our work, we find that the wood particle's transition from reflective to absorptive is caused by the consumption of cellulose and hemicellulose (carbohydrates). Since the particle continues to shrink after all carbohydrates have been consumed, our work agrees with Davidsson et al. that carbohydrate conversion is not the sole cause of particle shrinkage. Rather, the breakdown of lignin, which has been shown to react slowly over a large range of temperatures,[71] likely contributes to the shrinkage phenomenon. As heat is transferred upward through the particle, the visible reaction front (loss of reflectivity) would be accompanied by another reaction front, invisible to our STR-DRiSP technique, corresponding to the decomposition of lignin that would contribute to particle shrinkage.

Particle shrinkage has been tied to the collapse of the native structure within wood particles, which corresponds to breaking down of the solid biomass components as they react to form liquids and gases[70, 72]. Lignin is considered to be a key component in the structural integrity of biomass,[73] suggesting that lignin breakdown is an important aspect of particle shrinkage. Particle shrinkage is also related to the total non-volatilized mass/char yield of the particle (**Figure 2.4b**). As temperature increases, both total mass loss and total amount of shrinkage increase, while both shrinkage and weight loss reach a plateau at 600°C. This implies that shrinkage is directly related to the fraction of initial biomass that reacts to form solid char (less char equals more shrinkage) and that there is a minimum yield of char (~12%, asymptotic value, **Figure 2.4b**) that cannot be decreased by continuing to increase reactor temperatures.

It is noteworthy that including particle shrinkage in our model is critical to achieving reasonable agreement with experimental data. This is because the high nitrogen flow rate over the wood surface dramatically slows the heating of the upper portion of the particle (because the nitrogen contacting the wood is at a lower temperature than the ablative surface), delaying initiation of the decomposition reactions. Simulations without shrinkage exhibit a lag in predicted carbohydrate consumption compared to the experiment. At short times (1-2 seconds), before the particle has shrunk appreciably, this lag is not observed, but the lag grows with time as the difference in particle height increases between the shrinking-disabled model and the experimentally observed particle. Complete decomposition of carbohydrates within the 24-second simulation time is not observed when shrinking is disabled. The lag described above is eliminated when shrinkage is included in the model, because the driving force for heat transfer increases as the wood shrinks, increasing the speed of the thermal front moving from the bottom to the top of the particle.

In our model, particle shrinkage was achieved by defining a moving mesh that compressed all mesh elements in the wood domain with each time step. Despite the observed accuracy of this coarse approach, there are physically unrealistic aspects of our assumption arising from the fact that lignin decomposition is not uniformly distributed in position or time within the wood particle. In future work it would be interesting to systematically allow shrinking of individual mesh elements, based for example on their lignin mass fraction, so that the entire particle shrinks and warps asymmetrically as observed in experiments. While possible in principle, it is yet unknown if the increased computational cost of the more detailed simulation would be prohibitive, but it is worth considering due to the fact that there are as yet no known predictive shrinkage models for wood particle pyrolysis.

2.5.2 *Model Discussion*

Sensitivity analysis of the model of **Equations 5-26** shows that five parameters are important to know accurately, because variation in any of them changes the model results significantly. These parameters are the wood density, wood heat capacity, wood thermal conductivity, and the rate constants of the first reaction in both cellulose and hemicellulose decomposition. Our sensitivity analysis findings are in agreement with previously published work regarding the most important parameters.[10, 68, 74] In our case, all of these parameters except the wood density are general values for "wood" from literature, and are not specific to our yellow poplar. Determining species-specific certainty bounds on these five important parameters will be important for future modeling work, whereas accurate values for other model parameters are lower priority.

We made significant simplifying assumptions by neglecting the porosity of the wood sample and assuming instantaneous outflow of volatiles. Some researchers have chosen to account for the

pressure variation due to volatile flow within wood's pores using the Darcy law, especially in less computationally expensive one-dimensional models.[11] Our assumption of a non-porous solid was made in order to reduce the time and computational expense of running such a complicated model. However, including the porosity of wood will become more important in future comprehensive models when particle shrinkage is based on physical mechanisms.

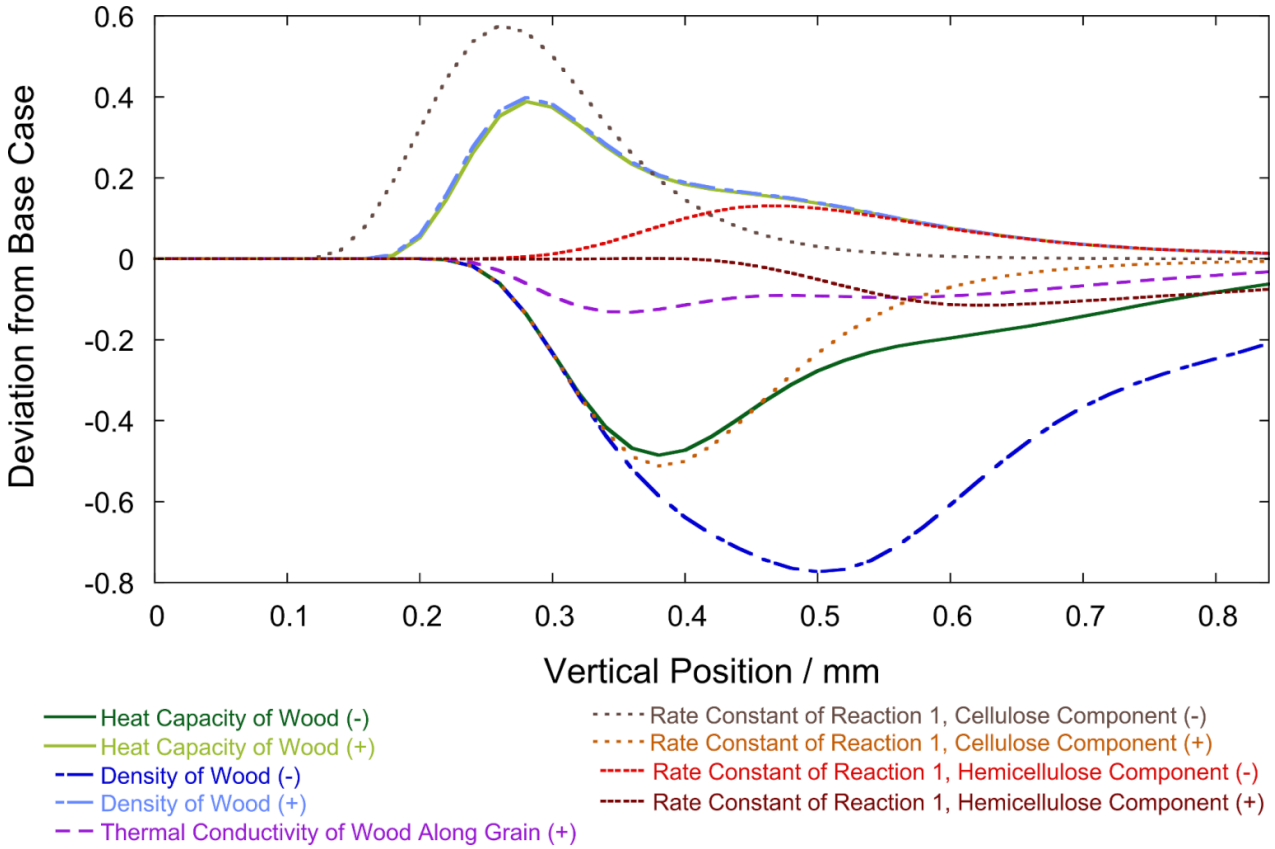


Figure 2.7: Reaction-transport model sensitivity analysis. Sensitivity of the 600˚C steel block simulation to the most important model parameters as a function of position (at three seconds). The deviation is defined as [carbohydrate mass fraction] – $[\text{carbohydrate mass fraction}]_{\text{base case}}$.

2.5.3 *FUTURE RESEARCH DIRECTIONS*

The experimental capability described here extends particle characterization beyond integral data sets for the first time and allows for validation using the fine structure of spatiotemporally-defined concentration gradients. However, the first generation of this technique is limited to monitoring the combined concentration of cellulose and hemicellulose and cannot distinguish between the two. Additionally, the concentration of lignin and char cannot be monitored. Future work is necessary including the modification of the detector setup such that it is capable of distinguishing between cellulose, hemicellulose, lignin, and char. This is possible in principle, because lignin has distinct absorption peaks at 1635 – 1825 nm and 1075 – 1250 nm while carbohydrates demonstrate absorbance peaks at 1400 – 1660 nm and 1890 – 2020 nm.[57] Also, lignin is clearly more reflective than char above 1000 nm (**Figure 2.3a**). Similarly, individual products and specific functional groups can be monitored online FTIR.[75] Future work will also be designed to better quantify and understand particle shrinkage. This will require positional tracking of specific sites on the wood surface during pyrolysis. By tracking multiple sites, one can identify the zones where shrinkage occurs and whether this zone coincides with changes in composition. This type of information will be important in identifying the cause(s) of biomass particle shrinkage (specifically how the decomposition of each species relates to shrinkage) which is needed for improvements in the reliability and accuracy of pyrolysis models.

2.6 CONCLUSIONS

We demonstrate and validate a new experimental technique (STR-DRiSP) capable of monitoring biomass composition during pyrolysis both temporally and spatially. Experimental design was provided for reaction of wood particles with one-millimeter thickness by direct ablation with a heated surface, such that heat transfer was overwhelmingly one-dimensional. For these thermal

conditions, characterization of the particle surface by spectroscopy was indicative of compositional changes throughout the particle. Spectroscopic characterization of a light source and camera detector along with the reflective properties of lignin, char, cellulose, and xylose supported the method of compositional characterization of carbohydrates with high spatial (ten micron) and temporal (one millisecond) resolution. Carbohydrate compositional profiles were generated for pyrolyzing yellow poplar wood particles at 500, 600 and 700°C, providing the first spatiotemporally resolved compositional profiles of reacting biomass particles at industrial conditions.

A comprehensive two-dimensional reaction-transport model of a single particle was developed that presents the most accurate model for reaction-transport modeling of wood particles under relevant heating conditions using carbohydrate composition data measured with STR-DRiSP. At lower temperatures the model results are in good agreement with experimental data; discrepancy at higher temperatures was due to the use of lumped kinetic reaction parameters for biomass fast pyrolysis. Given that only two parameters were fit in this pyrolysis model (the heat transfer coefficient and the shrinkage rate), it has utility in a predictive capacity for systems where the shrinkage rate is known or can be estimated. The combination of experimental STR-DRiSP and reaction-transport modeling lays the groundwork for addressing several longstanding challenges in the development of fast pyrolysis technology for biomass: (1) poor understanding of fast pyrolysis transport and reaction kinetics; (2) uncertainty about the mechanisms of particle shrinkage during fast pyrolysis; and (3) high variability in fast pyrolysis reaction-transport models.

2.7 NOMENCLATURE

α_i	Detector response of component I integrated over all wavelengths
ΔH_{H_2O}	Heat of vaporization of water
ΔH_1	Heat of Reaction 1
ΔH_2	Heat of Reaction 2
ΔH_3	Heat of Reaction 3
λ	Wavelength
μ	Dynamic viscosity of nitrogen gas
ρ_c	Density of char
ρ_w	Density of wood
A_i	Pre-exponential factor in Arrhenius expressions
c	Speed of light
$C_{p,c}$	Heat capacity of char
$C_{p,w}$	Heat capacity of wood
E_i	Activation energy
h	Planck's constant
h_{block}	Heat transfer coefficient for the wood-heated steel block
h_{air}	Heat transfer coefficient of air
i	Component; either cellulose, hemicellulose, lignin, or char
I	Identity matrix
I_λ	Intensity at wavelength λ
I_{carb}	Intensity from carbohydrates
I_{max}	Maximum experimental intensity
I_{min}	Minimum experimental intensity
k_B	Boltzmann's constant
k_c	Thermal conductivity of char
k_{H_2O}	Rate constant for vaporization of water
$k_{y,w}$	Thermal conductivity of wood with the grain
$k_{z,w}$	Thermal conductivity of wood across the grain
k_1^c	Rate constant of Reaction 1, cellulose component
k_1^H	Rate constant of Reaction 1, hemicellulose component

k_1^L	Rate constant of Reaction 1, lignin component
k_2^C	Rate constant of Reaction 2, cellulose component
k_2^H	Rate constant of Reaction 2, hemicellulose component
k_2^L	Rate constant of Reaction 2, lignin component
k_3^C	Rate constant of Reaction 3, cellulose component
k_3^H	Rate constant of Reaction 3, hemicellulose component
k_3^L	Rate constant of Reaction 3, lignin component
L_{entr}	Length of nitrogen inlet tube outside the modeled domain
m_i	Mass of component i
p	Pressure
Q	Heat source in energy equation
r^{λ}	Total percent of light reflected by wood at wavelength λ
r_i^{λ}	Total percent of light reflected by component i at wavelength λ
R	Universal gas constant
R_t	Total Detector Response
R_i^{λ}	Electrical response of the detector at wavelength λ
$R_{spectral}^{\lambda}$	Strength of detector response at wavelength λ
t	Time [sec]
T	The temperature at the coordinates being computed
T_{block}	The temperature of the heated steel block
T_{ext}	External air temperature around apparatus (20˚C)
T_{light}	Temperature of the light source
v	Velocity of nitrogen gas
X_{carb}	Carbohydrate mass fraction
X_i	Mass fraction of component i

Chapter 3

DETAILED KINETIC MODELING OF LIGNIN PYROLYSIS FOR PROCESS OPTIMIZATION

3.1 ABSTRACT

Biomass valorization through thermochemical conversion of lignocellulosic feedstocks is limited by our lack of detailed kinetic models. In addition to adding mechanistic understanding, more detailed models are needed to optimize industrial biomass pyrolysis processes for producing fuels and chemicals. To this end, a kinetic model for lignin pyrolysis involving ~100 species and 400 reactions is presented which is capable of predicting the temporal evolution of molecules and functional groups during lignin pyrolysis. The model provides information beyond the lumped yields of common pyrolysis models without any fitting, allowing it to cover a wider range of feedstocks and reaction conditions. Good agreement is observed with slow pyrolysis experiments, and an exhaustive global sensitivity analysis using over two million simulations sheds light on reactions that contribute most to the variance in model predictions. Model predictions for fast pyrolysis are available, however, recently developed experimental techniques for kinetically-controlled fast pyrolysis of biomass have yet to be applied to lignin.

3.2 INTRODUCTION

Lignin – a highly cross-linked amorphous copolymer of three randomly polymerized phenylpropane units – accounts for up to 35wt% of dry biomass, and carries the highest specific energy content of the three principal components of biomass.[76] In addition to its significant presence in biomass, lignin is a byproduct of the pulp and paper industry (more than 50 million

tons/year),[77] and of biorefinery processes that focus on utilizing the more easily convertible polysaccharide fractions of biomass (e.g. sugar fermentation to ethanol). Today, most lignin is burned for process heat,[78] however, processes like pyrolysis, gasification, combustion, and liquefaction are available to convert lignin or biomass into useful products. Of these, pyrolysis is the most promising method for converting lignin into lower molecular weight liquid and gas products.[2, 3, 79] A longstanding goal is the development of predictive models to optimize fast and slow pyrolysis processes that produce renewable fuels and aromatic chemicals from lignin; unfortunately, existing models lack sufficient detail to be useful for this type of optimization.

Empirical and global models of biomass or lignin pyrolysis using overall apparent kinetics derived from TGA studies are well established and can often predict yields of lumped products (tar, gases, char) within a small range of experimental conditions.[4, 10-12] The main limitations of these models are their inability to predict the composition and distribution of product tars and gases, or to cover a wide range of heating rates and temperatures with one set of kinetic parameters. A fully mechanistic model of lignin pyrolysis, comparable to that recently developed for glucose-based carbohydrates,[80] would be of great interest for these reasons, but developing such a model is challenging due to the random structure of lignin polymers.[17, 76, 81]

There have been two noteworthy attempts at making "less lumped" kinetic models of lignin pyrolysis in the last decade. Klein and Virk proposed a model[16] in which they assumed a structure based on Freudenberg's classic representation of spruce lignin.[82] They assigned independent probabilities that a given aromatic unit will contain different types of methoxyphenol or propanoid substituents, and determined the number of aromatic units of each type based on the spruce lignin model structure. Sixteen model compounds were selected from

these representative aromatic units and used in slow pyrolysis experiments to determine reaction products and the associated kinetic parameters. These kinetic pathways and parameters were used to simulate the constant volume batch pyrolysis of spruce lignin. Their work represents a significant improvement on earlier lumped models of lignin pyrolysis, but is limited because the analysis is based on a single representative structure of one lignin (spruce). Furthermore, without further experimental work it is not obvious how to extend this model to lignins from other feedstocks.

More recently, Faravelli et al. published a model[17] that moves toward addressing these limitations. Their model – described in more detail in the next section – bases its characterization of the starting lignin on functional group characteristics spanning a wide range of hardwood and softwood lignins (in contrast to the single representation of spruce lignin from Klein and Virk). This approach leads to a less detailed description of the initial lignin structure, but offers the capability to study a wide range of lignins. Kinetic parameters for pyrolysis reactions in the model are derived from gas phase reactions or taken from literature.

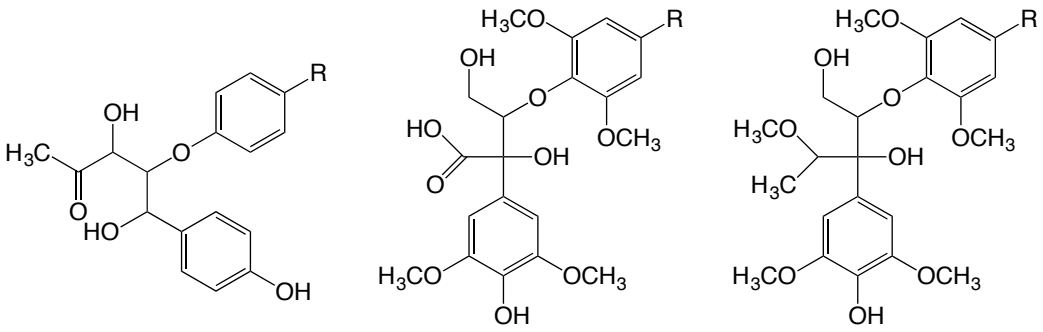
As a step toward a more mechanistic model, we explore the use of a semi-detailed kinetic model for lignin pyrolysis to predict the compositions of gases and tar components from the pyrolysis of lignin over a wide range of slow and fast pyrolysis conditions. Our kinetic scheme is based on that proposed by Faravelli et al., with key changes described in Section 3.3.2. The model combines elementary and lumped reaction steps to track approximately 100 real molecules, radicals, lumped heavy species, and functional groups linked to the degrading polymer. To better understand the limitations of the model or areas of focus for future refinement of the model, we also performed a detailed sensitivity analysis to understand the relationship between production

of permanent gases and functional groups in tar to experimental conditions. One significant obstacle in this field is the inability to utilize any of the many published kinetic models without investing significant time reproducing and troubleshooting them, so an additional goal of this work was to publish the complete model code and supporting files as a well-documented, open source Python module that users with basic knowledge of Python can begin using immediately.

3.3 DETAILED CHEMICAL KINETIC MODEL FOR LIGNIN PYROLYSIS

3.3.1 *MODEL COMPOUNDS TO APPROXIMATE LIGNIN'S COMPLEX STRUCTURE*

Challenges characterizing lignin and defining its structure are well documented,[16, 17, 76, 81] and in this work we follow the approach of Faravelli et al. in choosing model components to approximate different possible lignin structures. Complete characterization and modeling of lignin would require consideration of all the well-known inter-unit linkages and group distributions, but since little is known about the degradation mechanisms and intermediates in such a complex model some simplification and lumping is warranted. The most representative lignin substructure, making up about 60% of all the linkages in lignin, is the β-O-4 (β-aryl ether) linkage. For this reason, three model components (LIG-C, LIG-H, LIG-O, following the naming conventions of Faravelli et al.) based on a β-O-4 skeleton are chosen to characterize the initial lignin structure (**Figure 3.1**). Each model component has a different distribution of methoxyl and other functional groups, which are often used to characterize lignins (hardwoods contain more methoxyl groups than softwoods).



LIG-C
$C_{17}H_{17}O_5$

LIG-O
$C_{16}H_{11}O_6(OCH_3)_4$

LIG-H
$C_{18}H_{17}O_5(OCH_3)_4$

Figure 3.1: Model components used to characterize the initial lignin structure (modified from Faravelli et al.).

When describing lignins from various sources it would be ideal to have information about the distribution of different bonds and functional groups. Unfortunately, this level of detail is rarely available. Instead, in this work we translate commonly available elemental composition – mass fraction of carbon, hydrogen, and oxygen – measurements into an equivalent composition of the three reference units using a linear relationship based on conservation of mass:

$$\begin{bmatrix} \frac{204.187\ \text{g C}}{\text{mol LIG-C}} & \frac{240.22\ \text{g C}}{\text{mol LIG-O}} & \frac{264.242\ \text{g C}}{\text{mol LIG-H}} \\ \frac{17.1343\ \text{g H}}{\text{mol LIG-C}} & \frac{23.1817\ \text{g H}}{\text{mol LIG-O}} & \frac{29.2291\ \text{g H}}{\text{mol LIG-H}} \\ \frac{79.995\ \text{g O}}{\text{mol LIG-C}} & \frac{159.99\ \text{g O}}{\text{mol LIG-O}} & \frac{143.991\ \text{g O}}{\text{mol LIG-H}} \end{bmatrix}^{-1} \cdot \begin{bmatrix} w_C \\ w_H \\ w_O \end{bmatrix} = \begin{bmatrix} \text{mol LIG-C} \\ \text{mol LIG-O} \\ \text{mol LIG-H} \end{bmatrix} \quad \textbf{(3.5)}$$

In rare cases the experimentally determined elemental composition is outside the bounds of a triangle whose vertices are the three model components in "C-H-O-space". The result of this is that the mole fraction of one of the model components (**Figure 3.1**) is assigned a very small negative value. In such situations, we set the composition of that component to zero and normalize the values of other two components.

3.3.2 *THE KINETIC SCHEME AND MODEL EQUATIONS*

Details on the reactions and sources of kinetic parameters included in the original model are available in Faravelli et al.'s article.[17] We updated the model with a number of modifications, including the addition of eight reactions to their kinetic scheme (all of the changes are detailed in the Appendix B). For initiation reactions that are not analogous to existing reactions in the Faravelli et al. scheme we estimated the kinetic parameters as follows. First, we performed density functional theory (DFT) calculations using the M062X functional[83] with the 6-311G(2d,d,p) basis set (i.e., the so-called CBSB7 basis set[84]) in Gaussian[85] to calculate the heat of reaction of all the published initiation reactions using a standard geometry optimization and frequency calculation as in our prior work.[86] We then used the Blowers and Masel correlation to estimate the activation energy.[87] Our estimated activation energies were approximately two times larger than those for initiation reactions in the published scheme, so when we repeated this procedure with our proposed reactions we divided the result by this conversion factor. Frequency factors were chosen to match other reactions of the same class in Faravelli et al.'s scheme.

As this is a purely kinetic model, the model equations are straightforward ordinary differential equations for the concentrations of each species and the temperature in a perfectly mixed batch reactor:

$$\frac{dC_i}{dt} = R_i \tag{3.6}$$

$$\frac{dT}{dt} = \alpha \tag{3.7}$$

where C_i is the concentration of the i-th species, R_i is the net rate of formation, and α is the heating rate.

Our final kinetic scheme of 406 reactions involving 93 species is included in Appendix B (**Supporting Table B.1**). Programs were written in Python to automate all steps of the model-building process, and the resulting system of stiff ordinary differential equations was solved using DDASAC.[88] We have published all the software necessary to generate the model equations and couple them with any stiff ODE solver.[89]

3.4 RESULTS AND DISCUSSION

3.4.1 *MODEL VALIDATION*

The typical strategy for validating kinetic models of biomass slow pyrolysis is comparing predicted weight loss curves with experimental results from TGA experiments. In **Figure 3.2** we compare model predictions for total solid yields to the experiments of Jakab et al.[90] They examined 16 milled wood lignins from a variety of plants including hardwoods, softwoods, and grasses. Samples were pyrolyzed in an inert atmosphere and a heating rate of 20°C min^{-1}. Our model shows excellent temporal agreement with solid yield for lignins from the two species shown; we observed the same level of agreement with the other 14 species in the experimental study. It is worth noting that this agreement is achieved without fitting any parameters. This approach to validation is limited – only offering information on agreement with the devolatilization of lumped solid components. Technical limitations of most TGA units make the collection of temporal data on the yield of volatile components difficult, so they are routinely neglected in this type of comparison.

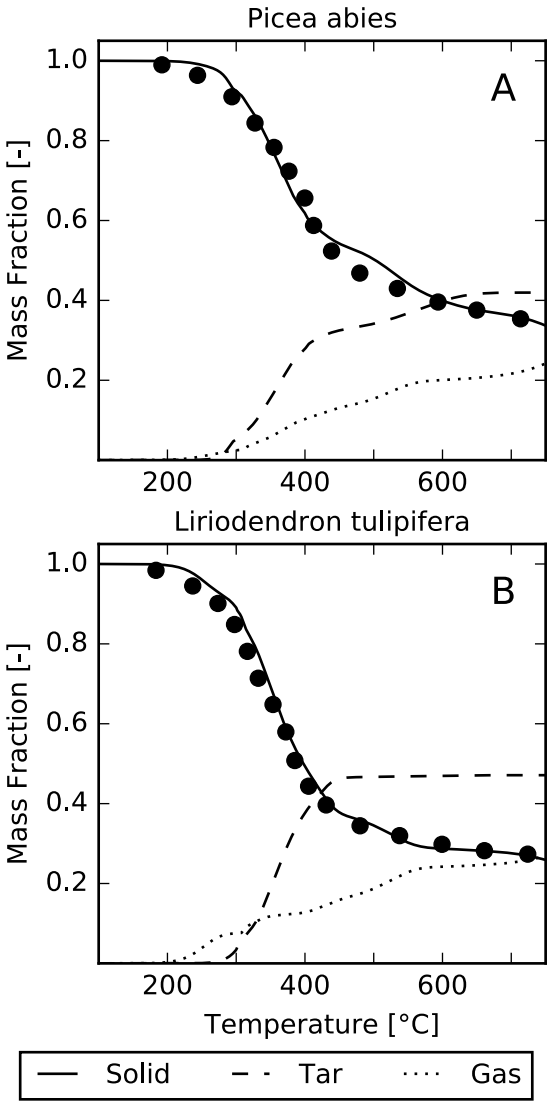


Figure 3.2: Yields of lumped products from the slow pyrolysis of two species of lignin. Predicted (lines) and experimental (points) yields of solid components are in excellent agreement. Experimental data is from Jakab et al. and does not include information on the yields of tars or gases.

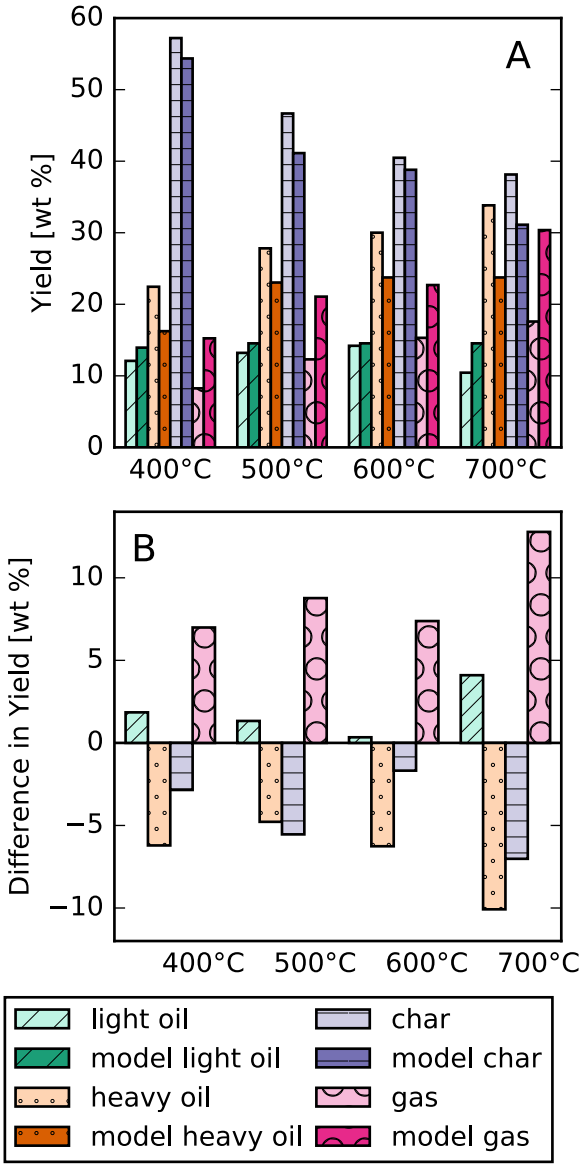


Figure 3.3: (A) Comparison of modeled integral yields of lumped components with experimental observations from Ben & Ragauskas.[23] Experimental and modeled data are for Kraft lignin heated at 2.7˚C/s to the maxmum temperatures shown. (B) The difference in the yield of lumped components between modeled and experimental values shown in panel A for each temperature.

In **Figure 3.3A** we compare the integral yields of all the lumped components with experimental observations from Ben and Ragauskas.[91] They investigated the pyrolysis of softwood kraft lignin

at a heating rate of 2.7°C s^{-1}. Volatiles were collected after completion of pyrolysis from condensers immersed in liquid nitrogen, and char was likewise collected. Gas yields were determined by mass difference. The liquid oil products contained two immiscible phases, referred to as heavy and light oil. To assign product species from our model to similar "phases" (gases, heavy oil, and light oil) we modeled each molecular species in Aspen Plus, v.8.6 (Aspen Technology Inc., USA). A feed stream of each pure component was fed into a flash reactor and run across a range of temperatures to determine the phase of the product stream. Parameters for the molecules were estimated using Aspen's NIST TDE evaluation. Species that are condensable at 30°C we defined as tar (oil), otherwise they are defined as gas. Within the tars, aromatic species were assigned to the heavy oil fraction and non-aromatics to the light oil. Our results generally show good agreement (less than 10% deviation from experiment) for the yields of all lumped components (**Figure 3.3B**). Two systematic trends are clear when comparing with this experimental dataset: our model over-predicts gas yield and under-predicts yields of heavy oil and char. Refining the kinetic parameters for the set of substitutive addition reactions in our scheme can reduce these deviations from experiment (not shown), but should be supported by additional theoretical or experimental work rather than being done for fitting purposes.

One of the primary reasons for making a more detailed kinetic model instead of the common "super lumped" models is in the ability to get more information about the composition of products. In our kinetic scheme the initial structure of lignin is not completely mechanistic for the reasons discussed earlier, so many product species do not represent actual molecules that would be observed experimentally. This makes it difficult to compare model predictions with experimental observations from techniques like GC-MS. We expect, however, that the chemical functionality of species in our model will approximate that of the precursors to real lignin

pyrolysis products. The addition of secondary tar cracking reactions will be necessary in later work to fully describe observed tar products.

However, with the level of detail present in our model we can describe the distribution of carbon functional groups in molecules that make up the tar fraction of our products. These functional groups are assigned from ^{13}C NMR chemical shift predictions made using the spectrum calculating tools at www.nmrdb.org. The molecules that make up heavy tars in our model should be considered tar precursors because they are not necessarily representative of the actual heavy tar species observed experimentally. For example, GC-MS results show that observed monophenols generally lack the three-carbon tails and associated functionality of our precursor components.[92-94] All of the experimentally observed carbon functional groups are possible in our model species with the exception of methyl aromatic groups (CH3-Ar and CH3-Ar').

Figure 3.4A compares the predicted distribution of carbon functional groups in heavy oils with quantitative ^{13}C NMR experiments by Ben and Ragauskas.[91] For most functional groups we find reasonable agreement, with carbonyl and general aliphatic C-C groups being the exception. General aliphatic C-C groups represent a significant fraction of our modeled light oil fraction, and we suspect that some of the species assigned to the light oil fraction in our model (based on molecular weight), are actually present in the heavy oil experimental fraction (defined by miscibility). If this is the case, and these species could be properly distributed between the light and heavy oil fractions, then we would expect to see an increased contribution from the model's general aliphatic C-C group accompanied by a concomitant decrease in the percentage of carbon attributed to the other functional groups.

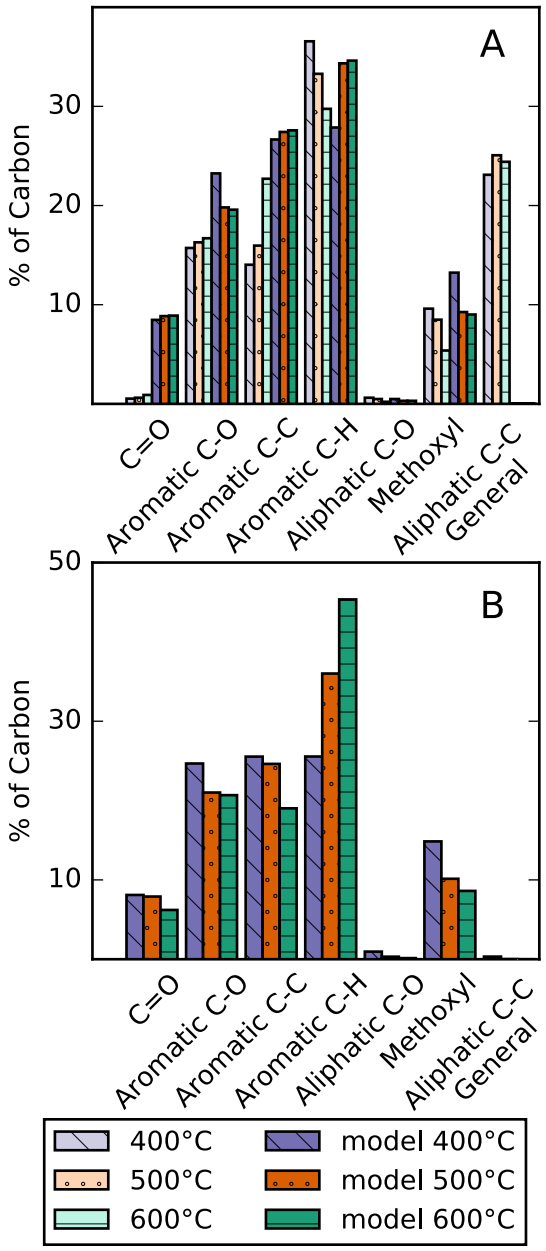


Figure 3.4: (A) Distribution of carbon functional groups present in the heavy oil fraction from slow pyrolysis (2.7˚C/s) of Kraft lignin. Experimental results (lighter colored bars) are the integrated totals of Ben and Ragauskas detected by ^{13}C NMR and the model results are assigned based on predicted ^{13}C NMR chemical shifts. Note that two experimentally observed functional groups, CH3-Ar, and CH3-Ar', are not possible in our model components and are not shown. (B) Model predictions for the distribution of carbon functional groups from the fast pyrolysis (2000˚C/s) of Kraft lignin.

Model parameters in our scheme are not fit to experimental observations from slow pyrolysis or TGA experiments, so we hope the model will be more generalizable to new reaction conditions including fast pyrolysis. However, as this is a purely kinetic model it must be validated with kinetically controlled experiments, which have historically been limited to slow pyrolysis reaction conditions. For slow pyrolysis with small particles, modelers generally assume a uniform temperature profile within particles and neglect mass transport effects.[11, 17, 95] Recent research calls into question this assumption that TGA experiments are in fact kinetically controlled,[12, 96] yet it remains standard practice to make this assumption when developing and using kinetic models. It is possible with our model to predict the products from fast pyrolysis reactions (**Figure 3.4B**), and we look forward to the availability of kinetically controlled lignin fast pyrolysis experiments to compare with. Experimental techniques to carry out such experiments have recently been developed, but have focused on generating results for cellulose pyrolysis studies.[97]

3.5 GLOBAL SENSITIVITY ANALYSIS

We carried out a variance-based global sensitivity analysis on the model to determine which parameters have the greatest impact on output measures of interest. A detailed description of variance-based sensitivity analyses and the particular methods used here are well covered by Saltelli[98-100] and Sobol.[101] In our analysis we examined the sensitivity of 30 model outputs to 410 input parameters: T_{max}, heating rate, carbon content, hydrogen content, and reaction rate constants for all 406 reactions. The range of parameter values explored was 400-950˚C for T_{max}, 5.0-120,000˚C/min for heating rate, 56.7-67.8 wt% C, and 5.4-6.6 wt% H. Rate constants are not usually adjustable parameters in our model, but we were interested in exploring how sensitive

the model results are to their values so they were allowed to vary one order of magnitude above and below their specified values in the model.

We used the Python package SALib to generate 2,055,000 parameter combinations based on the quasi-random Sobol sequence that uniformly cover the entire possible input parameter space.[102] After running simulations with all these parameter combinations, the variance in a given output measure is used to calculate first order, second order, and total order sensitivity indices with SALib. **Figure 3.5** shows the total sensitivity and second order indices for the top 10 most influential parameters on the yield of carbonyl groups in heavy and light oils. This is a representative sample of the type of information obtainable from this analysis, and the full sensitivity analysis results can be explored interactively using the Python package savvy, which we developed to visualize this type of high dimensional data.[103]

It is clear from **Figure 3.5** that if a certain set of output measures are important, this sensitivity analysis can inform decisions about the relative effort that should be spent accurately specifying specific input parameter values. In the example shown, if an accurate prediction of the carbonyl yield in the heavy oil fraction is desired, then it is important to precisely specify the values of certain rate constants, the initial lignin composition, and the maximum temperature. The specific rate constants that are important depends on the value of the sensitivity index that one chooses as a cutoff. If, on the other hand, the yield of carbonyl groups in light oils is all that matters, then only the hydrogen content is crucial and a user need not waste time or resources getting precise values of the other parameters.

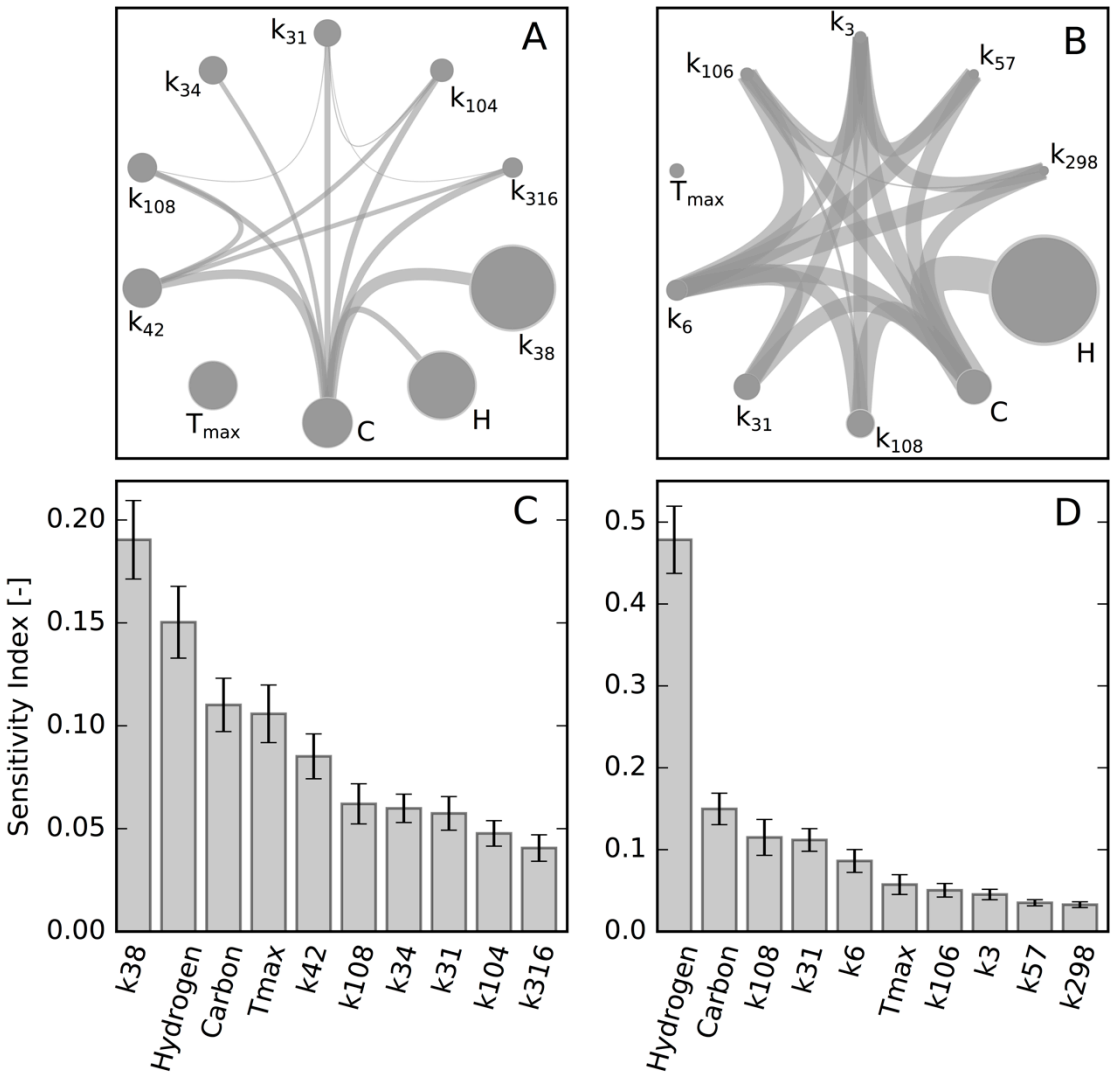


Figure 3.5: The total sensitivity indices for the yield of carbonyl groups in heavy (A, C) and light (B, D) pyrolysis oils. Only the 10 parameters with the highest total sensitivity indices are shown. In panels A and B the edge thickness between vertices corresponds to the second order sensitivity index, and the diameter of the vertex to the total order index.

Some general trends across output measures are observable in the sensitivity analysis results. We find that the lignin composition (carbon and hydrogen content) and maximum temperature are among the top 10 most influential parameters for nearly all of the 30 output measures examined. In this purely kinetic model, the value of the heating rate is rarely important in influencing output variance relative to other parameters, and only shows up in the top 10 for the yield of CO, total

solids, and total gases. Most of the kinetic parameters were assigned by Faravelli et al. from both literature values (which often vary widely for a given reaction within the literature), and from their experience with analogous reactions. Knowing which of these parameters the model is sensitive to enables us to recommend which reaction steps in the model are most important to measure or estimate more accurate parameters for. We find that the rate constants for reactions 31, 34, 38, and 48 are the most influential across all the output measures (they often appear in the top 6 parameters with the highest sensitivity indices). Reaction 31 is a molecular reaction, but the other three are lumped substitutive addition reactions built from lumped species (the full list of reactions is available in **Supporting Table B.1**). As the model is most sensitive to variations in the rate constants for these reactions, they are good candidates for further study and refinement, perhaps breaking them into less-lumped reactions, to ensure the values of their kinetic parameters are as accurate as possible.

An additional outcome of the sensitivity analysis was the identification of 64 reaction rate constants that have no influence on the variance of any of the 30 chosen output measures (**Supporting Table B.3**), which suggests those reactions could potentially be pruned from the kinetic scheme to reduce its computational cost as long as users are not interested in additional output measures. These 64 rate constants are all in the categories of recombination, char devolatilization, and hydrogen abstraction reactions. Finally, it is worth noting that the model is not highly sensitive to any of the reactions we added to Faravelli et al.'s original kinetic scheme, even though these reactions are critical to forming the intermediate species necessary for the kinetic scheme to proceed. Based on this observation we assume our calculated rate parameters for the new reactions are of the right order of magnitude.

3.6 ILLUSTRATIVE EXAMPLE

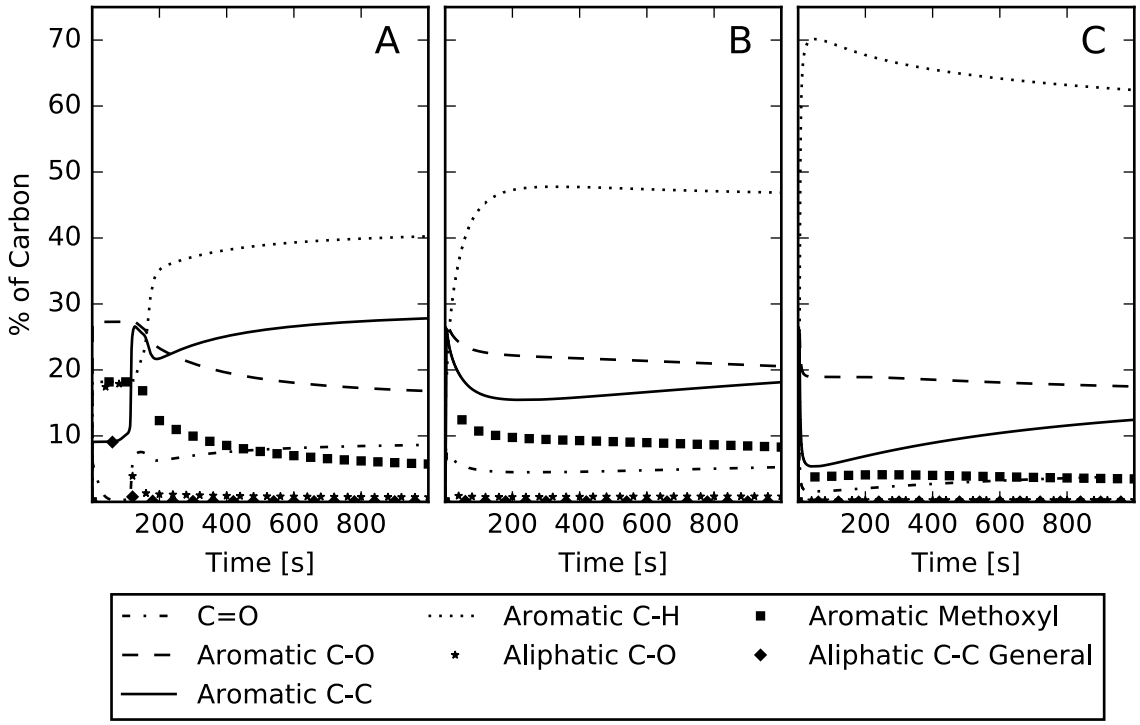


Figure 3.6: Predicted yields of different carbon functional groups in heavy oil during pyrolysis of Pseudotsuga menziesii. Heating rates and maximum temperatures are 162˚C/min to 500˚C (A); 2,000˚C/s to 500˚C (B); 2,000˚C/s to 600˚C (C).

Utilizing our model in practice is straightforward. The user specifies eight input parameters: the mass fractions of carbon and hydrogen, initial and maximum reactor temperatures, heating rate, and the durations of heating, isothermal, and cooling periods. Users that want to use an ODE solver other than DDASAC must pass the resulting ODEs to the solver of their choice. The range of model outputs includes any information that can be determined from the time distribution of the modeled molecular, radical, and lumped species. For example, **Figure 3.6** demonstrates how the model could be used to determine the impact of operating conditions like heating rate and maximum temperature on the yields of certain functional groups in heavy oils. To illustrate this more specifically, studies have shown that reacting ammonia with pyrolysis oils that have high carbonyl contents can produce a biodegradable slow-release fertilizer with lower leachability

than chemical fertilizers,[31] resulting in less ground water pollution and other benefits. By sampling various reaction conditions as in **Figure 3.6**, and looking at which of those conditions result in the highest yield of pyrolysis oils, a fertilizer manufacturer might select the operating condition with the best balance of oil yield and the desired property of higher carbonyl content.

3.7 CONCLUSIONS AND IMPLICATIONS

Existing lumped models for lignin pyrolysis lack the detail and flexibility necessary for use in optimizing industrial processes for making renewable fuels and chemicals. We presented a more detailed kinetic model that covers a wider range of feedstocks and reaction conditions than previous models. Without fitting any parameters, our model offers detailed descriptions of lignin pyrolysis products – beyond lumped yields – and our predictions agree well with available experimental data for slow pyrolysis. We carried out a comprehensive sensitivity analysis that identified specific reactions in the model that can benefit from further study.

A major challenge in this field is the difficulty reproducing and utilizing published models. Often the details included in literature are insufficient to quickly and reliably reproduce the authors' model, usually because key pieces are missing, are described vaguely, or have mistakes in them. Trying to reconstruct these models can take weeks, so there are few situations where any but the most basic lumped models are used and most researchers end up developing their own models. As mentioned earlier, to address this common problem we included the complete code to generate and run our model.[89] The code is written in Python, which is open source and readily available. We hope the ease of using this code will facilitate expansion and improvements to the model to include additional model components based on other bond types for characterizing

initial lignin structure, adding or refining the chemistry in the kinetic scheme, incorporating reactions for cellulose and hemicellulose, etc.

Chapter 4

APPLICATION OF MACHINE LEARNING TO PYROLYSIS REACTION NETWORKS: REDUCING MODEL SOLUTION TIME TO ENABLE PROCESS OPTIMIZATION

4.1 ABSTRACT

Comprehensive models of biomass pyrolysis are needed to develop renewable fuels and chemicals from biomass. Unfortunately, the detailed kinetic schemes required to optimize industrial biomass pyrolysis processes are too computationally expensive to include in models that account for both kinetics and transport within reacting particles. Here we present a machine learning approach using artificial neural networks and decision trees to reduce the computational expense of detailed kinetic models by four orders of magnitude, enabling their use in comprehensive models. The trained neural networks generalize very well, predicting the outputs of the detailed kinetic model with over 99.9% accuracy on new data. The machine learning approach we outline is not specific to kinetic modeling and can be applied to any set of input and output data, even if the underlying relationship between inputs and outputs is unknown.

4.2 INTRODUCTION

Dynamic mathematical models are commonly used in process engineering across disciplines as key tools for model-based control, product and process design, and optimization. In particular, detailed models based on the underlying physics of a scientific phenomenon can serve as tools that complement, and, in some cases, replace physical experiments. Model predictions can often

provide additional information that is difficult or impossible to measure, and are typically less expensive and take less time than experiments. Even so, in many cases models are still too computationally expensive, especially when modeling in two or three dimensions, so in this chapter we demonstrate a robust approach that drastically reduces the computational cost of evaluating a kinetic model for lignin pyrolysis. This approach is generalizable to any set of data with an underlying relationship – even if that relationship is unknown – so can be applied to experimental observations in addition to mathematical models.

Interest in developing comprehensive models that describe biomass pyrolysis is high due to growing scientific and industrial interest in lignocellulosic biomass as a feedstock for renewable fuels and chemicals. These models couple descriptions of chemical kinetic mechanisms for biomass degradation with conservation equations for heat and mass transport.[10, 11] Numerical solutions to such complex models, especially in two- or three-dimensions, are computationally expensive. To reduce the computation cost, published comprehensive models incorporate many simplifying assumptions for both the kinetic and transport processes – especially those that use CFD approaches to move beyond single particles and model whole reactors.[11, 96, 104-107]

All of these comprehensive models use kinetic schemes where both the reactions and the species involved are highly lumped. Such lumped kinetic schemes have a number of limitations, but from a process development and engineering standpoint the two key limitations are the facts that lumped models cannot predict the detailed product compositions necessary for optimizing pyrolysis reactors, nor are they accurate across ranges of reactor conditions or feedstocks.[10, 12] Research into developing more detailed or mechanistic kinetic schemes for biomass components

to address these limitations is ongoing,[15, 108] but the challenge of incorporating such schemes involving hundreds to thousands of reactions and species into comprehensive models, and solving them in a reasonable amount of time, has not been solved.[109]

We tackle this challenge using our previously developed kinetic model for lignin pyrolysis (Chapter 3). Lignin – a highly cross-linked amorphous copolymer of three randomly polymerized phenylpropane units – accounts for up to 35wt% of dry biomass, and carries the highest specific energy content of the three principal components of biomass.[76] Today, most lignin is burned for process heat,[78] however the economic viability of future lignocellulosic biorefineries will depend on making lignin a high-value intermediate instead of a low-value fuel. Doing so requires the development of predictive comprehensive models that can be used to optimize fast pyrolysis processes to produce valuable renewable fuels and chemicals from lignin. As discussed earlier, existing comprehensive models lack sufficient kinetic detail to be useful for this type of optimization.

Our detailed kinetic model, involving ~100 species and 400 reactions, predicts the evolution of molecules and functional groups during lignin pyrolysis (Chapter 3). The model shows good agreement with experiment, but solving the complete set of stiff ODEs is too computationally expensive to incorporate into comprehensive particle or CFD models. In this work we demonstrate an approach whereby machine learning algorithms called artificial neural networks (neural nets) can learn to predict the results of our kinetic model with high accuracy and a computational cost up to four orders of magnitude lower than solving the ODE model.

Neural nets are algorithms inspired by biological neurons for learning complex non-linear and multivariate relationships. They act as a "black box" which is fed a set of input data and produces some desired output, and are widely used across disciplines for tasks like prediction, classification, and pattern recognition. Neural nets have been applied to a number of chemical engineering problems,[110-114] and even to kinetic model reduction.[115, 116] In situations where detailed kinetic schemes can't be generated due to insufficient knowledge of governing mechanisms, for example when there are difficulties identifying and quantifying non-stable intermediates in a mechanism, neural nets also offer an attractive alternative.[116] To our knowledge neural nets have not been applied to the comprehensive modeling of biomass pyrolysis particles - where the computational cost is limiting further development.

Structurally, neural nets contain layers made up of simple processing elements called nodes that function as non-linear summing units. These nodes are interconnected with weighted connection lines between layers, where each layer receives its input from the previous layer and the output of each node feeds the next layer. The weights are adjusted iteratively during a training process, resulting in networks that can be highly accurate and extremely computationally efficient. The mathematical features of neural nets are explored in textbooks,[117, 118] and are not described here.

4.3 EXPERIMENTAL

With neural nets only the relevant input and output signals from the kinetic model are observed while training the network. A neural net approximates the complex functional dependence of the kinetic model's system of ODEs, without any explicit knowledge of these mathematical expressions. In this work, neural networks were constructed and trained with the following steps:

1. Choice of relevant input and output variables for lignin pyrolysis

2. Generating training and validation data from the detailed kinetic model
3. Choice of neural network architecture and learning algorithms
4. Training the neural networks

4.3.1 *CHOICE OF INPUT AND OUTPUT VARIABLES*

The kinetic model that we trained neural nets to reproduce predicts the temporal evolution of molecules and functional groups during lignin pyrolysis and is described in Chapter 3. Inputs for both the kinetic model and our neural nets are: the maximum pyrolysis temperature (T_{max}), the heating rate, and the mass fractions of carbon and hydrogen in the lignin feedstock. We trained our neural nets to predict the final values of 30 output measures after simulated pyrolysis in a batch reactor, so time was not considered as an input parameter, however we plan to extend the neural networks in future work to predict temporal changes.

The particular output measures we trained are listed in **Supporting Table C.1** and describe the yields of specific functional or lumped groups; we used these outputs because they were part of a large data set already generated for previous work, but we could have used any set of outputs that the kinetic model is able to generate including concentrations or yields of individual or lumped species, or values derived from those quantities.

4.3.2 *GENERATING TRAINING, VALIDATION, AND TEST DATA*

Training any supervised learning algorithm like a neural network involves exposing it to a sufficient amount of input-output training pairs. In order for the neural net to generalize well it is important that this training data covers the entire range of parameter values that it should be able to predict outputs for. We used the Python package SALib to generate 250,000 parameter combinations based on the quasi-random Sobol sequence that uniformly cover then entire

possible input parameter space.[102] The range of parameter values explored was 400 – 950˚C for T_{max}, 5 – 120,000˚C/min for heating rate, 56.7 – 67.8 wt% carbon, and 5.4 – 6.6 wt% hydrogen. These ranges for carbon and hydrogen cover the allowed values in the kinetic model. To generate the corresponding output data we ran the generated set of input parameters through our detailed kinetic model, resulting in 250,000 input-output learning pairs where each input contains the four parameter values, and each output contains the values of the 30 output measures.

It is common practice in supervised learning to split available learning data into three sets: a training set that is used to adjust the weights between nodes in the network; a validation set that is used to see how the trained network is performing on data that it has not been exposed to for the purpose of model selection and to decide when training is complete; and a test set to assess the performance of the final selected model on data it has never seen before. To set up these splits we first performed some pre-processing of the learning data by randomly shuffling the 250,000 learning pairs. After shuffling we divided the learning data into a training set with 140,000 input-output pairs, a validation set with 60,000 pairs, and a test set with 50,000 pairs. Then we standardized both the inputs and outputs to have a mean of zero and unit variance. We standardized the inputs in this way in order to prevent terms with higher magnitude, or larger ranges, from dominating the model, and to speed up convergence,[119] and did the same to the outputs because our error function (mean squared error) is scale-sensitive so variability in output values would otherwise impact how well the net learns different outputs.

4.3.3 *CHOICE OF NETWORK ARCHITECTURE AND LEARNING ALGORITHMS*

We took two approaches to designing multi-layer feed-forward neural networks that could approximate our kinetic model; one strategy was to train a single neural net to predict all 30

output measures – which we will call the "full net" (**Figure 4.1**A), while the other strategy was to train 30 independent neural nets, each of which predicts a single output measure – called the "single nets" (**Figure 4.1B**). For all of these neural nets we included two dense hidden layers, as we found that a single hidden layer was often insufficient to adequately approximate the kinetic model (predictions for some single-layer networks resulted in unacceptably low R^2 values, around 0.5, for some output measures). To determine the optimal number of nodes in each hidden layer we carried out an exhaustive grid search over the number of nodes in each layer, within the range of 5 to 32 nodes, using the grid search tools in Scikit-learn.[120] The network architectures with the lowest mean squared error on the validation set were selected for further training (**Supporting Table C.2**).

Activation functions in the hidden layers are hyperbolic tangents of the form recommended by LeCun et al.:[119]

$$f(a) = 1.7159 \tanh(\tfrac{2}{3} a) \qquad \textbf{(4.1)}$$

where a is the activation, a linear combination of the outputs from the previous layer in the network. Activation functions in the output layer are linear as is standard in regression problems. We used an adaptive learning rate algorithm, adagrad,[121] to adjust the weights and biases to minimize the error function during the training process. Weights were initialized using Glorot weight initialization with weights sampled from a uniform distribution.[122]

4.3.4 *TRAINING THE NEURAL NETS*

After selecting the optimal network architectures from the grid search results the neural nets were trained using a batch size of 128 for a maximum of 6000 epochs. Training ended earlier if

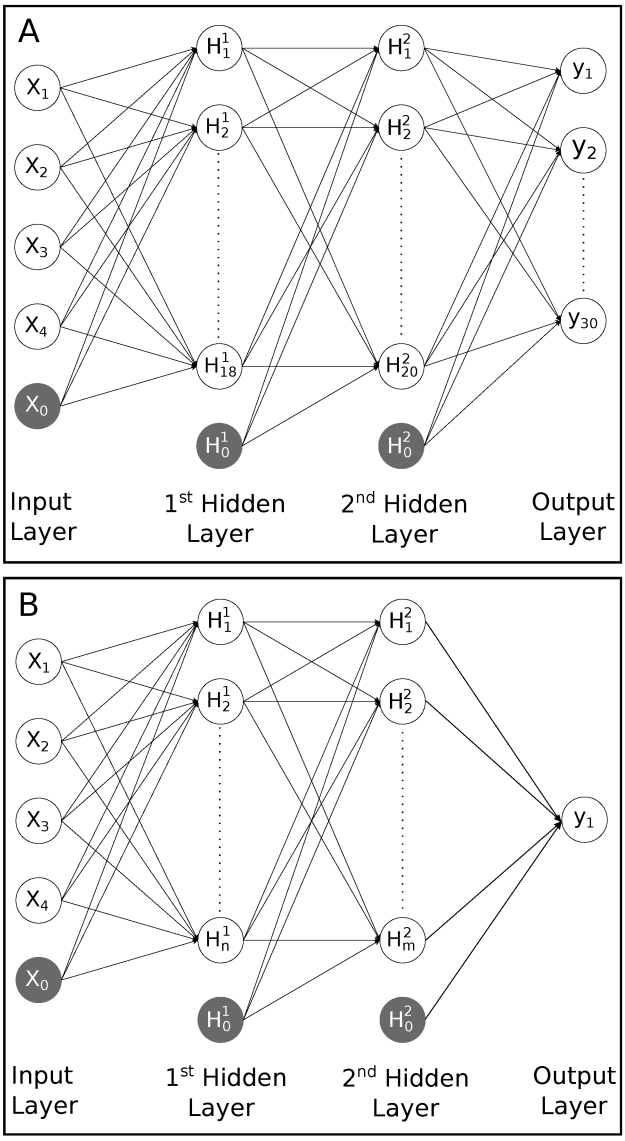


Figure 4.1: Network architectures for the full net (A), and single nets (B). The input, hidden, and output variables are represented by nodes, and the weight parameters by links between nodes. Bias parameters are represented by solid nodes, X_0 and H^i_0. Subscripts *n* and *m* refer to the number of nodes in the hidden layers of the single nets (**Supporting Table C.2**).

the error on the validation set increased for a sustained number of epochs. The Python package *nolearn* was used to build and train the neural nets.[123]

4.3.5 *Training a Decision Tree*

In addition to the neural nets we also trained another type of machine learning model, called a decision tree. We explored the use of decision trees because they are faster to train than neural nets (three seconds instead of one to several hours for the neural networks), and because they are less complicated so can be constructed by practitioners without the deep understanding required for training neural nets. We used the same learning data to train the decision tree as the neural networks, and followed a similar grid search procedure to optimize its architecture and determine the maximum depth and minimum number of samples in each leaf of the tree (**Supporting Table C.2**). Other hyperparameters of the decision tree were assigned default values from Scikit-learn's DecisionTreeRegressor class, and Scikit-learn was used to train the tree.[120] Our Python code to set up and train the neural nets and decision tree is available online.[124]

4.4 Results & Discussion

After the training procedure each machine learning model was evaluated based on how well it predicts the output of the ODE-based kinetic model. To this end each model was provided the input data from the test set, which none of the models had previously been exposed to, and model predictions were compared with the expected test set outputs. **Table 4.1** shows the coefficient of determination for the test set predictions of each output measure for each of the machine learning models. **Figure 4.2 - Figure 4.4** are representative parity plots showing two output measures with high R^2 values and two with low R^2 values for the single nets, full net, and decision tree, respectively. The complete set of parity plots for all output measures are available in the Appendix (**Supporting Figure C.1 - Supporting Figure C.3**).

Table 4.1: Coefficients of determination (R^2) of the predicted and actual values on the test set for each of the trained machine learning models. Max (green) and min (pink) values are highlighted.

Output measure	Single net	Full net	Decision tree
CO	0.9990	0.9909	0.9921
CO2	0.9999	0.9938	0.9977
H-C	0.9988	0.9930	0.9959
H2O	0.9992	0.9938	0.9952
Moisture	0.9983	0.9930	0.9957
heavy_C--O	0.9965	0.9862	0.9846
heavy_aliphatic C-C	0.9967	0.9935	0.9824
heavy_aliphatic_C-O	0.9974	0.9869	0.9879
heavy_aromatic_C-C	0.9987	0.9923	0.9892
heavy_aromatic_C-H	0.9996	0.9975	0.9970
heavy_aromatic_C-O	0.9996	0.9965	0.9972
heavy_aromatic_methoxyl	0.9995	0.9970	0.9974
heavytars	0.9981	0.9907	0.9912
light_C--O	0.9986	0.9910	0.9923
light_aliphatic_C-C	0.9997	0.9945	0.9960
light_aliphatic_C-O	0.9979	0.9860	0.9846
light_aromatic_C-H	0.9994	0.9950	0.9954
light_aromatic_C-O	0.9995	0.9954	0.9954
lighttars	0.9994	0.9979	0.9972
othergases	0.9989	0.9910	0.9957
solids	0.9982	0.9917	0.9909
tot_C--O	0.9993	0.9945	0.9979
tot_aliphatic_C-C	0.9998	0.9969	0.9979
tot_aliphatic_C-O	0.9982	0.9902	0.9933
tot_aromatic_C-C	0.9983	0.9935	0.9909
tot_aromatic_C-H	0.9995	0.9981	0.9974
tot_aromatic_C-O	0.9994	0.9965	0.9970
tot_aromatic_methoxyl	0.9994	0.9972	0.9973
totalgases	0.9992	0.9945	0.9938
totaltars	0.9984	0.9947	0.9946

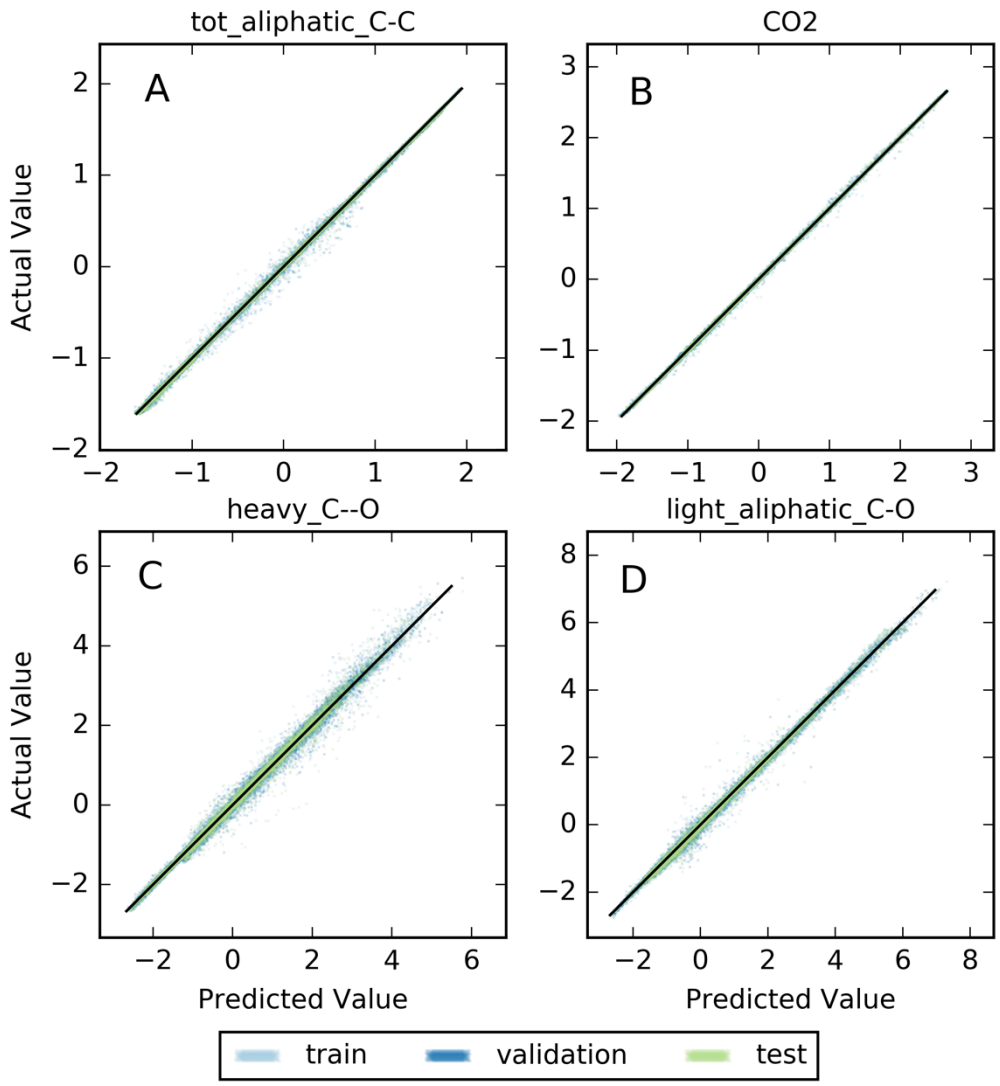


Figure 4.2: Representative parity plots for the training, validation, and test sets of four single net models. Panels (A) and (B) present two nets with high R^2 values on the test set, while (C) and (D) are nets with relatively low R^2 values.

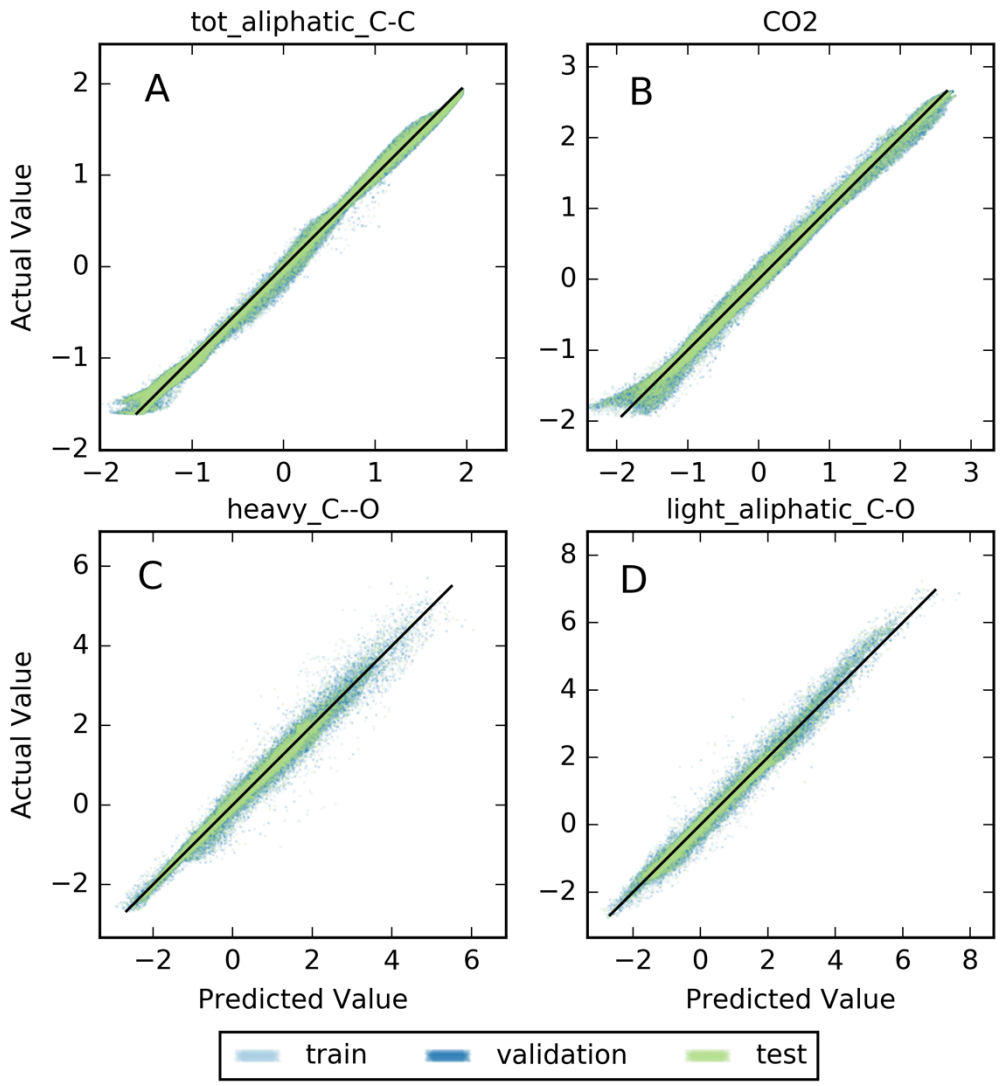


Figure 4.3: Representative parity plots for the training, validation, and test sets of four output measures in the full net. Panels (A) and (B) present two measures with high R^2 values on the test set, while (C) and (D) are measures with relatively low R^2 values.

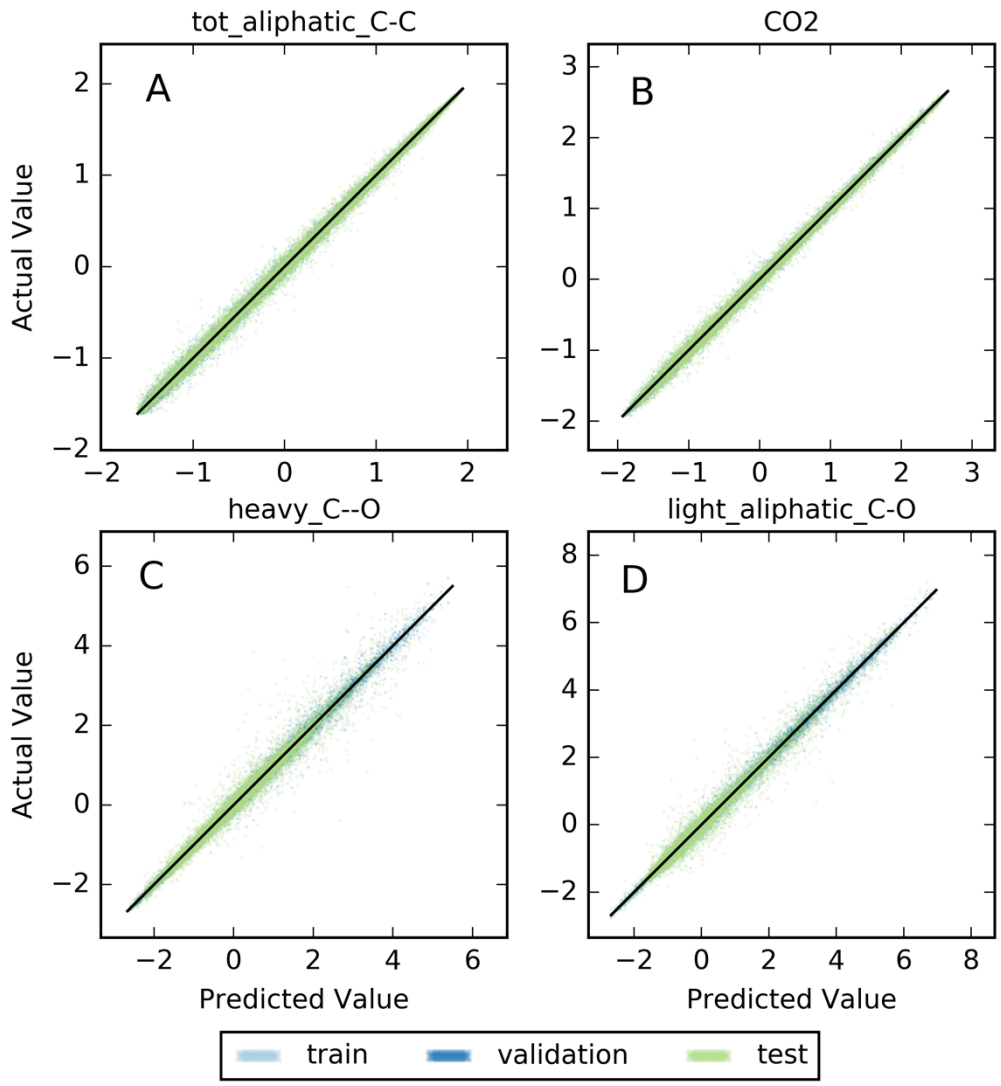


Figure 4.4: Representative parity plots for the training, validation, and test sets of four output measures in the decision tree. Panels (A) and (B) present two measures with high R^2 values on the test set, while (C) and (D) are measures with relatively low R^2 values.

These figures along with **Table 4.1** make it clear that the single nets are able to predict all the output measures with excellent accuracy, with over 99.9% of the variation in actual values predicted by the single net models for most outputs. Both the full net and decision tree also predict the output measures very well ($R^2 > 0.982$ for all outputs), and achieve comparable performance to each other on the test set.

Histograms showing the distribution of the mean squared error (MSE) for the training, validation, and test sets are available in **Supporting Figure C.4** – **Supporting Figure C.6**. The neural net models show a relatively even distribution of errors between the training, validation, and test sets – indicating good generalization – while the decision tree has very low error on the training set and relatively more error in the test set than the neural nets, which can be a sign of poor generalization (**Supporting Figure C.6**). In our case, the very large training set compensates for this tendency of decision trees to overfit, and the spread of MSE's for the decision tree is consequently nearly the same as that of the full net. This is not the case when less training data is available (**Figure 4.5**), and we see that while the average MSE of the neural net on both the training and validation sets levels-off with around 35,000 training pairs, the decision tree requires 105,000 more training pairs to approach the same validation error. More advanced approaches to using decision trees, such as using them within an ensemble, have not been explored here but might mitigate this issue.

The purpose of training these machine learning models was to reduce the computational cost of predicting the outputs of a complicated kinetic model so that detailed kinetics can be used in comprehensive particle modeling. To assess the reduction in computational cost we performed benchmarking tests of all the models (**Table 4.2**). Benchmarking accounted for the time to predict the output measures and convert them from standardized to true values. For the complete ODE model only the portion of the code where the system of ODEs is solved was included; we ignored steps to set up and write model equations. Each model was evaluated on a randomly chosen set of inputs 1000 times on a Linux workstation with two four core 2.27 GHz Intel Xeon E5607 processors and 5.7 GB of memory.

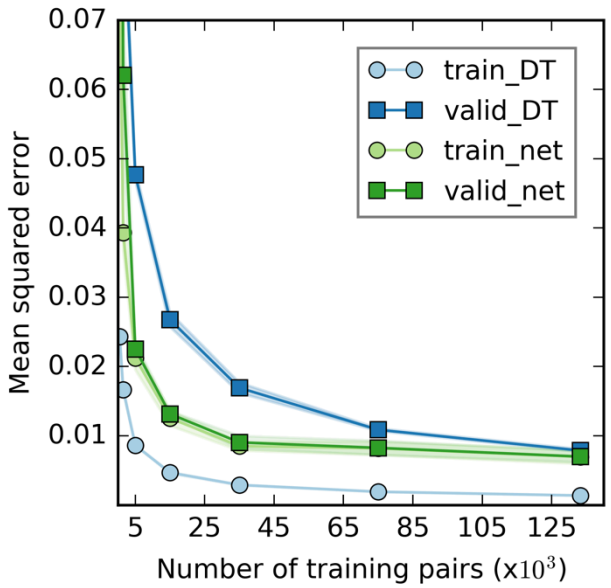


Figure 4.5: Learning curve depicting the change in mean squared error with the amount of available training data for the training and validation sets of a decision tree (DT) and a neural net (net) with the architectures described herein. 3-fold cross-validation was used to split the training data; lines connect the average MSE while the shaded regions represent the entire range of MSE scores at each training subset size.

Table 4.2: Benchmarking results for solving the kinetic model. The time reported is the average of 1000 model calls.

Kinetic model format	Average code execution time (s)
Decision tree	1.106×10^{-4}
Full net	1.690×10^{-4}
Single net	1.747×10^{-4}
30 single nets (run serially)	4.236×10^{-3}
Complete ODE model[108]	4.725

Using the machine learning models we are able to reduce the computation expense by three to four orders of magnitude, depending on the level of accuracy that is desired and the number of output parameters required. With these speeds it is possible to evaluate detailed reaction kinetics

in CFD simulations even faster than using the simple lumped ODE-based kinetics that are the current standard.

The choice of which machine learning model to use depends on the level of accuracy required and the amount of training data available. When accurate model predictions are most important, the single nets are the best choice. By itself, each single net is faster to train than the full net (though predicting the output from one trained net takes essentially the same amount of time as for the full net), and it is not necessary to predict all 30 output measures if they are not all of interest at a given time. If there is more room for error in the predicted output measures, and a sufficient amount of training data is available such that the decision tree performs comparably to the full net, then using the decision tree is the best choice because it is faster to solve and faster to train. While not a concern on most modern computers, it should be noted that the trained decision tree is over 100 times larger on disk than a trained neural net (170 MB vs. 1.3 MB).

It should be stressed that at this time the machine learning models presented in this work cannot replace full ODE model evaluation in general for the current system because our neural nets and decision tree do not account for temporal changes in the reacting system. Extending these models to include time as an additional input parameter is possible, and mathematically trivial – it would involve repeating all the steps outlined in the Experimental section, but also adding time to the four existing input parameters, so the input layer of each neural net will have five nodes instead of four. A significant challenge with this approach will be the large file sizes required for the learning data, because each set of our four existing inputs will need to be associated with descriptions of potentially thousands of time points, depending on the desired temporal

resolution. It is also likely that neural nets for such systems will be more complex (more nodes in the hidden layers). While such nets would be more expensive to train, a one-time cost, their execution times for predicting the temporal output of the kinetic model would be similar to the benchmarking times in **Table 4.2**.

4.5 CONCLUSIONS

In this work we've outlined an approach to kinetic model reduction that reduces the computational expense of computing a detailed kinetic model for lignin pyrolysis by up to four orders of magnitude. Our neural network and decision tree models demonstrate excellent agreement with expected final values for the outputs from simulated reactions in a batch reactor. In future work we plan to extend the models to predict the temporal values of these outputs as discussed above.

The machine learning approach described is inherently generalizable and is not specific to kinetic modeling. Any set of input and output data that is based on some underlying relationship, even if that relationship is unknown, can be approximated with a neural network. Therefore, using this approach we need not fully understand the underlying physical phenomena to develop a robust, computationally efficient predictor for a given system.

Within the biomass pyrolysis realm, this means that we might skip kinetic modeling altogether, and instead generate neural networks that predict output measures of a specific experimental or industrial reactor based entirely on observed experimental data. The challenge with such an approach will be in generating sufficient training data that spans the range of conditions we might model. When sufficient data is not available ahead of time, online learning can be used to

train the network as new data is generated, so as a reactor is operated the neural net predictions will improve continuously. For a computational model, generating this data is a one-time cost after which training the neural net(s) is relatively quick, and predicting the output values from a trained network is extremely fast, as we've demonstrated. In situations where the model will only be evaluated a limited number of times, or if it does not have a pre-determined parameter range, such machine learning approaches will not improve the overall computational efficiency.

Chapter 5

CONCLUSIONS

This chapter summarizes the research described in this dissertation and its significance, and provides recommendations for future work.

5.1 SIGNIFICANCE AND CONTRIBUTIONS OF THIS WORK

As described in Chapter 1 and the introductions to each subsequent chapter, there are many challenges facing the biomass pyrolysis community that need to be solved before comprehensive models of pyrolyzing particles and reactors become useful for industry-scale optimization. We sought to address some of these challenges in the work described in this dissertation:

1. Existing experimental data sets have been unable to capture the spatiotemporally-resolved composition data needed to validate comprehensive particle models.
2. Particle shrinkage is rarely included in comprehensive models because the mechanisms driving it are unknown.
3. The lumped models commonly used today cannot predict detailed product properties or compositions.
4. Lumped models parameterized from TGA experiments are not generalizable and can't be applied to new feedstocks or outside the range of conditions of the original experiments.
5. Detailed kinetic schemes, necessary for addressing the previous two points, are much too computationally expensive for use in comprehensive particle models.

In Chapter 2 we used a new experimental technique, STR-DRiSP, to develop a comprehensive 2D reaction-transport model for fast pyrolysis that addresses challenges 1 and 2. Ours is the first comprehensive wood particle model to confirm the presence of a visible carbohydrate reaction front during fast pyrolysis under industrial conditions. We contributed a new method for

modeling the shrinkage of particles, which is generally ignored in existing models but was critical for agreement with the experimental results in our study. Our findings support the hypothesis that lignin decomposition contributes more to particle shrinkage than carbohydrate decomposition.[70] In developing this model we experienced challenges 3-5 directly; since no detailed kinetic models for whole biomass (as opposed to individual components like cellulose, lignin) exist, we used a common lumped model for this project, capable only of predicting the amounts of char, volatiles, and gas. The kinetic model we used is derived from slow pyrolysis TGA data, and while it performed well enough to validate the new experimental technique, it exhibits discrepancies at higher temperatures. Even if a detailed kinetic model were available, incorporating it into the CFD model would be impractical: our simple lumped model involving 10 reactions with 16 species resulted in simulations that took 6 days each to run.

In Chapter 3 we applied a detailed kinetic model for lignin pyrolysis to address challenges 3 and 4. Our model, involving ~100 species and 406 reactions, covers a wider range of feedstocks and reaction conditions than previous models. Without fitting any parameters, our model predicts detailed descriptions of lignin pyrolysis products and shows great agreement with available lumped experimental results across a wide range of conditions. Our detailed global sensitivity analysis is the most comprehensive we have seen in the pyrolysis modeling literature, and allowed us to identify specific reactions that may benefit from further study. We found that the kinetic model is most sensitive to variations in the composition of the feedstock and the maximum pyrolysis temperature, but that changes in the value of the heating rate don't contribute much to variance in the model outputs.

Finally, in Chapter 4 we used machine learning to tackle challenge 5. Using our detailed kinetic model from Chapter 3, we trained neural networks and decision trees to reproduce the kinetic model outputs with excellent agreement, but a computational cost four orders of magnitude lower. With this reduction in computational cost it is now possible to evaluate detailed pyrolysis reaction kinetics in comprehensive CFD particle simulations even faster than using simple lumped models like the one we used in Chapter 2. The machine learning approach that we present is generalizable – it does not require a detailed kinetic model, or even any physics-based understanding of a process to model that process.

5.2 Additional Research Deliverables

A major challenge in this field is the difficulty reproducing and utilizing published models or making efficient use of existing work. To address this challenge we made an effort to document and make available online the software written while carrying out this research. Each project's code is citable with a DOI and written in Python, which is open source and readily available. Citations to the relevant code are included in the preceding chapters, and the projects listed here:

ligpy: the detailed kinetic model for lignin pyrolysis developed in Chapter 3
https://github.com/houghb/ligpy

savvy: a tool to visualize high dimensionality sensitivity analysis data
https://github.com/houghb/savvy

lignet: scripts to create, train, and assess neural networks and decision trees for Chapter 4
https://github.com/houghb/lignet

5.3 Recommendations for Future Work

The work described in this dissertation represents a few steps on a long journey to robust biomass pyrolysis models. Here we recommend logical next steps for each of the projects.

5.3.1 *CHAPTER 2 RECOMMENDATIONS*

Correlations between the developed model and experiment provided new insights into particle shrinkage during fast pyrolysis that suggested better approaches to incorporating predictive shrinkage into future particle models. We recommend future CFD work should attempt to model shrinkage by allowing the height of individual mesh elements to change based on their composition, starting with tying the height to the lignin content, so that the entire particle shrinks and warps asymmetrically as observed in experiments. While the computational cost of doing so may be significant, it could offer the first accurate approach to predictive shrinkage modeling.

Our sensitivity analysis identified five parameters to accurately specify in the model: wood density, heat capacity, thermal conductivity, and the rate constants for the initiation reactions for cellulose and hemicellulose. The issue of the rate constants will be addressed by ongoing work in kinetic model development as it moves toward more mechanistic models. Significant effort needs to be put into developing a database of the physical properties of different species of wood, including those listed above, with a particular focus on how those properties change as the particle reacts. Existing transport models, like that developed in Chapter 2, rely on limited and often obscure values for these parameters that are based on a small set of available measurements for limited species of wood where measurements – if made at all – are done on virgin biomass, and sometimes on the resulting char, but don't account for changes during pyrolysis. These values range wildly from one source to another, even for the same species of wood, and the choice of a particular value to use has significant impacts on model behavior.

5.3.2 *CHAPTER 3 RECOMMENDATIONS*

Two systematic trends are clear when comparing the predictions of our kinetic model to lumped yields from slow pyrolysis experiments: the model over-predicts the gas yield and under-predicts the yields of heavy oil and char. We found that adjustments to the kinetic parameters for the substitutive addition reactions in the kinetic scheme can be made to achieve a better fit to the data. We recommend that any such changes are based on future work to better understand these reactions and perhaps split them into more mechanistic steps that might be studied computationally or experimentally to determine more accurate values for their kinetic parameters. The same should be done for the set of reactions that our sensitivity analysis identified as having the highest sensitivity indices.

Our model does well predicting the distribution of functional groups in tars from lignin pyrolysis, but the tar molecules produced in the model should be considered "tar precursors" as described in section 3.4.1. Future work should add secondary tar cracking reactions to the kinetic scheme so that predicted products more closely resemble the species observed by GCMS.

A recently developed experimental technique is making it possible to gather millisecond-resolved data on the composition of volatiles during kinetically-controlled fast pyrolysis for the first time.[97] As this technique is very new it has not yet been used to characterize the products of lignin fast pyrolysis, but when that data is available it will provide the first opportunity to compare our kinetic model's fast pyrolysis predictions with kinetically-controlled experimental data. This model should then be re-evaluated to make sure it is accurately capturing the kinetics at those conditions.

5.3.3 *CHAPTER 4 RECOMMENDATIONS*

The next step in developing these neural network models for biomass pyrolysis is extending them to include temporal changes in output values. As outlined in Chapter 4, the steps to do this are straightforward, but they require a new set of skills for dealing with "big data". The training data sets for these future neural networks (or decision trees) will be very large and should be stored in databases in the cloud. Training of the neural nets may also be more expensive. A cloud-based platform such as Amazon's EC2 will allow training of many networks in parallel and provide the infrastructure for accessing the training data database.

Once the neural nets are trained to include time as an input parameter, they should be incorporated into a comprehensive CFD simulation to demonstrate the possibility of using detailed reaction kinetics in this type of simulation for the first time.

A parallel extension of this work should be to train neural nets or decision trees based on just experimental observations for pyrolyzing systems – skipping any physics-based models entirely. This approach holds great promise for modeling, prediction, and control of industrial reactors where it may not be necessary to understand exactly why certain operating conditions lead to a specific outcome, as long as you can reliably predict that outcome. As mentioned in Chapter 4, the limiting factor to this approach will be generating sufficient experimental data to train a reliable model, so it may be necessary to switch from the batch learning used in this work to an online learning approach.

BIBLIOGRAPHY

1. Huber, G. W.; Iborra, S.; Corma, A., Synthesis of transportation fuels from biomass: Chemistry, catalysts, and engineering. *Chemical Reviews* **2006,** 106, (9), 4044-4098.
2. Bridgwater, A. V., Review of fast pyrolysis of biomass and product upgrading. *Biomass & Bioenergy* **2012,** 38, 68-94.
3. Mohan, D.; Pittman, C. U.; Steele, P. H., Pyrolysis of wood/biomass for bio-oil: A critical review. *Energy & Fuels* **2006,** 20, (3), 848-889.
4. Babu, B. V., Biomass pyrolysis: a state-of-the-art review. *Biofuels Bioproducts & Biorefining-Biofpr* **2008,** 2, (5), 393-414.
5. Kersten, S.; Garcia-Perez, M., Recent developments in fast pyrolysis of ligno-cellulosic materials. *Current Opinion in Biotechnology* **2013,** 24, (3), 414-420.
6. Mettler, M. S.; Vlachos, D. G.; Dauenhauer, P. J., Top ten fundamental challenges of biomass pyrolysis for biofuels. *Energy & Environmental Science* **2012,** 5, (7), 7797-7809.
7. Thomas, R. J., Wood Structure and Chemical Composition. *Goldstein, Irving S. (Ed.). Acs (American Chemical Society) Symposium Series, Vol. 43. Wood Technology: Chemical Aspects. San Francisco, Calif., USA, Aug. 31-Sept. 2, 1976. X+372p. Illus. American Chemical Society: Washington, D.C., USA. Isbn 0-8412-0373-3* **1977**, 1-23.
8. Laboratory, F. P., *Wood handbook : wood as an engineering material*. Forest Products Society: Madison, Wis., 2011.
9. Gibson, L. J., The hierarchical structure and mechanics of plant materials. *Journal of the Royal Society Interface* **2012,** 9, (76), 2749-2766.
10. Di Blasi, C., Modeling chemical and physical processes of wood and biomass pyrolysis. *Progress in Energy and Combustion Science* **2008,** 34, (1), 47-90.
11. Moghtaderi, B., The state-of-the-art in pyrolysis modelling of lignocellulosic solid fuels. *Fire and Materials* **2006,** 30, (1), 1-34.
12. White, J. E.; Catallo, W. J.; Legendre, B. L., Biomass pyrolysis kinetics: A comparative critical review with relevant agricultural residue case studies. *Journal of Analytical and Applied Pyrolysis* **2011,** 91, (1), 1-33.
13. Bradbury, A. G. W.; Sakai, Y.; Shafizadeh, F., Kinetic-Model for Pyrolysis of Cellulose. *Journal of Applied Polymer Science* **1979,** 23, (11), 3271-3280.
14. Miller, R.; Bellan, J., A generalized biomass pyrolysis model based on superimposed cellulose, hemicellulose and lignin kinetics. *Combustion Science and Technology* **1997,** 126, (1-6), 97-137.
15. Vinu, R.; Broadbelt, L. J., A mechanistic model of fast pyrolysis of glucose-based carbohydrates to predict bio-oil composition. *Energy & Environmental Science* **2012,** 5, (12), 9808-9826.
16. Klein, M. T.; Virk, P. S., Modeling of lignin thermolysis. *Energy & Fuels* **2008,** 22, (4), 2175-2182.
17. Faravelli, T.; Frassoldati, A.; Migliavacca, G.; Ranzi, E., Detailed kinetic modeling of the thermal degradation of lignins. *Biomass & Bioenergy* **2010,** 34, (3), 290-301.
18. Pyle, D. L.; Zaror, C. A., Heat-transfer and Kinetics in the Low-temperature Pyrolysis of Solids. *Chemical Engineering Science* **1984,** 39, (1), 147-158.
19. Bryden, K. M.; Hagge, M. J., Modeling the combined impact of moisture and char shrinkage on the pyrolysis of a biomass particle. *Fuel* **2003,** 82, (13), 1633-1644.

20. Mamleev, V.; Bourbigot, S.; Le Bras, M.; Yvon, J., The facts and hypotheses relating to the phenomenological model of cellulose pyrolysis Interdependence of the steps. *Journal of Analytical and Applied Pyrolysis* **2009,** 84, (1), 1-17.
21. Teixeira, A. R.; Mooney, K. G.; Kruger, J. S.; Williams, C. L.; Suszynski, W. J.; Schmidt, L. D.; Schmidt, D. P.; Dauenhauer, P. J., Aerosol generation by reactive boiling ejection of molten cellulose. *Energy & Environmental Science* **2011,** 4, (10), 4306-4321.
22. Bellais, M.; Davidsson, K. O.; Liliedahl, T.; Sjostrom, K.; Pettersson, J. B. C., Pyrolysis of large wood particles: a study of shrinkage importance in simulations. *Fuel* **2003,** 82, (12), 1541-1548.
23. Ragauskas, A. J.; Williams, C. K.; Davison, B. H.; Britovsek, G.; Cairney, J.; Eckert, C. A.; Frederick, W. J.; Hallett, J. P.; Leak, D. J.; Liotta, C. L.; Mielenz, J. R.; Murphy, R.; Templer, R.; Tschaplinski, T., The path forward for biofuels and biomaterials. *Science* **2006,** 311, (5760), 484-489.
24. Sanderson, K., A field in ferment. *Nature* **2006,** 444, (7120), 673-676.
25. Perlack, R. D.; Wright, L. L.; Turhollow, A. F.; Graham, R. L.; Stokes, B. J.; Erbach, D. C. *Biomass as a Feedstock for a Bioenergy and Bioproducts Industry: The Technical Feasibility of a Billion-Ton Annual Supply*; Oak Ridge National Laboratory, Oak Ridge, TN: 2005.
26. Perlack, R. D.; Stokes, B. J. *U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry*; Oak Ridge National Laboratory, Oak Ridge, TN: 2011.
27. Anex, R. P.; Aden, A.; Kazi, F. K.; Fortman, J.; Swanson, R. M.; Wright, M. M.; Satrio, J. A.; Brown, R. C.; Daugaard, D. E.; Platon, A.; Kothandaraman, G.; Hsu, D. D.; Dutta, A., Techno-economic comparison of biomass-to-transportation fuels via pyrolysis, gasification, and biochemical pathways. *Fuel* **2010,** 89, S29-S35.
28. Dauenhauer, P. J.; Dreyer, B. J.; Degenstein, N. J.; Schmidt, L. D., Millisecond reforming of solid biomass for sustainable fuels. *Angewandte Chemie-International Edition* **2007,** 46, (31), 5864-5867.
29. Sutton, D.; Kelleher, B.; Ross, J. R. H., Review of literature on catalysts for biomass gasification. *Fuel Processing Technology* **2001,** 73, (3), 155-173.
30. Bridgwater, A. V.; Meier, D.; Radlein, D., An overview of fast pyrolysis of biomass. *Organic Geochemistry* **1999,** 30, (12), 1479-1493.
31. Czernik, S.; Bridgwater, A. V., Overview of applications of biomass fast pyrolysis oil. *Energy & Fuels* **2004,** 18, (2), 590-598.
32. Vispute, T. P.; Zhang, H.; Sanna, A.; Xiao, R.; Huber, G. W., Renewable Chemical Commodity Feedstocks from Integrated Catalytic Processing of Pyrolysis Oils. *Science* **2010,** 330, (6008), 1222-1227.
33. Brammer, J. G.; Lauer, M.; Bridgwater, A. V., Opportunities for biomass-derived "bio-oil" in European heat and power markets. *Energy Policy* **2006,** 34, (17).
34. Carlson, T. R.; Vispute, T. R.; Huber, G. W., Green gasoline by catalytic fast pyrolysis of solid biomass derived compounds. *Chemsuschem* **2008,** 1, (5), 397-400.
35. Carlson, T. R.; Tompsett, G. A.; Conner, W. C.; Huber, G. W., Aromatic Production from Catalytic Fast Pyrolysis of Biomass-Derived Feedstocks. *Topics in Catalysis* **2009,** 52, (3), 241-252.
36. Adjaye, J. D.; Bakhshi, N. N., Production of hydrocarbons by catalytic upgrading of a fast pyrolysis bio-oil .1. Conversion over various catalysts. *Fuel Processing Technology* **1995,** 45, (3), 161-183.

37. Zhang, S. P.; Yan, Y. J.; Li, T. C.; Ren, Z. W., Upgrading of liquid fuel from the pyrolysis of biomass. *Bioresource Technology* **2005,** 96, (5), 545-550.
38. Kersten, S.; Garcia-Perez, M., Recent developments in fast pyrolysis of ligno-cellulosic materials *Current opinion in biotechnology* **2013,** 24, (3), 414-20.
39. Agarwal, V.; Dauenhauer, P. J.; Huber, G. W.; Auerbach, S. M., Ab initio dynamics of cellulose pyrolysis: nascent decomposition pathways at 327 and 600°C. *Journal of the American Chemical Society* **2012,** 134, (36).
40. Cho, J.; Chu, S.; Dauenhauer, P. J.; Huber, G. W., Kinetics and reaction chemistry for slow pyrolysis of enzymatic hydrolysis lignin and organosolv extracted lignin derived from maplewood. *Green Chemistry* **2012,** 14, (2), 428-439.
41. Dauenhauer, P. J.; Colby, J. L.; Balonek, C. M.; Suszynski, W. J.; Schmidt, L. D., Reactive boiling of cellulose for integrated catalysis through an intermediate liquid. *Green Chemistry* **2009,** 11, (10), 1555-1561.
42. Broido, A.; Nelson, M. A., Char Yield on Pyrolysis of Cellulose. *Combustion and Flame* **1975,** 24, (2), 263-268.
43. Mettler, M. S.; Mushrif, S. H.; Paulsen, A. D.; Javadekar, A. D.; Vlachos, D. G.; Dauenhauer, P. J., Revealing pyrolysis chemistry for biofuels production: Conversion of cellulose to furans and small oxygenates. *Energy & Environmental Science* **2012,** 5, (1), 5414-5424.
44. Mettler, M. S.; Paulsen, A. D.; Vlachos, D. G.; Dauenhauer, P. J., The chain length effect in pyrolysis: bridging the gap between glucose and cellulose. *Green Chemistry* **2012,** 14, 1284-1288.
45. Paulsen, A. D., Mettler, M. S., Vlachos, D. G., & Dauenhauer, P. J., The Role of Sample Dimension and Temperature in Cellulose Pyrolysis. *Energy & Fuels* **2013**.
46. Mettler, M. S.; Paulsen, A. D.; Vlachos, D. G.; Dauenhauer, P. J., Pyrolytic conversion of cellulose to fuels: levoglucosan deoxygenation via elimination and cyclization within molten biomass. *Energy & Environmental Science* **2012**.
47. Matthews, J. F.; Bergenstrahle, M.; Beckham, G. T.; Himmel, M. E.; Nimlos, M. R.; Brady, J. W.; Crowley, M. F., High-Temperature Behavior of Cellulose I. *Journal of Physical Chemistry B* **2011,** 115, (10), 2155-2166.
48. Agarwal, V.; Huber, G. W.; Conner, W. C., Jr.; Auerbach, S. M., Simulating infrared spectra and hydrogen bonding in cellulose I beta at elevated temperatures. *Journal of Chemical Physics* **2011,** 135, (13).
49. Mayes, H. B.; Broadbelt, L. J., Unraveling the Reactions that Unravel Cellulose. *Journal of Physical Chemistry A* **2012,** 116, (26), 7098-7106.
50. Seshadri, V.; Westmoreland, P. R., Concerted Reactions and Mechanism of Glucose Pyrolysis and Implications for Cellulose Kinetics. *Journal of Physical Chemistry A* **2012,** 116, (49), 11997-12013.
51. Colby, J. L.; Dauenhauer, P. J.; Michael, B. C.; Bhan, A.; Schmidt, L. D., Improved utilization of biomass-derived carbon by co-processing with hydrogen-rich feedstocks in millisecond reactors. *Green Chemistry* **2010,** 12, (3), 378-380.
52. Varhegyi, G.; Jakab, E.; Antal, M. J., Is the Broido-Shafizadeh Model for Cellulose Pyrolysis True. *Energy & Fuels* **1994,** 8, (6), 1345-1352.
53. Kim, D. Y.; Nishiyama, Y.; Wada, M.; Kuga, S. S., *Wood and Cellulose: Building Blocks for Chemicals and Fuels and Advanced Materials*. In Syracuse, New York, 2000.

54. Cho, J.; Davis, J. M.; Huber, G. W., The Intrinsic Kinetics and Heats of Reactions for Cellulose Pyrolysis and Char Formation. *Chemsuschem* **2010,** 3, (10), 1162-1165.
55. Soukupova, J.; Rock, B. N.; Albrechtova, J., Spectral characteristics of lignin and soluble phenolics in the near infrared - a comparative study. *International Journal of Remote Sensing* **2002,** 23, (15), 3039-3055.
56. Sanderson, M. A.; Agblevor, F.; Collins, M.; Johnson, D. K., Compositional analysis of biomass feedstocks by near infrared reflectance spectroscopy. *Biomass & Bioenergy* **1996,** 11, (5), 365-370.
57. Kelley, S. S.; Rials, T. G.; Snell, R.; Groom, L. H.; Sluiter, A., Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood. *Wood Science and Technology* **2004,** 38, (4), 257-276.
58. Axrup, L.; Markides, K.; Nilsson, T., Using miniature diode array NIR spectrometers for analyzing wood chips and bark samples in motion. *Journal of Chemometrics* **2000,** 14, (5-6), 561-572.
59. Bilbao, R.; Millera, A.; Murillo, M. B., Temperature profiles and weight-loss in the thermal-decomposition of large spherical wood particlesS. *Industrial & Engineering Chemistry Research* **1993,** 32, (9), 1811-1817.
60. Malkavaara, P.; Alen, R., A spectroscopic method for determining lignin content of softwood and hardwood kraft pulps. *Chemometrics and Intelligent Laboratory Systems* **1998,** 44, 287-292.
61. Forestville Builders & Supply. http://forestvillebuilders.com/ (July 30, 2014),
62. Rasband, W. S. ImageJ. http://imagej.nih.gov/ij/ (July 30, 2014),
63. Jackson, M.; Yang, D.; Parkin, R., Analysis of wood surface waviness with a two-image photometric stereo method. *Proceedings of the Institution of Mechanical Engineers Part I-Journal of Systems and Control Engineering* **2007,** 221, (I8), 1091-1099.
64. Mehrabian, R.; Zahirovic, S.; Scharler, R.; Obernberger, I.; Kleditzsch, S.; Wirtz, S.; Scherer, V.; Lu, H.; Baxter, L. L., A CFD model for thermal conversion of thermally thick biomass particles. *Fuel Processing Technology* **2012,** 95, 96-108.
65. Products, L. F.; Products, S. F., *Wood handbook: wood as an engineering material*. Forest Products Society: Madison, Wis., 2011.
66. Ward, S. M.; Braslaw, J., Experimental weight loss kinetics of wood pyrolysis under vacuum. *Combustion and Flame* **1985,** 61, (3), 261-269.
67. Branca, C.; Albano, A.; Di Blasi, C., Critical evaluation of global mechanisms of wood devolatilization. *Thermochimica Acta* **2005,** 429, (2), 133-141.
68. Lu, H. Experimental and Modeling Investigations of Biomass Particle Combustion. Brigham Young University, 2006.
69. Bryden, K. M.; Ragland, K. W.; Rutland, C. J., Modeling thermally thick pyrolysis of wood. *Biomass & Bioenergy* **2002,** 22, (1), 41-53.
70. Davidsson, K. O.; Pettersson, J. B. C., Birch wood particle shrinkage during rapid pyrolysis. *Fuel* **2002,** 81, (3), 263-270.
71. Antal, M. J.; Varhegyi, G., Cellulose Pyrolysis Kinetics - The Current State Knowledge. *Industrial & Engineering Chemistry Research* **1995,** 34, (3), 703-717.
72. DiBlasi, C., Heat, momentum and mass transport through a shrinking biomass particle exposed to thermal radiation. *Chemical Engineering Science* **1996,** 51, (7), 1121-1132.
73. Vanholme, R.; Demedts, B.; Morreel, K.; Ralph, J.; Boerjan, W., Lignin Biosynthesis and Structure. *Plant Physiology* **2010,** 153, (3), 895-905.

74. Gronli, M. G.; Melaaen, M. C., Mathematical model for wood pyrolysis - Comparison of experimental measurements with model predictions. *Energy & Fuels* **2000,** 14, (4), 791-800.
75. Bassilakis, R.; Carangelo, R. M.; Wojtowicz, M. A., TG-FTIR analysis of biomass pyrolysis. *Fuel* **2001,** 80, (12), 1765-1786.
76. Azadi, P.; Inderwildi, O. R.; Farnood, R.; King, D. A., Liquid fuels, hydrogen and chemicals from lignin: A critical review. *Renewable & Sustainable Energy Reviews* **2013,** 21, 506-523.
77. Suhas; Carrott, P. J. M.; Carrott, M., Lignin - from natural adsorbent to activated carbon: A review. *Bioresource Technology* **2007,** 98, (12), 2301-2312.
78. Vishtal, A.; Kraslawski, A., Challenges in industrial applications of technical lignins. *Bioresources* **2011,** 6, (3), 3547-3568.
79. Goyal, H. B.; Seal, D.; Saxena, R. C., Bio-fuels from thermochemical conversion of renewable resources: A review. *Renewable & Sustainable Energy Reviews* **2008,** 12, (2), 504-517.
80. Zhou, X. W.; Nolte, M. W.; Mayes, H. B.; Shanks, B. H.; Broadbelt, L. J., Experimental and Mechanistic Modeling of Fast Pyrolysis of Neat Glucose-Based Carbohydrates. 1. Experiments and Development of a Detailed Mechanistic Model. *Industrial & Engineering Chemistry Research* **2014,** 53, (34), 13274-13289.
81. Hatakeyama, H.; Hatakeyama, T., Lignin Structure, Properties, and Applications. In *Biopolymers: Lignin, Proteins, Bioactive Nanocomposites*, Abe, A.; Dusek, K.; Kobayashi, S., Eds. Springer-Verlag Berlin: Berlin, 2010; Vol. 232, pp 1-63.
82. Freudenberg, K., Biosynthesis and constitution of lignin. *Nature* **1959,** 183, (4669), 1152-1155.
83. Zhao, Y.; Truhlar, D. G., The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theoretical Chemistry Accounts* **2008,** 120, (1-3), 215-241.
84. Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A., A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. *Journal of Chemical Physics* **1999,** 110, (6), 2822-2827.
85. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M.; A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima; T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers; E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi; J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross; B., J.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann; E., R.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski; W., J.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas; Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.

86. Pfaendtner, J.; Broadbelt, L. J., Mechanistic modeling of lubricant degradation. 1. Structure-reactivity relationships for free-radical oxidation. *Industrial & Engineering Chemistry Research* **2008,** 47, (9), 2886-2896.
87. Blowers, P.; Masel, R. I., An extension of the Marcus equation for atom transfer reactions. *Journal of Physical Chemistry A* **1999,** 103, (35), 7047-7054.
88. Stewart, W. E.; Caracotsios, M.; Sorensen, J. P. *Double Precision Differential/Algebraic Sensitivity Analysis Code (DDASAC)*, 1997; Madison, WI, 1997.
89. Hough, B. *ligpy: a kinetic model for lignin pyrolysis*, 2016. http://dx.doi.org/10.5281/zenodo.53202
90. Jakab, E.; Faix, O.; Till, F., Thermal decomposition of milled wood lignins studied by thermogravimetry mass spectrometry. *Journal of Analytical and Applied Pyrolysis* **1997,** 40-1, 171-186.
91. Ben, H. X.; Ragauskas, A. J., NMR Characterization of Pyrolysis Oils from Kraft Lignin. *Energy & Fuels* **2011,** 25, (5), 2322-2332.
92. Patwardhan, P. R.; Brown, R. C.; Shanks, B. H., Understanding the Fast Pyrolysis of Lignin. *Chemsuschem* **2011,** 4, (11), 1629-1636.
93. Jiang, G. Z.; Nowakowski, D. J.; Bridgwater, A. V., Effect of the Temperature on the Composition of Lignin Pyrolysis Products. *Energy & Fuels* **2010,** 24, 4470-4475.
94. Lou, R.; Wu, S. B.; Lyu, G. J., Quantified monophenols in the bio-oil derived from lignin fast pyrolysis. *Journal of Analytical and Applied Pyrolysis* **2015,** 111, 27-32.
95. Ranzi, E.; Cuoci, A.; Faravelli, T.; Frassoldati, A.; Migliavacca, G.; Pierucci, S.; Sommariva, S., Chemical Kinetics of Biomass Pyrolysis. *Energy & Fuels* **2008,** 22, (6), 4292-4300.
96. Anca-Couce, A., Reaction mechanisms and multi-scale modelling of lignocellulosic biomass pyrolysis. *Progress in Energy and Combustion Science* **2016,** 53, 41-79.
97. Krumm, C.; Pfaendtner, J.; Dauenhauer, P. J., Millisecond Pulsed Films Unify the Mechanisms of Cellulose Fragmentation. *Chemistry of Materials* **2016,** 28, (9), 3108-3114.
98. Saltelli, A.; Annoni, P.; Azzini, I.; Campolongo, F.; Ratto, M.; Tarantola, S., Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications* **2010,** 181, (2), 259-270.
99. Saltelli, A., Global sensitivity analysis the primer. In Wiley: Chichester [u.a.], 2008.
100. Saltelli, A., Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications* **2002,** 145, (2), 280-297.
101. Sobol, I. M., Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* **2001,** 55, (1-3), 271-280.
102. Herman, J.; Hadka, D.; xantares; Fernando; bernardoct; Mutel, C.; Usher, W. *SALib: New documentation, doc strings and installation requirements*, 0.7; 2015. http://dx.doi.org/10.5281/zenodo.46450
103. Hough, B.; Fu, C.; Paliwal, S. *savvy: visualize high dimensionality sensitivity analysis data*, 2016. http://dx.doi.org/10.5281/zenodo.53099
104. Paulsen, A. D.; Hough, B. R.; Williams, C. L.; Teixeira, A. R.; Schwartz, D. T.; Pfaendtner, J.; Dauenhauer, P. J., Fast Pyrolysis of Wood for Biofuels: Spatiotemporally Resolved Diffuse Reflectance In situ Spectroscopy of Particles. *Chemsuschem* **2014,** 7, (3), 765-776.
105. Anca-Couce, A.; Zobel, N., Numerical analysis of a biomass pyrolysis particle model: Solution method optimized for the coupling to reactor models. *Fuel* **2012,** 97, 80-88.

106. Papadikis, K.; Gu, S.; Bridgwater, A. V.; Gerhauser, H., Application of CFD to model fast pyrolysis of biomass. *Fuel Processing Technology* **2009,** 90, (4), 504-512.
107. Mehrabian, R.; Zahirovic, S.; Scharler, R.; Obernberger, I.; Kleditzsch, S.; Wirtz, S.; Scherer, V.; Lu, H.; Baxter, L., A CFD model for thermal conversion of thermally thick biomass particles. *Fuel Processing Technology* **2012,** 95, 96-108.
108. Hough, B. R.; Schwartz, D. T.; Pfaendtner, J., Detailed kinetic modeling of lignin pyrolysis for process optimization. *Industrial & Engineering Chemistry Research* **2016**.
109. Reyniers, P. A.; Schietekat, C. M.; Van Cauwenberge, D. J.; Vandewalle, L. A.; Van Geem, K. M.; Marin, G. B., Necessity and Feasibility of 3D Simulations of Steam Cracking Reactors. *Industrial & Engineering Chemistry Research* **2015,** 54, (49), 12270-12282.
110. Asadi-Eydivand, M.; Solati-Hashjin, M.; Farzadi, A.; Abu Osman, N. A., Artificial neural network approach to estimate the composition of chemically synthesized biphasic calcium phosphate powders. *Ceramics International* **2014,** 40, (8), 12439-12448.
111. Molga, E. J., Neural network approach to support modelling of chemical reactors: problems, resolutions, criteria of application. *Chemical Engineering and Processing* **2003,** 42, (8-9), 675-695.
112. Rocabruno-Valdes, C. I.; Ramirez-Verduzco, L. F.; Hernandez, J. A., Artificial neural network models to predict density, dynamic viscosity, and cetane number of biodiesel. *Fuel* **2015,** 147, 9-17.
113. Kayala, M. A.; Baldi, P., ReactionPredictor: Prediction of Complex Chemical Reactions at the Mechanistic Level Using Machine Learning. *Journal of Chemical Information and Modeling* **2012,** 52, (10), 2526-2540.
114. Beck, D. A. C.; Carothers, J. M.; Subramanian, V. R.; Pfaendtner, J., Data Science: Accelerating Innovation and Discovery in Chemical Engineering. *Aiche Journal* **2016,** 62, (5), 1402-1416.
115. Conesa, J. A.; Caballero, J. A.; Reyes-Labarta, J. A., Artificial neural network for modelling thermal decompositions. *Journal of Analytical and Applied Pyrolysis* **2004,** 71, (1), 343-352.
116. Molga, E. J.; van Woezik, B. A. A.; Westerterp, K. R., Neural networks for modelling of chemical reaction systems with complex kinetics: oxidation of 2-octanol with nitric acid. *Chemical Engineering and Processing* **2000,** 39, (4), 323-334.
117. Bishop, C. M., *Pattern recognition and machine learning*. Springer: New York, 2006; p xx, 738 pages.
118. Haykin, S. S., *Neural Networks and Learning Machines*. Prentice Hall: 2009.
119. LeCun, Y.; Bottou, L.; Orr, G. B.; Muller, K. R., Efficient backprop. *Neural Networks: Tricks of the Trade* **1998,** 1524, 9-50.
120. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E., Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011,** 12, 2825-2830.
121. Duchi, J.; Hazan, E.; Singer, Y., Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *Journal of Machine Learning Research* **2011,** 12, 2121-2159.
122. Glorot, X.; Bengio, Y. In *Understanding the difficulty of training deep feedforward neural networks*, International conference on artificial intelligence and statistics, 2010, 2010; 2010; pp 249-256.
123. Nouri, D. *nolearn*, 0.7; Github, 2015.

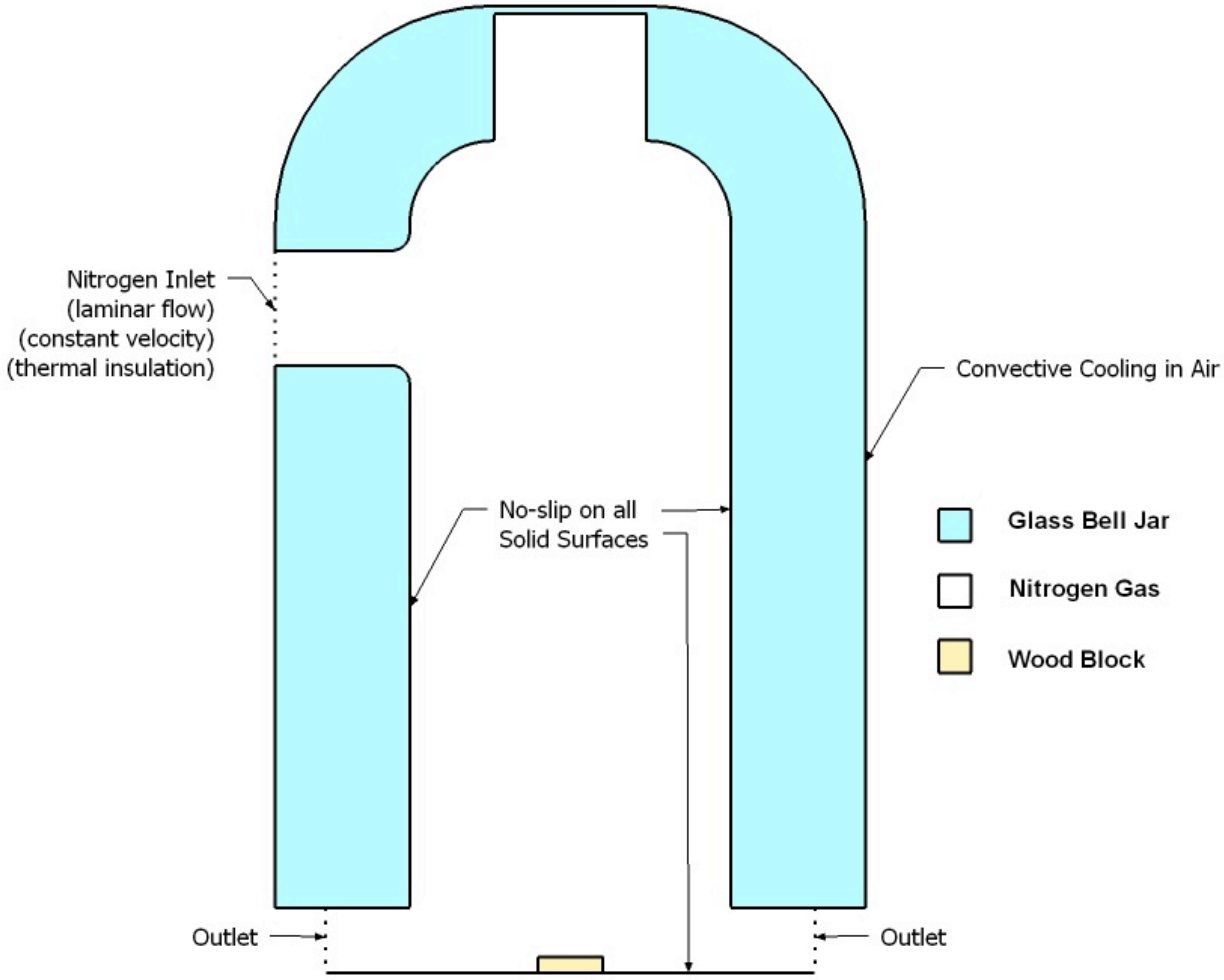
124. Hough, B. *lignet: artificial neural networks for modeling lignin pyrolysis*, 0.1; Zenodo: 2016. http://dx.doi.org/10.5281/zenodo.56733
125. Hanberry, B. B., Compositional changes in selected forest ecosystems of the western United States. *Applied Geography* **2014,** 52, 90-98.
126. Stephens, S. L.; Ruth, L. W., Federal forest-fire policy in the United States. *Ecological Applications* **2005,** 15, (2), 532-542.
127. Noss, R. F.; Franklin, J. F.; Baker, W. L.; Schoennagel, T.; Moyle, P. B., Managing fire-prone forests in the western United States. *Frontiers in Ecology and the Environment* **2006,** 4, (9), 481-487.
128. Hessburg, P. F.; Agee, J. K.; Franklin, J. F., Dry forests and wildland fires of the inland Northwest USA: Contrasting the landscape ecology of the pre-settlement and modem eras. *Forest Ecology and Management* **2005,** 211, (1-2), 117-139.
129. Hagmann, R. K.; Franklin, J. F.; Johnson, K. N., Historical conditions in mixed-conifer forests on the eastern slopes of the northern Oregon Cascade Range, USA. *Forest Ecology and Management* **2014,** 330, 158-170.
130. van Mantgem, P. J.; Stephenson, N. L.; Byrne, J. C.; Daniels, L. D.; Franklin, J. F.; Fule, P. Z.; Harmon, M. E.; Larson, A. J.; Smith, J. M.; Taylor, A. H.; Veblen, T. T., Widespread Increase of Tree Mortality Rates in the Western United States. *Science* **2009,** 323, (5913), 521-524.
131. Dennison, P. E.; Brewer, S. C.; Arnold, J. D.; Moritz, M. A., Large wildfire trends in the western United States, 1984-2011. *Geophysical Research Letters* **2014,** 41, (8), 2928-2933.
132. Stephens, S. L.; Agee, J. K.; Fule, P. Z.; North, M. P.; Romme, W. H.; Swetnam, T. W.; Turner, M. G., Managing Forests and Fire in Changing Climates. *Science* **2013,** 342, (6154), 41-42.
133. van Mantgem, P. J.; Nesmith, J. C. B.; Keifer, M.; Knapp, E. E.; Flint, A.; Flint, L., Climatic stress increases forest fire severity across the western United States. *Ecology Letters* **2013,** 16, (9), 1151-1156.
134. Westerling, A. L.; Hidalgo, H. G.; Cayan, D. R.; Swetnam, T. W., Warming and earlier spring increase western US forest wildfire activity. *Science* **2006,** 313, (5789), 940-943.
135. Wu, T.; Kim, Y. S., Pricing ecosystem resilience in frequent-fire ponderosa pine forests. *Forest Policy and Economics* **2013,** 27, 8-12.
136. Krieger, D. J., *The economic value of forest ecosystem services: a review*. Wilderness Society Washington, DC, USA: 2001.
137. Hand, M. S.; Gebert, K. M.; Liang, J.; Calkin, D. E.; Thompson, M. P.; Zhou, M., *Economics of wildfire management : the development and application of suppression expenditure models*. 2014.
138. Forest Service *Fiscal Year 2015 Budget Justification*; United States Department of Agriculture, Forest Service: Washington, D.C., 2014; p 547.
139. USDOI *Wildland fire management program benefit-cost analysis: A review of relevant literature*; USDOI Office of Policy Analysis: 2012.
140. Reinhardt, E. D.; Keane, R. E.; Calkin, D. E.; Cohen, J. D., Objectives and considerations for wildland fuel treatment in forested ecosystems of the interior western United States. *Forest Ecology and Management* **2008,** 256, (12), 1997-2006.
141. Forest Service *A strategic assessment of forest biomass and fuel reduction treatments in Western States*; United States Department of Agriculture, Forest Service, Rocky Mountain Research Station: Fort Collins, CO, 2005; p 17.

142. Huang, C. H.; Finkral, A.; Sorensen, C.; Kolb, T., Toward full economic valuation of forest fuels-reduction treatments. *Journal of Environmental Management* **2013,** 130, 221-231.
143. Calkin, D.; Gebert, K., Modeling fuel treatment costs on forest service lands in the western United States. *Western Journal of Applied Forestry* **2006,** 21, (4), 217-221.
144. Prestemon, J. P.; Abt, K. L.; Barbour, R. J., Quantifying the net economic benefits of mechanical wildfire hazard treatments on timberlands of the western United States. *Forest Policy and Economics* **2012,** 21, 44-53.
145. Long, J. N., Emulating natural disturbance regimes as a basis for forest management: A North American view. *Forest Ecology and Management* **2009,** 257, (9), 1868-1873.
146. Brown, R. T.; Agee, J. K.; Franklin, J. F., Forest restoration and fire: Principles in the context of place. *Conservation Biology* **2004,** 18, (4), 903-912.
147. Nackley, L. L.; Lieu, V. H.; Garcia, B. B.; Richardson, J. J.; Isaac, E.; Spies, K.; Rigdon, S.; Schwartz, D. T., Bioenergy that supports ecological restoration. *Frontiers in Ecology and the Environment* **2013,** 11, (10), 535-540.
148. Rockwell, D.; Auld, F.; Ball, S.; Barce, G.; Becker, R.; Beyer, J.; DelSordo, D.; Dupuis, D.; Gobeille, J.; Hansen, B.; Harwood, T.; Makepeace, S.; McDonald, T.; Tanner, T.; Thomas, P.; Trosper, B.; White, G., Flathead Indian Reservation Forest Management Plan. In Confederated Salish and Kootenai Tribes: Pablo, MT, 2000; p 303.
149. He, L. X.; English, B. C.; Ugarte, D. G. D.; Hodges, D. G., Woody biomass potential for energy feedstock in United States. *Journal of Forest Economics* **2014,** 20, (2), 174-191.
150. Perlack, R. D.; Stokes, B. J., U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry. In Energy, U. S. D. o., Ed. Oak Ridge National Laboratory, Oak Ridge, TN, 2011; p 227p.
151. Ordinance No. 95: Cultural Resource Protection Ordinance. In Reservation, T. C. o. t. C. S. K. T. o. t. F., Ed. 1995; Vol. Ordinance No. 95, p 40.
152. Morgan, T. A. *An Assessment of Forest-based Woody Biomass Supply and Use in Montana*; University of Montana Bureau of Business and Economic Research: Missoula, MT, 29 April, 2009, 2009; p 22.
153. Abbas, D.; Current, D.; Ryans, M.; Taff, S.; Hoganson, H.; Brooks, K. N., Harvesting forest biomass for energy - An alternative to conventional fuel treatments: Trials in the Superior National Forest, USA. *Biomass & Bioenergy* **2011,** 35, (11), 4557-4564.
154. Delfino, R. J.; Brummel, S.; Wu, J.; Stern, H.; Ostro, B.; Lipsett, M.; Winer, A.; Street, D. H.; Zhang, L.; Tjoa, T.; Gillen, D. L., The relationship of respiratory and cardiovascular hospital admissions to the southern California wildfires of 2003. *Occupational and Environmental Medicine* **2009,** 66, (3), 189-197.
155. Henderson, S. B.; Brauer, M.; MacNab, Y. C.; Kennedy, S. M., Three Measures of Forest Fire Smoke Exposure and Their Associations with Respiratory and Cardiovascular Health Outcomes in a Population-Based Cohort. *Environmental Health Perspectives* **2011,** 119, (9), 1266-1271.
156. Kochi, I.; Donovan, G. H.; Champ, P. A.; Loomis, J. B., The economic cost of adverse health effects from wildfire-smoke exposure: a review. *International Journal of Wildland Fire* **2010,** 19, (7), 803-817.
157. Henderson, S. B.; Johnston, F. H., Measures of forest fire smoke exposure and their associations with respiratory health outcomes. *Current Opinion in Allergy and Clinical Immunology* **2012,** 12, (3), 221-227.

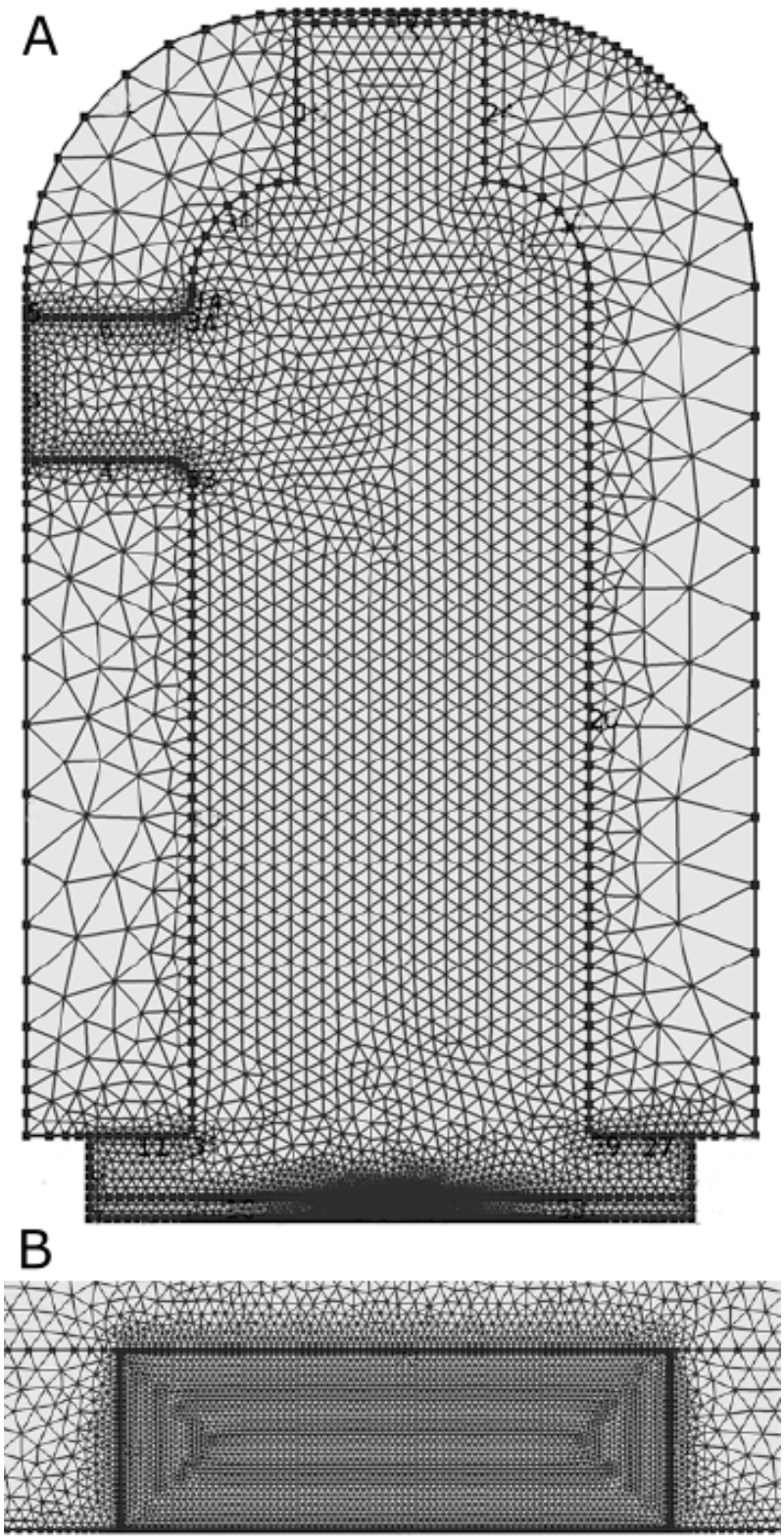
158. Springsteen, B.; Christofk, T.; Eubanks, S.; Mason, T.; Clavin, C.; Storey, B., Emission Reductions from Woody Biomass Waste for Energy as an Alternative to Open Burning. *Journal of the Air & Waste Management Association* **2011,** 61, (1), 63-68.
159. McKinley, D. C.; Ryan, M. G.; Birdsey, R. A.; Giardina, C. P.; Harmon, M. E.; Heath, L. S.; Houghton, R. A.; Jackson, R. B.; Morrison, J. F.; Murray, B. C.; Pataki, D. E.; Skog, K. E., A synthesis of current knowledge on forests and carbon storage in the United States. *Ecological Applications* **2011,** 21, (6), 1902-1924.
160. Lippke, B.; Gustafson, R.; Venditti, R.; Steele, P.; Volk, T. A.; Oneil, E.; Johnson, L.; Puettmann, M. E.; Skog, K., Comparing Life-Cycle Carbon and Energy Impacts for Biofuel, Wood Product, and Forest Management Alternatives. *Forest Products Journal* **2012,** 62, (4), 247-257.
161. Sunde, K.; Brekke, A.; Solberg, B., Environmental impacts and costs of woody Biomass-to-Liquid (BTL) production and use - A review. *Forest Policy and Economics* **2011,** 13, (8), 591-602.
162. Lippke, B.; Puettmann, M. E.; Johnson, L.; Gustafson, R.; Venditti, R.; Steele, P.; Katers, J. F.; Taylor, A.; Volk, T. A.; Oneil, E.; Skog, K.; Budsberg, E.; Daystar, J.; Caputo, J., Carbon Emission Reduction Impacts from Alternative Biofuels. *Forest Products Journal* **2012,** 62, (4), 296-304.
163. Richardson, J. J.; Spies, K. A.; Rigdon, S.; York, S.; Lieu, V.; Nackley, L.; Garcia, B. B.; Cawston, R.; Schwartz, D. T., Uncertainty in biomass supply estimates: Lessons from a Yakama Nation case study. *Biomass & Bioenergy* **2011,** 35, (8), 3698-3707.
164. Skog, K.; Barbour, J.; Buford, M.; Dykstra, D.; Lebow, P.; Miles, P.; Perlack, B.; Stokes, B., Forest-Based Biomass Supply Curves for the United States. *Journal of Sustainable Forestry* **2013,** 32, (1-2), 14-27.
165. Howard, J. O. *Ratios for Estimating Logging Residue in the Pacific Northwest*; PNW-288; USDA Forest Service: Portland, OR, 1981; p 32.
166. Perez-Garcia, J.; Oneil, E.; Hansen, T.; Mason, T.; McCarter, J.; Rogers, L.; Cooke, A.; Comnick, J.; McLaughlin, M. *Washington Forest Biomass Supply Assessment*; Washington Department of Natural Resources: Olympia, WA, 2012; p 183.
167. Rorstad, P. K.; Tromborg, E.; Bergseng, E.; Solberg, B., Combining GIS and Forest Modelling in Estimating Regional Supply of Harvest Residues in Norway. *Silva Fennica* **2010,** 44, (3), 435-451.
168. Cozzi, M.; Di Napoli, F.; Viccaro, M.; Romano, S., Use of Forest Residues for Building Forest Biomass Supply Chains: Technical and Economic Analysis of the Production Process. *Forests* **2013,** 4, (4), 1121-1140.
169. Alam, M. B.; Pulkki, R.; Shahi, C., Woody biomass availability for bioenergy production using forest depletion spatial data in northwestern Ontario. *Canadian Journal of Forest Research* **2012,** 42, (3), 506-516.
170. Bouriaud, O.; Stefan, G.; Flocea, M., Predictive models of forest logging residues in Romanian spruce and beech forests. *Biomass & Bioenergy* **2013,** 54, 59-66.

APPENDIX A

SUPPLEMENTARY MATERIAL FOR CHAPTER 2



Supporting Figure A.1: Model geometry showing the three regimes in the reaction-transport model. The geometry of the model and the locations where boundary conditions are applied is shown here (the locations of additional boundary conditions are described in Chapter 2).



Supporting Figure A.2: Computational mesh used to solve the model system. (A) Free triangular mesh with 14,936 elements. (B) Zoom of the wood block region of the mesh

Supporting Table A.1: Parameter variations for the model sensitivity analysis. All values are for the 600°C heating block at t=0. X^i is the char formation mass ratio of component i.

Parameter	Units	Base case	Lower variation	Upper variation	Notes
ρ_{wood}	[kg m^{-3}]	573	300	728	Exploring the range of literature values for each of these parameters
ρ_{char}	[kg m^{-3}]	200	100	437	
$K_{y,wood}$	[W m^{-1} K^{-1}]	0.14	0.087	0.323	
$K_{y,char}$	[W m^{-1} K^{-1}]	0.071	0.052	0.105	
$C_{p,wood}$	[J kg^{-1} K^{-1}]	1772.9	1260	3480	
$C_{p,char}$	[J kg^{-1} K^{-1}]	1091.2	980	2760	
ΔH_2	[kJ kg^{-1}]	255	204	306	+/- 20% of literature value
ΔH_3	[kJ kg^{-1}]	-20	-16	-24	
X^{Cell}	[-]	0.35	0.28	0.42	
X^{Hemi}	[-]	0.60	0.48	0.72	
X^{Lig}	[-]	0.75	0.60	0.90	
$k_{1,cell}$	[s^{-1}]	2.8E19[s^{-1}]*exp(-242.4[kJ mol^{-1}]/(R*T))	Multiply by 0.1	Multiply by 10	Changed by 1 order of magnitude up/down
$k_{1,hemi}$	[s^{-1}]	2.1E16[s^{-1}]*exp(-186.7[kJ mol^{-1}]/(R*T))			
$k_{1,lig}$	[s^{-1}]	9.6E8[s^{-1}]*exp(-107.6[kJ mol^{-1}]/(R*T))			
$k_{2,cell}$	[s^{-1}]	3.28E14[s^{-1}]*exp(-196.5[kJ mol^{-1}]/(R*T))			

$k_{2,hemi}$	[s^{-1}]	8.75E15[s^{-1}]*exp(-202.4[kJ mol^{-1}]/(R*T))			
$k_{2,lig}$	[s^{-1}]	1.5E9[s^{-1}]*exp(-143.8[kJ mol^{-1}]/(R*T))			
$k_{3,cell}$	[s^{-1}]	1.3E10[s^{-1}]*exp(-150.5[kJ mol^{-1}]/(R*T))			
$k_{3,hemi}$	[s^{-1}]	2.6E11[s^{-1}]*exp(-145.7[kJ mol^{-1}]/(R*T))			
$k_{3,lig}$	[s^{-1}]	7.7E6[s^{-1}]*exp(-111.4[kJ mol^{-1}]/(R*T))			

Supporting Table A.2: Time dependent solver settings. Settings used to solve the model system

Relative tolerance: 0.002
Absolute tolerance (global, scaled): 0.005
Absolute tolerances were specified for the following variables:

Variable	**Absolute tolerance**
Nitrogen velocity	0.07 [m s^{-1}]
Temperature	2 [K]
Pressure	1E4 [Pa]
Masses of all species	0.1 [µg]

Time stepping:
Method: BDF with Free steps
Max step size: 0.005 s

Solver: PARDISO

APPENDIX B

SUPPLEMENTARY MATERIAL FOR CHAPTER 3

We have published the software programs that were used to generate and solve our kinetic model on github, and archived them at zenodo.org. The python code for building and solving the model and generating the results presented in this paper, including changes we made from Faravelli et al.'s original scheme, is archived on zenodo and can be accessed there or on github:

> ligpy model code: http://dx.doi.org/10.5281/zenodo.53202
> on github: https://github.com/houghb/ligpy

We recommend getting the code from github as that will be the most up to date, and it links directly to the documentation with instructions on using the program. The copy at zenodo is static (it is an archive of Release 1.0 on github), so will always reflect the exact code used in this paper.

We have also written a python package to visualize high dimensionality sensitivity analysis results, and included with that package the complete sensitivity analysis from the ligpy model described in this paper so readers can explore the sensitivity analysis results themselves. This package is available here:

> savvy code: http://dx.doi.org/10.5281/zenodo.53099
> on github: https://github.com/houghb/savvy

Supporting Table B.1: Kinetic scheme for lignin pyrolysis with changes from Faravelli et al.

- The names of groups linked to the solid residue start with P, where radical species start with R (radicals linked to the degrading polymer matrix start with PR).
- Values in the Rxn # column are zero-indexed to correspond to the python programs. Reaction numbers for H-abstraction reactions are not listed in this table because the reactions are built iteratively from the values in this table (see the ligpy program). **Orange** reaction numbers indicate our sensitivity analysis found these reactions do not contribute to any output variance for the output measures we examined.
- Rate constants are expressed as $k(T) = A\ T^n \exp(-E/RT)$
- Reactions are written as follows:

  stoic-coeff _ species , stoic-coeff _ species2 , …

  where stoic-coeff is negative for reactants and positive for products
- We added reactions with shaded backgrounds to the Faravelli et al. scheme
- Species in **red** were changed from the Faravelli model
- Other changes from the Faravelli model are in **blue** (however, reactions and species that were removed completely are not shown).

Rxn #	Reaction	A [s^{-1}] or [L mol^{-1} s^{-1}]	n	E [J mol^{-1}]
	Initiation Reactions			
104	-1_PLIGH,1_LIGH	1.00E+13	0	154812
105	-1_PLIGC,1_LIGC	1.00E+13	0	168458
106	-1_PLIGO,1_LIGO	1.00E+13	0	154812
0	-1_PLIGH,1_PRLIGH	1.00E+13	0	163254
1	-1_LIGH,1_C3H6,1_OH,1_RLIGM2A	1.00E+13	0	163254
2	-1_LIGM2,1_RPHENOXM2,1_RADIOM2	1.00E+13	0	163254
3	-1_PLIGM2,1_RPHENOXM2,1_PRADIOM2	1.00E+13	0	163254
4	-1_LIG,1_RPHENOX,1_RADIO	1.00E+13	0	184184
5	-1_PLIG,1_RPHENOX,1_PRADIO	1.00E+13	0	188370
6	-1_PADIOM2,1_PRADIOM2	1.00E+13	0	171626
7	-1_PADIO,1_PRADIO	1.00E+13	0	179998
	Decomposition Reactions			
8	-1_PKETM2,1_PRKETM2	1.00E+13	0	167440
9	-1_PRKETM2,1_KETDM2,1_OH	1.00E+13	0	121394
10	-1_RPHENOXM2,1_CO,0.5_C10H2M4,0.5_H2	4.00E+10	0	209300
11	-1_RPHENOX,1_CO,0.5_C10H2,1.5_H2	4.00E+10	0	209300
12	-1_RLIGH,1_RLIGM2A,1_ALD3	1.00E+13	0	133952

13	-1_PRLIGH2,1_PRLIGM2A,1_ALD3	1.00E+13	0	133952
14	-1_RADIOM2,1_RMGUAI,1_C3H6O2	1.00E+13	0	133952
15	-1_RLIGM2A,1_KETM2,1_RPHENOXM2	5.00E+12	0	133952
16	-1_PRLIGM2A,1_PKETM2,1_RPHENOXM2	5.00E+12	0	133952
17	-1_RLIGM2B,1_RMGUAI,1_PFET3M2	1.00E+13	0	163254
18	-1_PRFET3M2,1_RPHENOXM2,1_C3H4O2	1.00E+13	0	133952
19	-1_RADIO,1_RPHENOL,1_C3H6O2	1.00E+13	0	163254
20	-1_RLIGA,1_KET,1_RPHENOX	1.00E+13	0	138138
21	-1_RLIGB,1_RPHENOL,1_PFET3	1.00E+13	0	163254
22	-1_PRFET3,1_RPHENOX,1_C3H4O2	1.00E+13	0	138138
23	-1_RADIOM2,1_OH,1_SYNAPYL	3.00E+11	0	104650
24	-1_RKETM2,1_OH,1_KETDM2	3.00E+11	0	104650
25	-1_RADIO,1_OH,1_COUMARYL	3.00E+11	0	104650
26	-1_RKET,1_OH,1_KETD	3.00E+11	0	113022
27	-1_RC3H7O2,1_RCH3O,1_CH3CHO	1.00E+13	0	129766
28	-1_C10H2M4,1_RCH3O,0.5_C10H2M4,1_RPHENOX	1.00E+13	0	196742
29	-1_C10H2M2,1_RCH3O,0.5_C10H2,1_RPHENOX	1.00E+13	0	196742
	Molecular Reactions			
30	-1_PLIGC,1_PLIG,1_CH2CO	1.00E+08	0	121394
31	-1_PLIGO,1_PLIGM2,1_CO2	1.00E+09	0	108836
108	-1_LIGC,1_LIG,1_CH2CO	1.00E+08	0	121394
107	-1_LIGO,1_LIGM2,1_CO2	1.00E+09	0	108836
	Substitutive Addition Reactions			
32	-1_ADIOM2,-1_RPHENOXM2,1_C10H2M4,0.5_H2,1_PH2,2_PCOH,1_RC3H7O2	1.00E+09	0	121394
33	-1_KETM2,-1_RPHENOXM2,1_C10H2M4,0.5_H2,1_PH2,2_PCOH,1_RC3H5O2	1.00E+09	0	121394
34	-1_KETDM2,-1_RPHENOXM2,1_C10H2M4,0.5_H2,1_PH2,2_PCOH,1_RC3H3O	1.00E+09	0	121394
35	-1_SYNAPYL,-1_RPHENOXM2,1_C10H2M4,1_H2,1.5_PH2,2_PCOH,1_RC3H3O	1.00E+09	0	121394
36	-1_ADIOM2,-1_RPHENOX,0.5_C10H2M4,0.5_C10H2_1.5_H2,1_PH2,1_PCOS,1_PCOH,1_RC3H7O2	1.00E+09	0	121394
37	-1_KETM2,-1_RPHENOX,0.5_C10H2M4,0.5_C10H2,0.5_H2,2_PH2,2_PCOS,1_RC3H5O2	1.00E+09	0	121394
38	-1_KETDM2,-1_RPHENOX,0.5_C10H2M4,0.5_C10H2,0.5_H2,2_PH2,1_PCOS,1_PCOH,1_RC3H3O	1.00E+09	0	121394

39	-1_SYNAPYL,-1_RPHENOX,0.5_C10H2,0.5_C10H2M4,1.5_H2,2_PH2,1_PCOH,1_PCOS,1_RC3H3O	1.00E+09	0	121394
40	-1_ADIO,-1_RPHENOX,1_C10H2,1.5_H2,2_PH2,2_PCOS,1_RC3H7O2	1.00E+09	0	121394
41	-1_KET,-1_RPHENOX,1_C10H2,1.5_H2,2_PCOS,1_RC3H5O2,2_PH2	1.00E+09	0	121394
42	-1_KETD,-1_RPHENOX,1_C10H2,1.5_H2,2_PCOS,1_RC3H3O,2_PH2	1.00E+09	0	121394
43	-1_COUMARYL,-1_RPHENOX,1_C10H2,2.5_H2,2_PCOS,1_RC3H3O,2_PH2	1.00E+09	0	121394
44	-1_ADIO,-1_RPHENOXM2,0.5_C10H2,0.5_C10H2M4,1.5_H2,1_PH2,1_PCOS,1_PCOH,1_RC3H7O2	1.00E+09	0	121394
45	-1_KET,-1_RPHENOXM2,0.5_C10H2,0.5_C10H2M4,1.5_H2,1_PH2,1_PCOS,1_PCOH,1_RC3H5O2	1.00E+09	0	121394
46	-1_KETD,-1_RPHENOXM2,0.5_C10H2,0.5_C10H2M4,1.5_H2,1_PH2,1_PCOS,1_PCOH,1_RC3H3O	1.00E+09	0	121394
47	-1_COUMARYL,-1_RPHENOXM2,0.5_C10H2,0.5_C10H2M4,1.5_H2,2_PH2,1_PCOS,1_PCOH,1_RC3H3O	1.00E+09	0	121394
48	-1_C10H2M4,-1_RPHENOXM2,0.5_C10H2M2,1_C10H2M4,1_RCH3O,0.5_H2,1_PCOH	1.00E+09	0	115115
49	-1_C10H2M2,-1_RPHENOXM2,0.5_C10H2M2,0.5_C10H2,1_RCH3O,0.5_C10H2M4,0.5_H2,1_PCOH	1.00E+09	0	115115
50	-1_C10H2M4,-1_RPHENOX,0.5_C10H2M2,0.5_C10H2M4,1_RCH3O,0.5_C10H2,1.5_H2,1_PCOS	1.00E+09	0	117208
51	-1_C10H2M2,-1_RPHENOX,0.5_C10H2M2,1_RCH3O,1_C10H2,1.5_H2,1_PCOS	1.00E+09	0	117208
	Recombination Reactions			
52	-1_RCH3O,-1_RPHENOX,1_RCH3,1_CO2,0.5_C10H2,1.5_H2	1.00E+08	0	54418
53	-1_RCH3O,-1_RPHENOXM2,1_RCH3,1_CO2,0.5_C10H2M4,0.5_H2	1.00E+08	0	54418
54	-1_RPHENOXM2,-1_RCH3,1_PCOH,0.5_C10H2M4,0.5_PH2,1_PCH3	1.00E+08	0	48139
55	-1_RPHENOX,-1_RCH3,1_PCOH,0.5_C10H2,1.5_PH2,1_PCH3	1.00E+08	0	48139
56	-1_RCH3O,-1_RCH3,1_ETOH	1.00E+08	0	12558
57	-2_RCH3O,1_CH3CHO,1_H2O	1.00E+08	0	12558
58	-1_RCH3O,-1_PRADIOM2,1_PCH2OH,1_PADIOM2	1.00E+08	0	52325
59	-1_RCH3O,-1_PRKETM2,1_PCH2OH,1_PKETM2	1.00E+08	0	52325
60	-2_RADIOM2,2_PCH2OH,2_PCHP2,2_PCHOHP,2_PH2,1_C10H2M4,2_PCOH	3.16E+07	0	83720
61	-2_RLIGM2B,0.2_CHAR,2_PCH2OH,2_PCHOHP,3_PH2,2_C10H2M4,4_PCOH	3.16E+07	0	83720

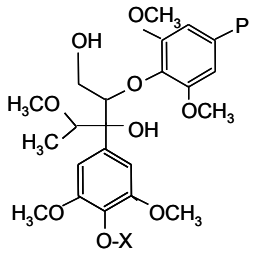
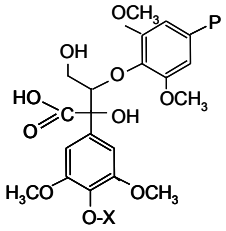
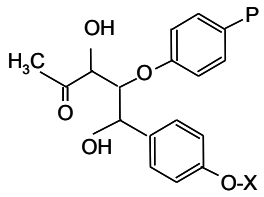
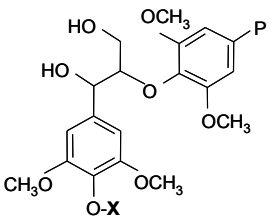
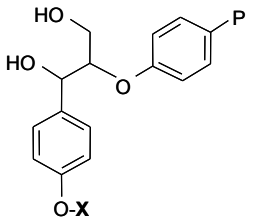
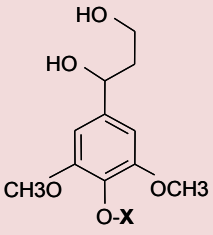
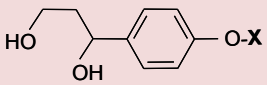
62	-2_RLIGM2A,2_PCOHP2,2_H2O,2_PC2H2,4_PCOH,3_PH2,2_C10H2M4	3.16E+07	0	83720
63	-2_RMGUAI,2_PH2,1_C10H2M4,2_PCOH	3.16E+07	0	83720
64	-2_RKETM2,2_PCHP2,2_PCH2OH,2_PH2,1_C10H2M4,4_PCOH	3.16E+07	0	83720
65	-1_PRFET3M2,-1_PRFET3M2,2_PCOH,1_C10H2M4,2_PCHP2,2_PCHO,2_PCHOHP,1_PH2	3.16E+07	0	83720
66	-2_RC3H7O2,2_H2O,2_PC2H2,2_PCHOHP,1_PH2	1.00E+08	0	12558
67	-2_RC3H5O2,2_H2O,2_PC2H2,1_PH2,2_PCOH	1.00E+08	0	12558
68	-2_RC3H3O,2_PC2H2,1_PH2,2_PCOH	1.00E+08	0	12558
69	-1_OH,-1_RCH3,1_CH3OH	1.00E+08	0	12558
70	-2_RCH3,1_C2H6	1.00E+08	0	12558
71	-1_RCH3,-1_PRADIOM2,1_PCH3,1_PADIOM2	1.00E+08	0	48139
72	-1_RCH3,-1_PRKETM2,1_PCH3,1_PKETM2	1.00E+08	0	48139
73	-1_RPHENOX,-1_RLIGB,3_PCOS,1_PCOHP2,1_H2O,1_PC2H2,1.5_C10H2,3_PH2,2_H2	3.16E+07	0	83720
74	-2_RADIO,2_PCH2OH,2_PCHP2,2_PCHOHP,1_PCOS,1_PCOH,1_C10H2,1_H2,3_PH2	3.16E+07	0	92092
75	-1_RLIGB,-1_RLIGB,1_PCOS,0.2_CHAR,2_PCH2OH,2_PCHOHP,2_C10H2,3_PCOS,5_PH2,2_H2	3.16E+07	0	83720
76	-2_RLIGA,2_PCOHP2,2_H2O,2_PC2H2,2_C10H2,4_PCOS,5_PH2,2_H2	3.16E+07	0	83720
77	-2_RKET,3_PCOH,2_PCHP2,2_PCH2OH,1_PCOS,1_C10H2,2_H2,2_PH2	3.16E+07	0	83720
78	-1_PRFET3,-1_PRFET3,2_PCHP2,2_PCHO,2_PCHOHP,2_PCOS,1_C10H2,1_H2,2_PH2	3.16E+07	0	83720
79	-1_RLIGH,-1_RLIGH,2_PCOHP2,4_H2O,2_PC2H2,4_PCOH,2_C10H2M4,1_H2,2_PCH3,2_PCHP2,2_PCH2P	3.16E+07	0	83720
80	-1_RPHENOX,-1_RPHENOL,2_PCOS,1_C10H2,1.5_H2,2_PH2	5.00E+07	0	83720
81	-1_RPHENOX,-1_RC3H3O,1_PCOS,0.5_C10H2,1_H2,1_PH2,1_PC2H2,1_PCOH	1.00E+08	0	48139
82	-1_RPHENOX,-1_CHAR,1_PCOS,0.5_C10H2,1.5_H2,1_CHAR	3.00E+07	0	104650
83	-1_RPHENOXM2,-1_CHAR,1_PCOS,0.5_C10H2M4,0.5_H2,1_CHAR	3.00E+07	0	104650
109	-2_PRADIO,1_PCOH,1_PCOS,2_PCH2OH,2_PCHP2,2_PCHOHP,1_C10H2,2_PH2,1_H2	3.16E+07	0	92092
110	-2_PRADIOM2,2_PCOH,2_PCH2OH,2_PCHP2,2_PCHOHP,1_PH2,1_C10H2M4	3.16E+07	0	83720
	Volatilization Reactions			
84	-1_ADIOM2,1_VADIOM2	1	1	62790
85	-1_KETM2,1_VKETM2	1	1	62790

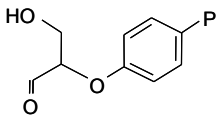
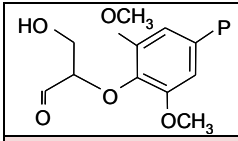
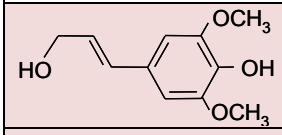
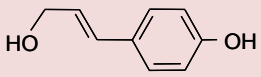
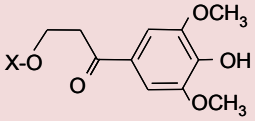
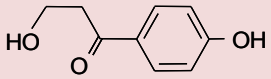
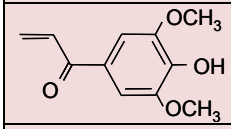
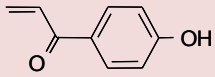
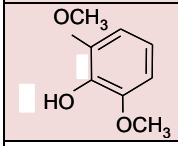
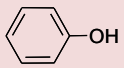
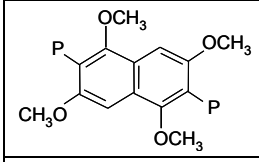
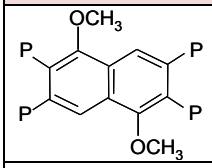
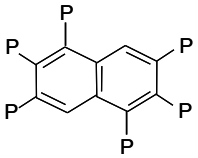
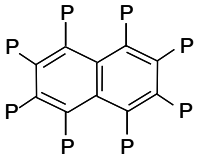
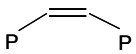
86	-1_KETDM2,1_VKETDM2	1	1	58604
87	-1_SYNAPYL,1_VSYNAPYL	1	1	62790
88	-1_MGUAI,1_VMGUAI	1	1	66976
89	-1_COUMARYL,1_VCOUMARYL	1	1	83720
90	-1_ADIO,1_VADIO	1	1	66976
91	-1_KET,1_VKET	1	1	66976
92	-1_KETD,1_VKETD	1	1	83720
93	-1_PHENOL,1_VPHENOL	1	1	58604
	Char Devolatilization Reactions			
94	-1_PCOS,1_CO	2.00E+08	0	209300
95	-1_PCOH,1_CO	1.00E+07	0	138138
96	-1_PH2,1_H2	1.00E+10	0	209300
97	-1_PC2H2,1_H2,0.2_CHAR	5.00E+08	0	205114
98	-1_PCH2OH,1_OH,1_PCH2P	5.00E+08	0	0
99	-1_PCHOHP,1_OH,0.5_PC2H2	5.00E+08	0	0
100	-1_PCOHP2,1_OH,0.1_CHAR	5.00E+08	0	0
101	-1_PCH3,1_RCH3	1.00E+13	0	125580
102	-1_PCH2P,1_PCHP2,0.5_H2	5.00E+08	0	209300
103	-1_PCHP2,0.1_CHAR,0.5_H2	5.00E+08	0	209300
111	-1_PCHO,1_CO	1.00E+07	0	138138
	Hydrogen Abstraction Reactions			
	-1_R,-1_LIGH,1_RH,1_RLIGH	2	0	4186
	-1_R,-1_PLIGH,1_RH,1_PRLIGH2	2	0	4186
	-1_R,-1_PLIGM2,1_RH,1_PRLIGM2A	2	0	4186
	-1_R,-1_LIGM2,1_RH,1_RLIGM2A	1	0	4186
	-1_R,-1_LIGM2,1_RH,1_RLIGM2B	1	0	4186
	-1_R,-1_PFET3M2,1_RH,1_PRFET3M2	2	0	4186
	-1_R,-1_ADIOM2,1_RH,1_RADIOM2	1	0	4186
	-1_R,-1_KETM2,1_RH,1_RKETM2	1	0	4186
	-1_R,-1_C10H2,1_RH,0.5_C10H2,0.5_CHAR	1	0	-20930
	-1_R,-1_LIG,1_RH,1_RLIGA	1	0	4186
	-1_R,-1_LIG,1_RH,1_RLIGB	1	0	4186
	-1_R,-1_PFET3,1_RH,1_PRFET3	2	0	4186
	-1_R,-1_ADIO,1_RH,1_RADIO	1	0	4186
	-1_R,-1_KET,1_RH,1_RKET	1	0	4186

	Kinetic Paramaters for Abstracting Radical (each radical in this list takes the place of R above)			
	OH	1.00E+9.5	0	14651
	RCH3	1.00E+8.5	0	48139
	RCH3O	1.00E+8.5	0	33069
	RKET, RPHENOL	1.00E+8.5	0	54418
	RLIGA, RLIGB	1.00E+08	0	56511
	RC3H7O2	1.00E+08	0	62790
	RC3H5O2	1.00E+08	0	71162
	RC3H3O	1.00E+08	0	83720
	PRADIO	1.00E+08	0	46046
	PRLIGH, PRADIOM2	1.00E+8.5	0	41860
	RADIOM2, RADIO, RMGUAI, RLIGM2A, RLIGM2B, PRFET3M2, PRKETM2, PRFET3	1.00E+08	0	54418

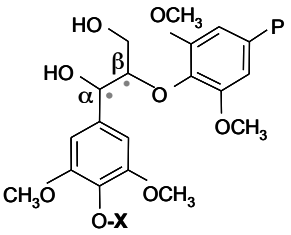
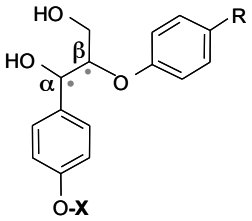
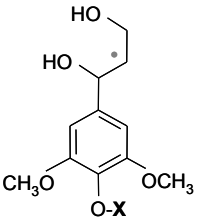
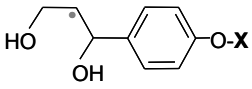
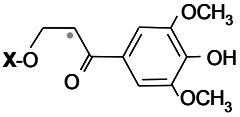
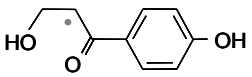
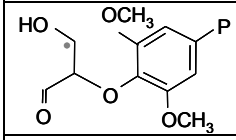
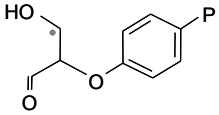
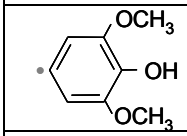
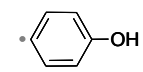
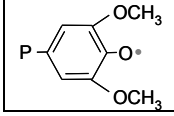
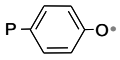
Supporting Table B.2: Molecules in the kinetic scheme

- This table is modified from Faravelli et al.'s Supplementary Information
- Heavy tar components are highlighted in red.

Structure	Name, Composition And MW	Structure	Name, Composition and MW
	X= P PLIGH $C_{22}H_{28}O_9$ 436 X=H LIGH $C_{22}H_{29}O_9$ 437		X=P PLIGO $C_{20}H_{22}O_{10}$ 422 X=H LIGO $C_{20}H_{23}O_{10}$ 423
	X=P PLIGC $C_{17}H_{16}O_5$ 300 X=H LIGC $C_{17}H_{17}O_5$ 301		X=P PLIGM2 $C_{19}H_{22}O_8$ 378 X=H LIGM2 $C_{19}H_{23}O_8$ 379
	X=P PLIG C15H14O4 258 X=H LIG C15H15O4 259		X=P PADIOM2 $C_{11}H_{15}O_5$ 227 X=H ADIOM2 $C_{11}H_{16}O_5$ 228
	X=P PADIO $C_9H_{11}O_3$ 167 X=H ADIO $C_9H_{12}O_3$ 168		

	PFET3 $C_9H_9O_3$ 165			PFET3M2 $C_{11}H_{13}O_5$ 225
	Synapyl $C_{11}H_{14}O_4$ 210			Coumaryl $C_9H_{10}O_2$ 150
	X=P PKETM2 $C_{11}H_{13}O_5$ 225 X=H KETM2 $C_{11}H_{14}O_5$ 226			X=P PKET $C_9H_9O_3$ 165 X=H KET $C_9H_{10}O_3$ 166
	KETDM2 $C_{11}H_{12}O_4$ 208			KETD $C_9H_8O_2$ 148
	MGUAI $C_8H_{10}O_3$ 154			Phenol C_6H_6O 94
	C10H2M4 $C_{14}H_{14}O_4$ 244			C10H2M2 $C_{12}H_8O_2$ 182
	C10H2 $C_{10}H_2$ 122			Char C_{10} 120
	PC2H2 C_2H_2 26			PCH2P CH_2 14
	PCOS CO 28			PCOH CO 28
	PCH2OH CH_3O 31			PCHO CHO 29
	PCHOHP CH_2O 30			

Radical Species

Structure	Name, Composition And MW		Structure	Name, Composition and MW
	X=P PRLIGM2A, PRLIGM2B $C_{19}H_{21}O_8$ 377 X=H RLIGM2A, RLIGM2B $C_{19}H_{22}O_8$ 378			X=P PRLIGA, PRLIGB $C_{15}H_{13}O_4$ 257 X=H RLIGA, RLIGB $C_{15}H_{14}O_4$ 258
	X=P PRADIOM2 $C_{11}H_{14}O_5$ 226 X=H RADIOM2 $C_{11}H_{15}O_5$ 227			X=P PRADIO $C_9H_{10}O_3$ 166 X=H RADIO $C_9H_{11}O_3$ 167
	X=P PRKETM2 $C_{11}H_{12}O_5$ 224 X=H RKETM2 $C_{11}H_{13}O_5$ 225			RKET $C_9H_9O_3$ 165
	PRFET3M2 $C_{11}H_{12}O_5$ 224			PRFET3 $C_9H_8O_3$ 164
	Rguai $C_8H_9O_3$ 153			Rphenol C_6H_5O 91
	RphenoxM2 $C_8H_8O_3$ 152			Rphenox C_6H_4O 90

Supporting Table B.3: Rate constants with sensitivity indices of exactly zero. These rate constants have sensitivity indices (total and first order) of exactly zero for all 30 output measures we examined.

Parameter	Reaction
k60	-2_RADIOM2,2_PCH2OH,2_PCHP2,2_PCHOHP,2_PH2,1_C10H2M4,2_PCOH
k64	-2_RKETM2,2_PCHP2,2_PCH2OH,2_PH2,1_C10H2M4,4_PCOH
k72	-1_RCH3,-1_PRKETM2,1_PCH3,1_PKETM2
k74	-2_RADIO,2_PCH2OH,2_PCHP2,2_PCHOHP,1_PCOS,1_PCOH,1_C10H2,1_H2,3_PH2
k76	-2_RLIGA,2_PCOHP2,2_H2O,2_PC2H2,2_C10H2,4_PCOS,5_PH2,2_H2
k98	-1_PCH2OH,1_OH,1_PCH2P
k99	-1_PCHOHP,1_OH,0.5_PC2H2
k100	-1_PCOHP2,1_OH,0.1_CHAR
k113	-1_PRFET3,-1_LIGH,1_PFET3,1_RLIGH
k120	-1_RCH3,-1_LIGH,1_CH4,1_RLIGH
k122	-1_RKET,-1_LIGH,1_KET,1_RLIGH
k123	-1_PRADIO,-1_LIGH,1_PADIO,1_RLIGH
k125	-1_RLIGB,-1_LIGH,1_LIG,1_RLIGH
k126	-1_RLIGA,-1_LIGH,1_LIG,1_RLIGH
k131	-1_RPHENOL,-1_LIGH,1_PHENOL,1_RLIGH
k132	-1_RADIO,-1_LIGH,1_ADIO,1_RLIGH
k133	-1_RC3H5O2,-1_PLIGH,1_C3H6O2,1_PRLIGH2
k134	-1_PRFET3,-1_PLIGH,1_PFET3,1_PRLIGH2
k135	-1_RC3H7O2,-1_PLIGH,1_C3H8O2,1_PRLIGH2
k137	-1_PRFET3M2,-1_PLIGH,1_PFET3M2,1_PRLIGH2
k143	-1_RKET,-1_PLIGH,1_KET,1_PRLIGH2
k144	-1_PRADIO,-1_PLIGH,1_PADIO,1_PRLIGH2
k145	-1_RC3H3O,-1_PLIGH,1_C3H4O,1_PRLIGH2
k147	-1_RLIGA,-1_PLIGH,1_LIG,1_PRLIGH2
k151	-1_RCH3O,-1_PLIGH,1_CH3OH,1_PRLIGH2
k152	-1_RPHENOL,-1_PLIGH,1_PHENOL,1_PRLIGH2
k153	-1_RADIO,-1_PLIGH,1_ADIO,1_PRLIGH2
k154	-1_RC3H5O2,-1_PLIGM2,1_C3H6O2,1_PRLIGM2A
k155	-1_PRFET3,-1_PLIGM2,1_PFET3,1_PRLIGM2A
k182	-1_RLIGM2A,-1_LIGM2,1_LIGM2,1_RLIGM2A
k195	-1_RADIO,-1_LIGM2,1_ADIO,1_RLIGM2A

k202	-1_RLIGM2B,-1_LIGM2,1_LIGM2,1_RLIGM2B
k216	-1_RPHENOL,-1_LIGM2,1_PHENOL,1_RLIGM2B
k217	-1_RC3H5O2,-1_PFET3M2,1_C3H6O2,1_PRFET3M2
k218	-1_PRFET3,-1_PFET3M2,1_PFET3,1_PRFET3M2
k219	-1_RC3H7O2,-1_PFET3M2,1_C3H8O2,1_PRFET3M2
k220	-1_RADIOM2,-1_PFET3M2,1_ADIOM2,1_PRFET3M2
k221	-1_PRFET3M2,-1_PFET3M2,1_PFET3M2,1_PRFET3M2
k222	-1_PRLIGH,-1_PFET3M2,1_PLIGH,1_PRFET3M2
k231	-1_RLIGA,-1_PFET3M2,1_LIG,1_PRFET3M2
k237	-1_RADIO,-1_PFET3M2,1_ADIO,1_PRFET3M2
k241	-1_RC3H7O2,-1_ADIOM2,1_C3H8O2,1_RADIOM2
k246	-1_RLIGM2A,-1_ADIOM2,1_LIGM2,1_RADIOM2
k249	-1_PRADIO,-1_ADIOM2,1_PADIO,1_RADIOM2
k258	-1_RPHENOL,-1_ADIOM2,1_PHENOL,1_RADIOM2
k260	-1_PRFET3,-1_KETM2,1_PFET3,1_RKETM2
k279	-1_RADIO,-1_KETM2,1_ADIO,1_RKETM2
k282	-1_RC3H7O2,-1_C10H2,1_C3H8O2,0.5_C10H2,0.5_CHAR
k283	-1_RADIOM2,-1_C10H2,1_ADIOM2,0.5_C10H2,0.5_CHAR
k285	-1_PRLIGH,-1_C10H2,1_PLIGH,0.5_C10H2,0.5_CHAR
k306	-1_PRLIGH,-1_LIG,1_PLIGH,1_RLIGA
k315	-1_RLIGA,-1_LIG,1_LIG,1_RLIGA
k335	-1_RLIGB,-1_LIG,1_LIG,1_RLIGB
k344	-1_PRFET3,-1_PFET3,1_PFET3,1_PRFET3
k346	-1_RADIOM2,-1_PFET3,1_ADIOM2,1_PRFET3
k367	-1_RADIOM2,-1_ADIO,1_ADIOM2,1_RADIO
k368	-1_PRFET3M2,-1_ADIO,1_PFET3M2,1_RADIO
k369	-1_PRLIGH,-1_ADIO,1_PLIGH,1_RADIO
k370	-1_RLIGM2B,-1_ADIO,1_LIGM2,1_RADIO
k372	-1_RCH3,-1_ADIO,1_CH4,1_RADIO
k380	-1_RMGUAI,-1_ADIO,1_MGUAI,1_RADIO
k384	-1_RADIO,-1_ADIO,1_ADIO,1_RADIO
k390	-1_PRLIGH,-1_KET,1_PLIGH,1_RKET
k395	-1_PRKETM2,-1_KET,1_PKETM2,1_RKET

APPENDIX C

SUPPLEMENTARY MATERIAL FOR CHAPTER 4

We have published our Python code written to set up, train, and analyze all the neural networks and the decision tree on github, and archived it at zenodo.org. It can be accessed at either of the following links:

on zenodo: http://dx.doi.org/10.5281/zenodo.56733
on github: https://github.com/houghb/lignet

Supporting Table C.1: Output measures from the detailed kinetic model of lignin pyrolysis that we trained our machine learning models to predict.

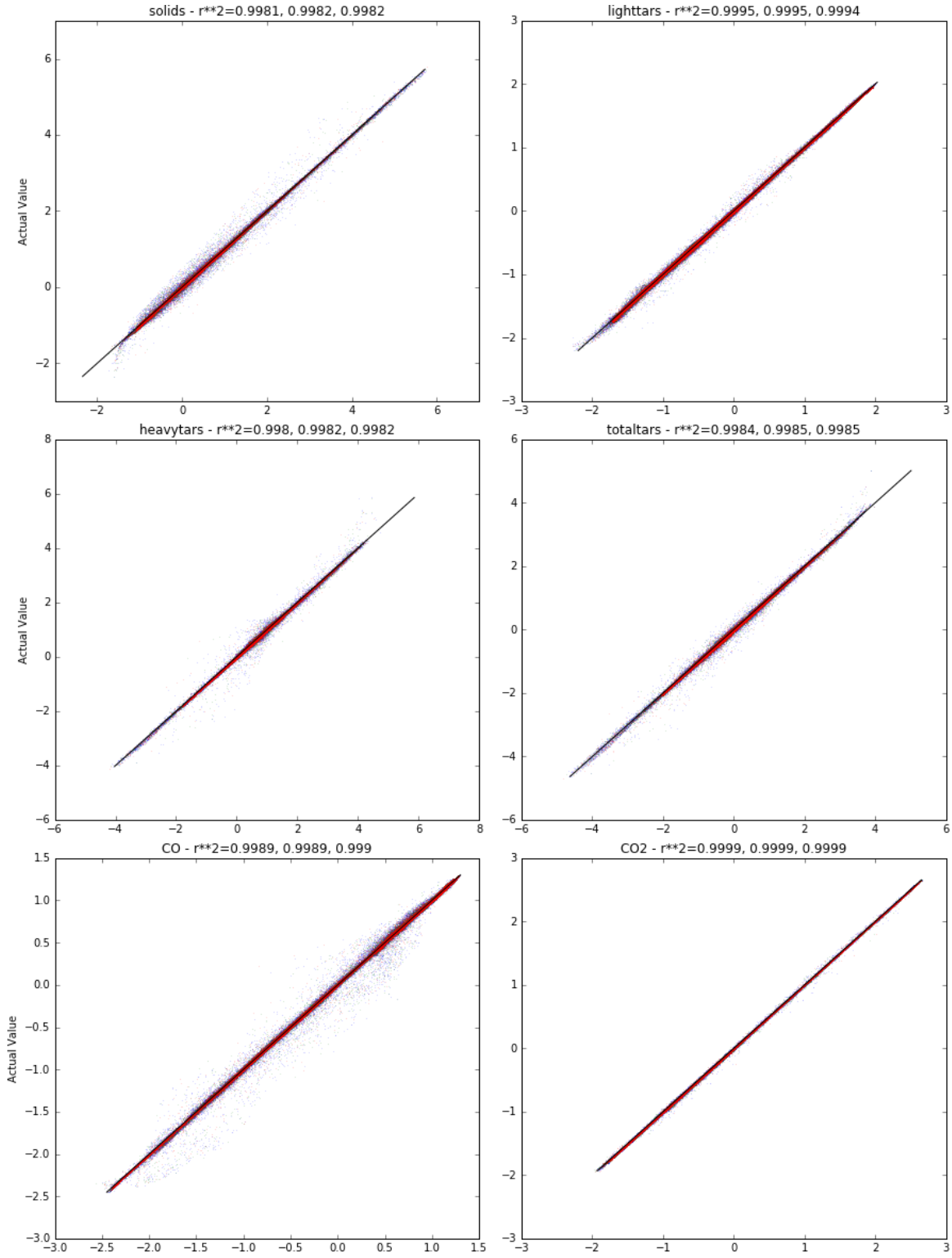
Label	Description
solids	wt% of solid species
lighttars	wt% of light tars
heavytars	wt% of heavy tars
totaltars	wt% of total tars
CO	wt% of CO
CO2	wt% of CO_2
othergases	wt% of gases that are not CO or CO_2
totalgases	wt% of all gases
H2O	wt% of water
H-C-ratio	Ratio of hydrogen to carbon
Moisture	Moisture content of tar (wt%)
tot_C--O	Distribution of functional groups (shown as a % of total carbon) in tars ('tot_'), heavy tars ('heavy_'), and light tars ('light_')
tot_aromatic_C-O	
tot_aromatic_C-C	
tot_aromatic_C-H	
tot_aliphatic_C-O	
tot_aromatic_methoxyl	
tot_aliphatic_C-C	
heavy_C--O	
heavy_aromatic_C-O	
heavy_aromatic_C-C	
heavy_aromatic_C-H	
heavy_aliphatic_C-O	
heavy_aromatic_methoxyl	
heavy_aliphatic_C-C	
light_C--O	
light_aromatic_C-O	
light_aromatic_C-H	
light_aliphatic_C-O	
light_aliphatic_C-C	

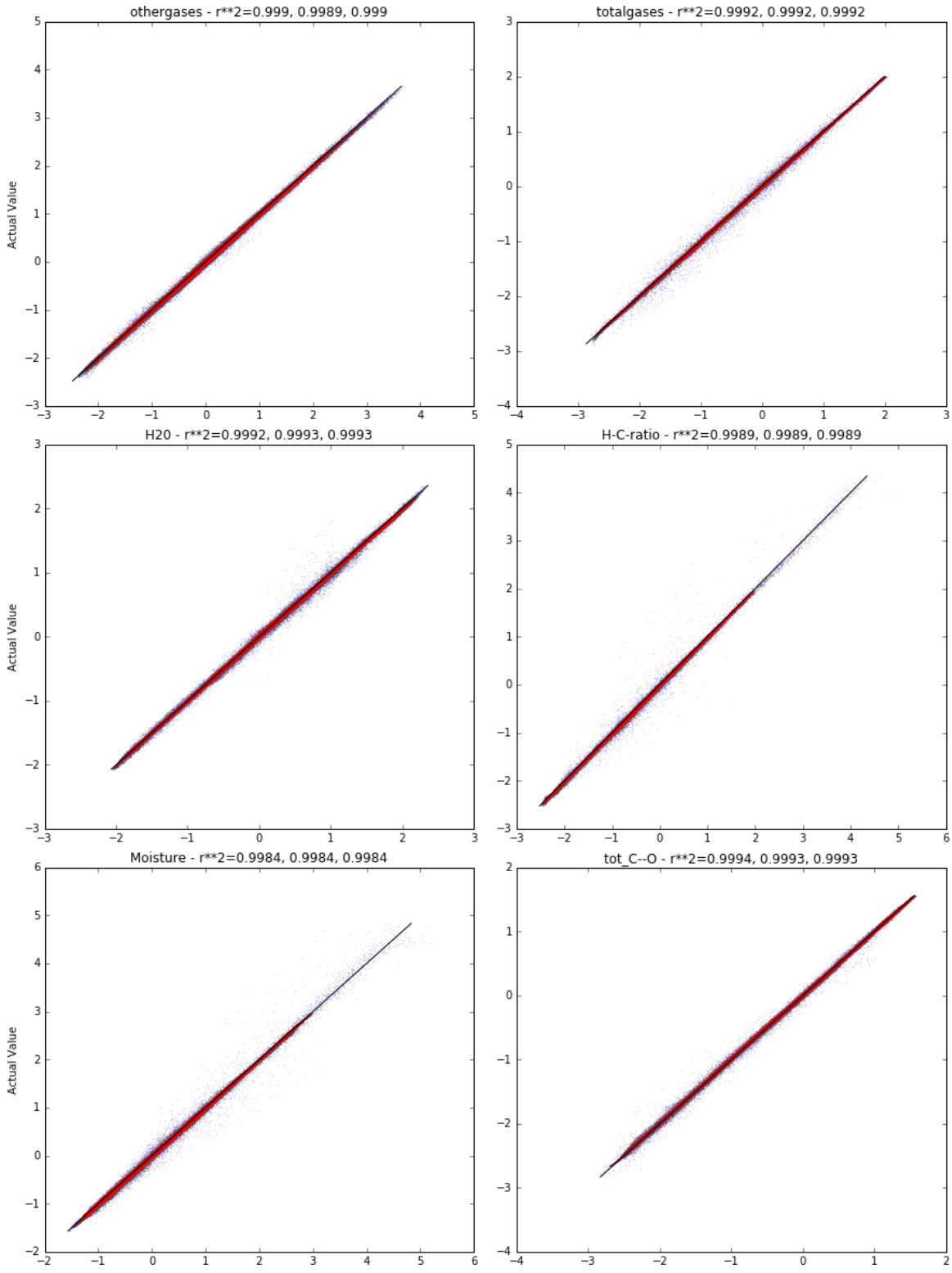
Supporting Table C.2: Network architectures for neural nets and decision tree. All neural nets have four nodes in the input layer. Input and hidden layers also have one bias node that is not included in the numbers shown here.

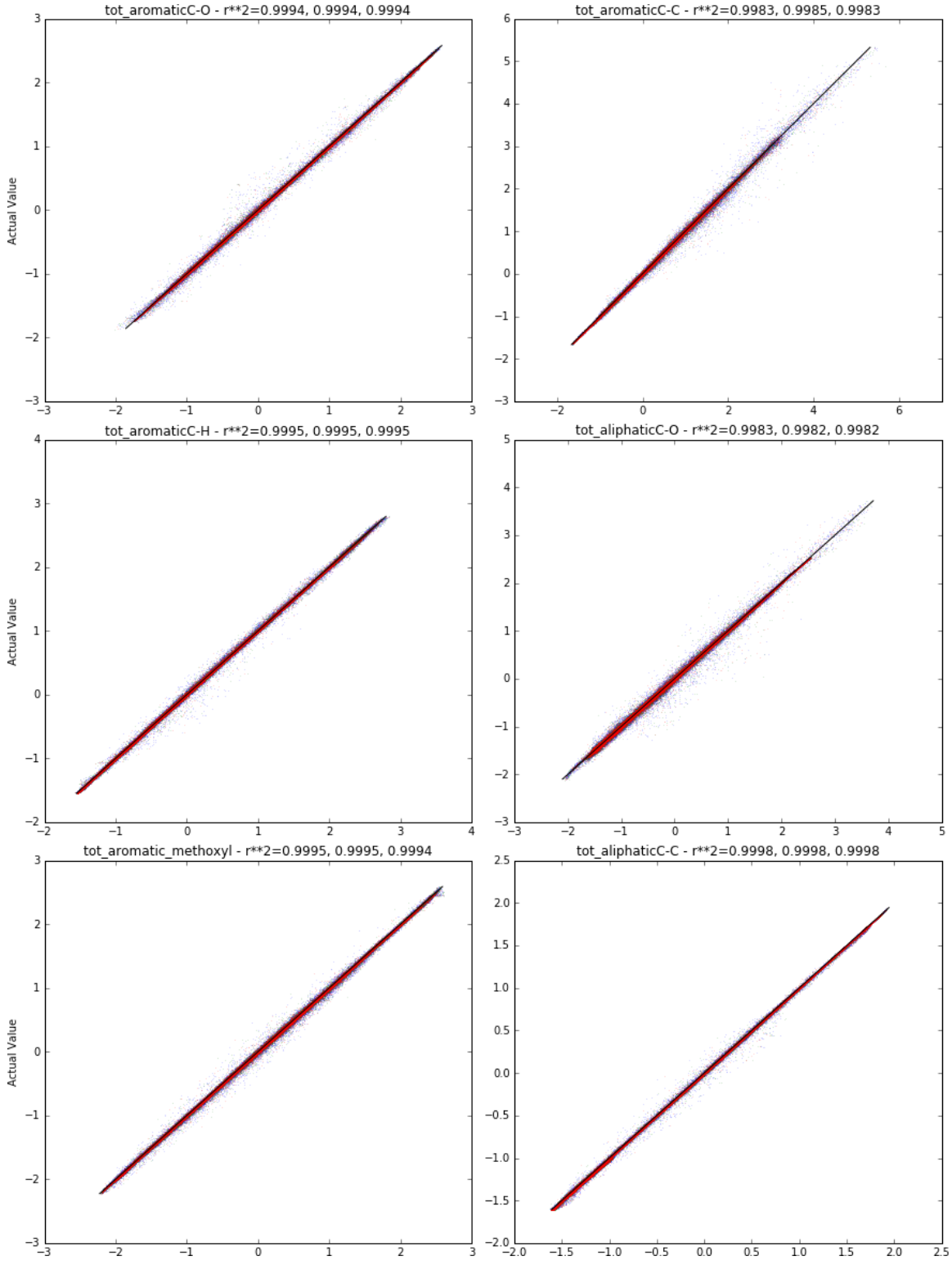
Neural Net Name	Number of nodes in:		
	1st Hidden Layer	2nd Hidden Layer	Output Layer
full_net	18	20	30
final_0_solids_net	30	8	1
final_1_lighttars_net	11	23	1
final_2_heavytars_net	28	11	1
final_3_totaltars_net	31	18	1
final_4_CO_net	22	14	1
final_5_CO2_net	18	20	1
final_6_othergases_net	19	6	1
final_7_totalgases_net	21	20	1
final_8_H2O_net	22	6	1
final_9_H-C-ratio_net	18	20	1
final_10_Moisture_net	9	23	1
final_11_tot_C--O_net	10	16	1
final_12_tot_aromaticC-O_net	26	18	1
final_13_tot_aromaticC-C_net	23	26	1
final_14_tot_aromaticC-H_net	21	27	1
final_15_tot_aliphaticC-O_net	16	22	1
final_16_tot_aromatic_methoxyl_net	31	30	1
final_17_tot_aliphaticC-C_net	28	17	1
final_18_heavy_C--O_net	21	7	1
final_19_heavy_aromatic_C-O_net	24	25	1
final_20_heavy_aromatic_C-C_net	18	27	1
final_21_heavy_aromatic_C-H_net	17	28	1
final_22_heavy_aliphatic_C-O_net	26	19	1
final_23_heavy_aromatic_methoxyl_net	18	20	1
final_24_heavy_aliphatic_C-C_net	22	26	1
final_25_light_C--O_net	19	26	1
final_26_light_aromatic_C-O_net	24	16	1
final_27_light_aromatic_C-H_net	12	30	1
final_28_light_aromatic_C-O_net	26	14	1
final_29_light_aliphatic_C-C_net	20	7	1

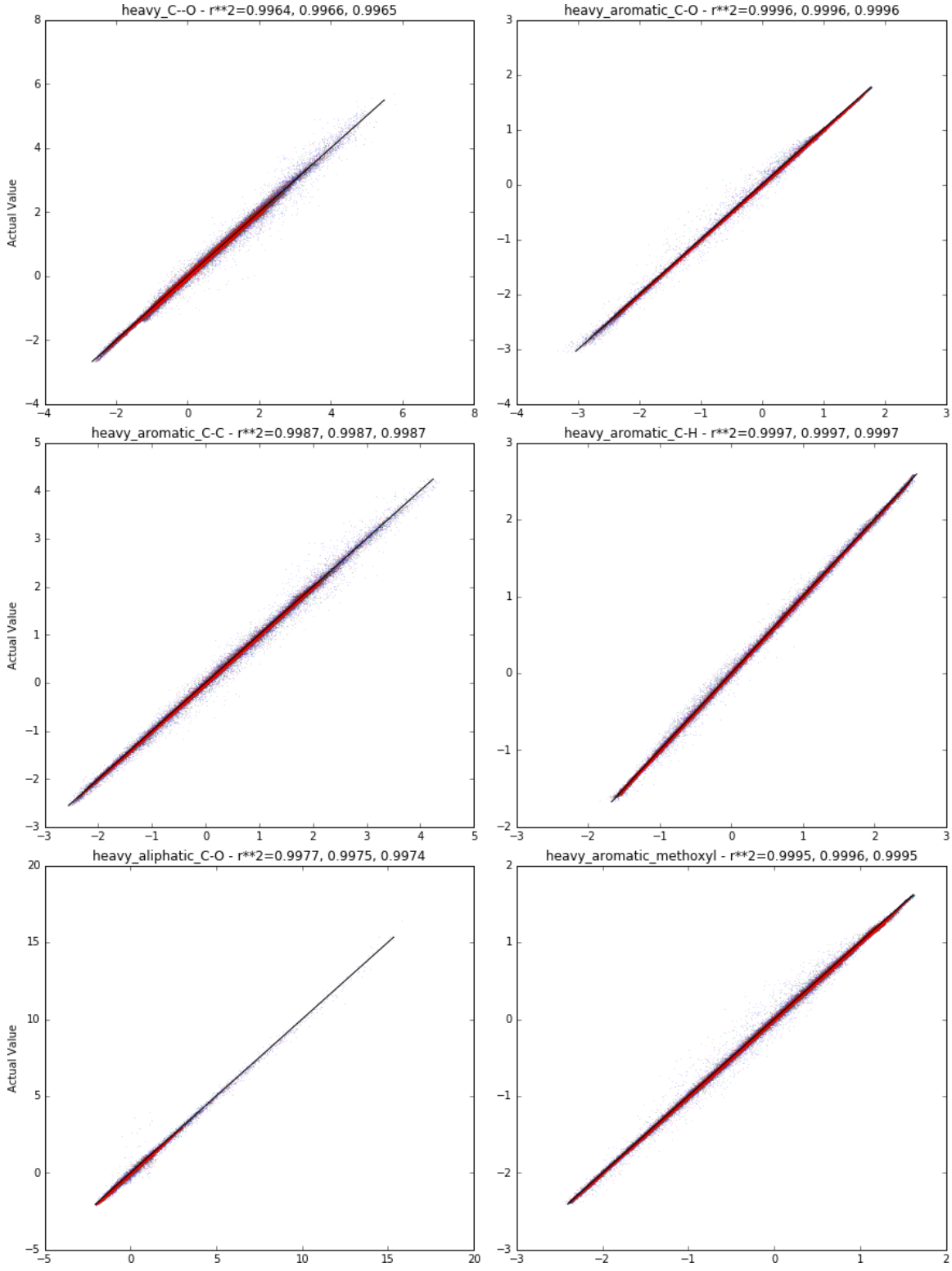
Decision Tree Name	Depth	Min Samples in a Leaf
decision_tree	26	2

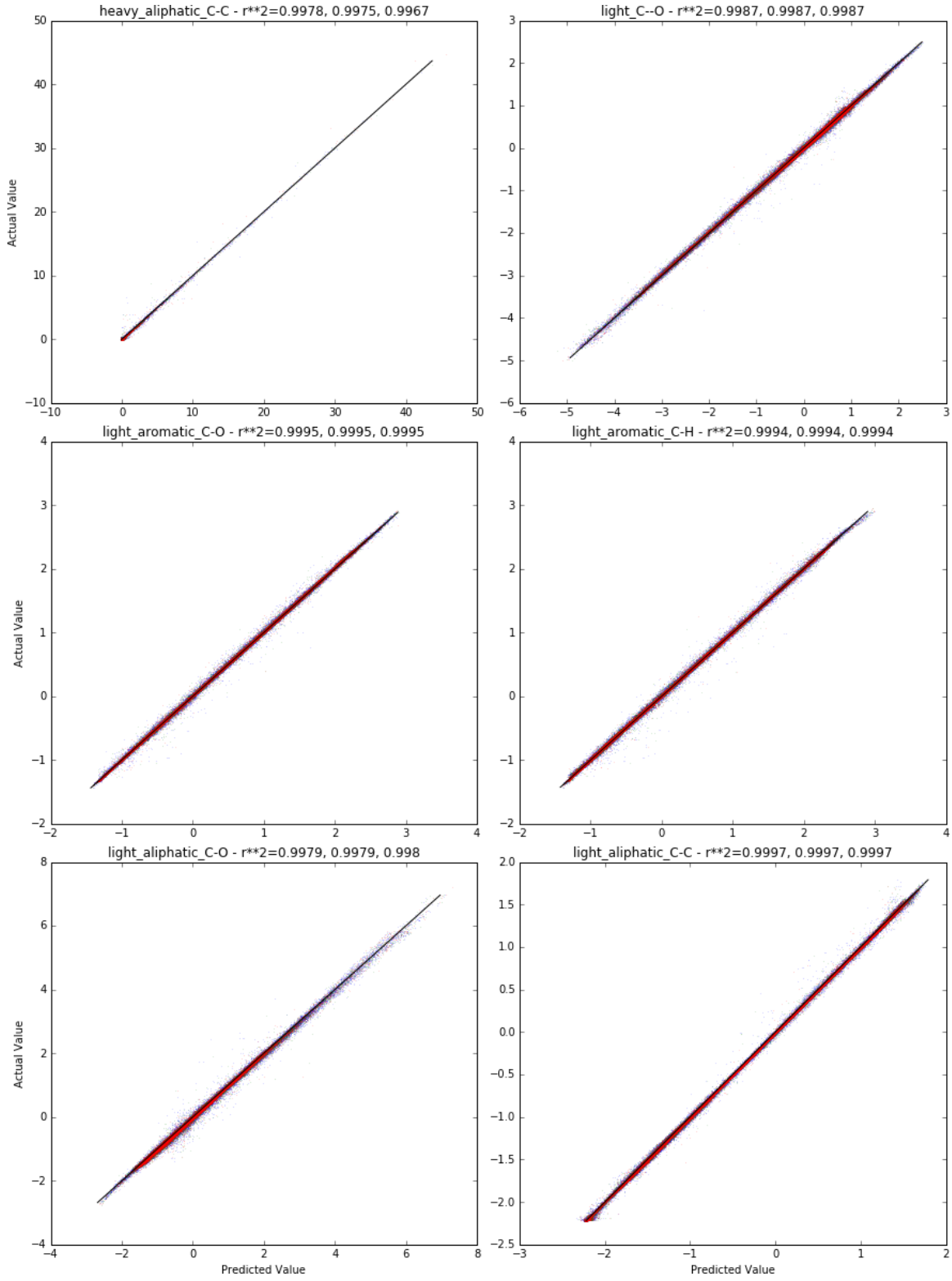
Supporting Figure C.1: Parity plots for the single nets. Training data (blue), validation data (green), and test data (red) are shown. Plot titles indicate the output measure trained, followed by the R^2 values for the training, validation, and test data.



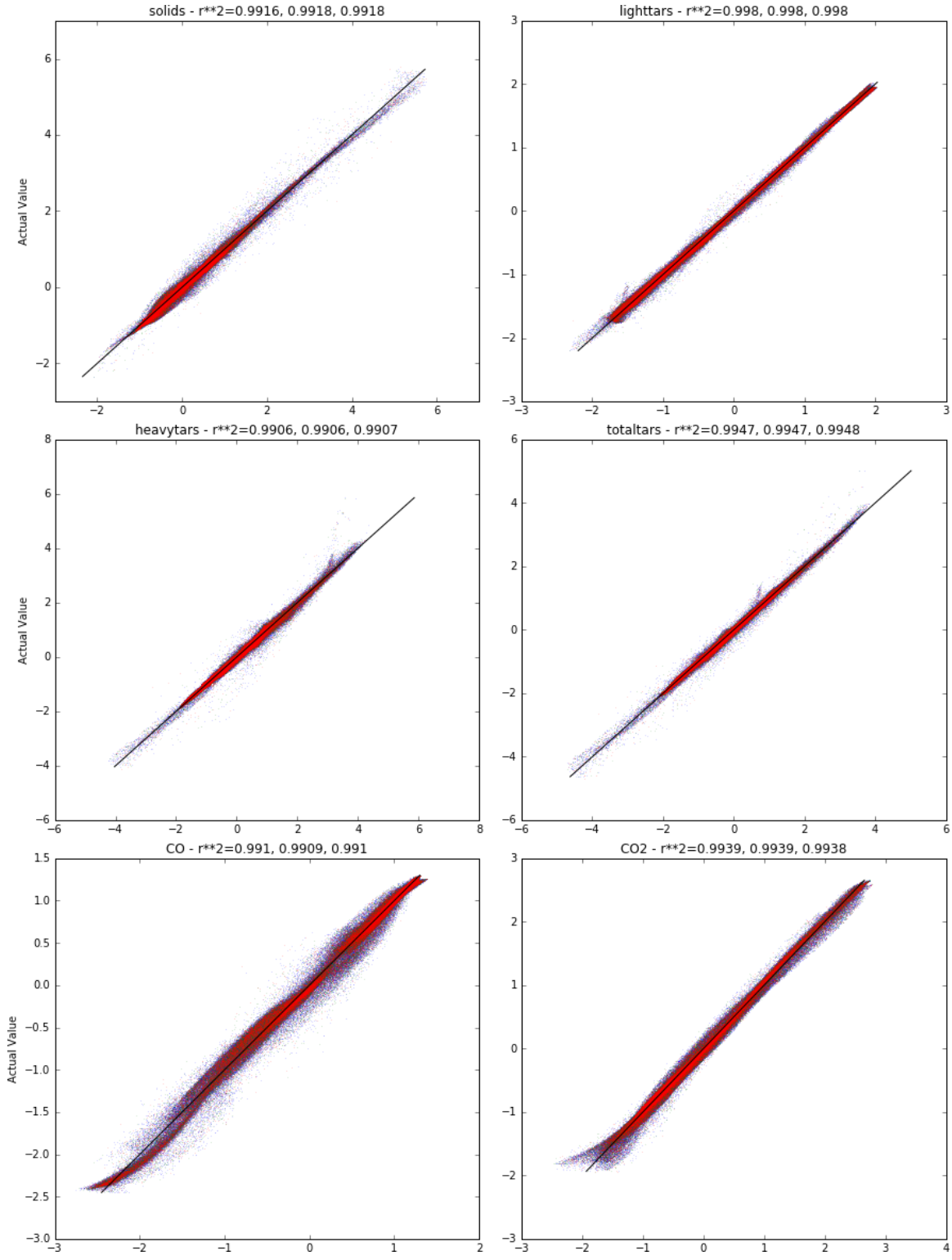

othergases - r**2=0.999, 0.9989, 0.999
totalgases - r**2=0.9992, 0.9992, 0.9992
H20 - r**2=0.9992, 0.9993, 0.9993
H-C-ratio - r**2=0.9989, 0.9989, 0.9989
Moisture - r**2=0.9984, 0.9984, 0.9984
tot_C--O - r**2=0.9994, 0.9993, 0.9993
Actual Value
Actual Value

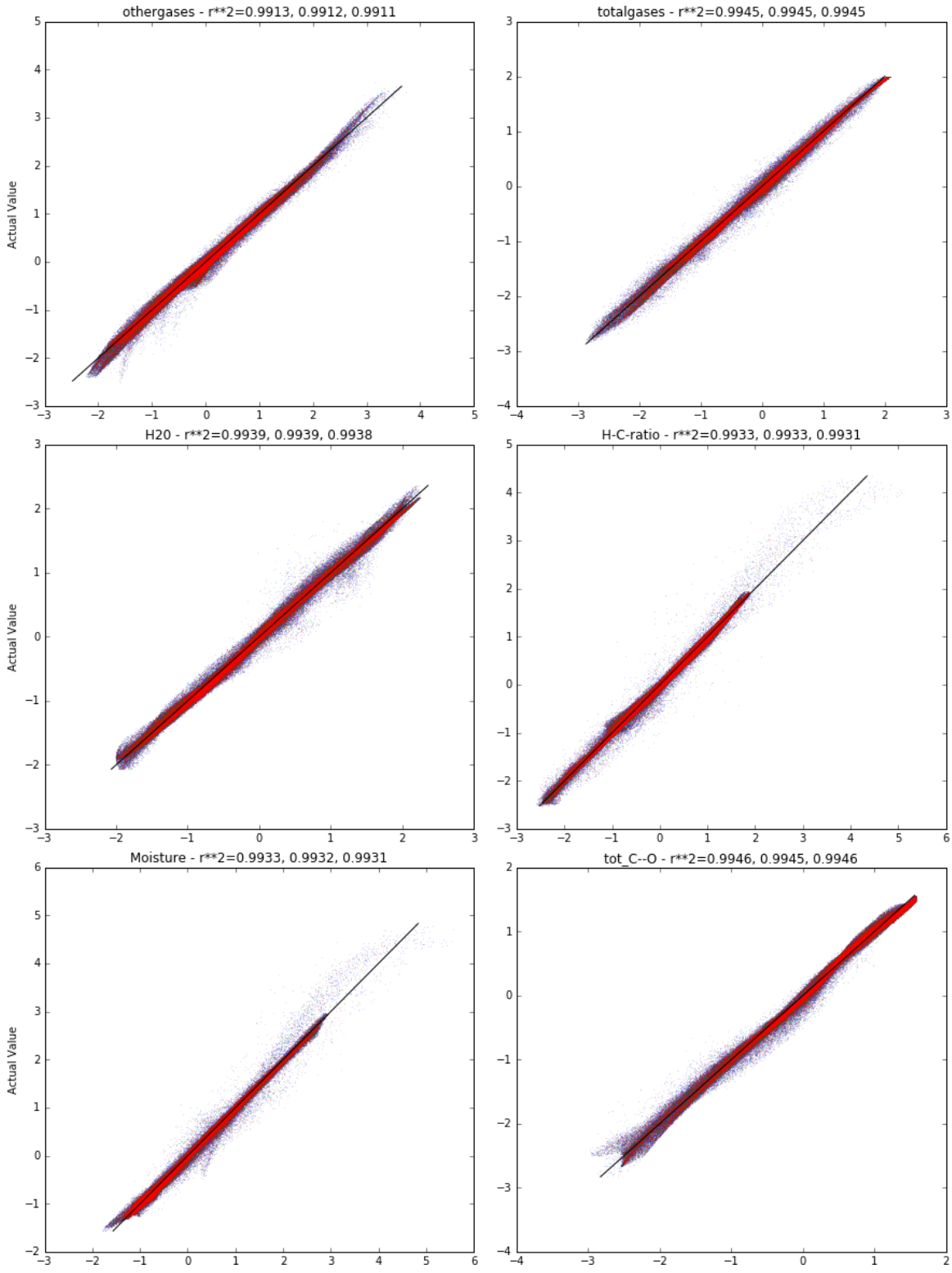

tot_aromaticC-O - r**2=0.9994, 0.9994, 0.9994
tot_aromaticC-C - r**2=0.9983, 0.9985, 0.9983
tot_aromaticC-H - r**2=0.9995, 0.9995, 0.9995
tot_aliphaticC-O - r**2=0.9983, 0.9982, 0.9982
tot_aromatic_methoxyl - r**2=0.9995, 0.9995, 0.9994
tot_aliphaticC-C - r**2=0.9998, 0.9998, 0.9998
Actual Value
Actual Value

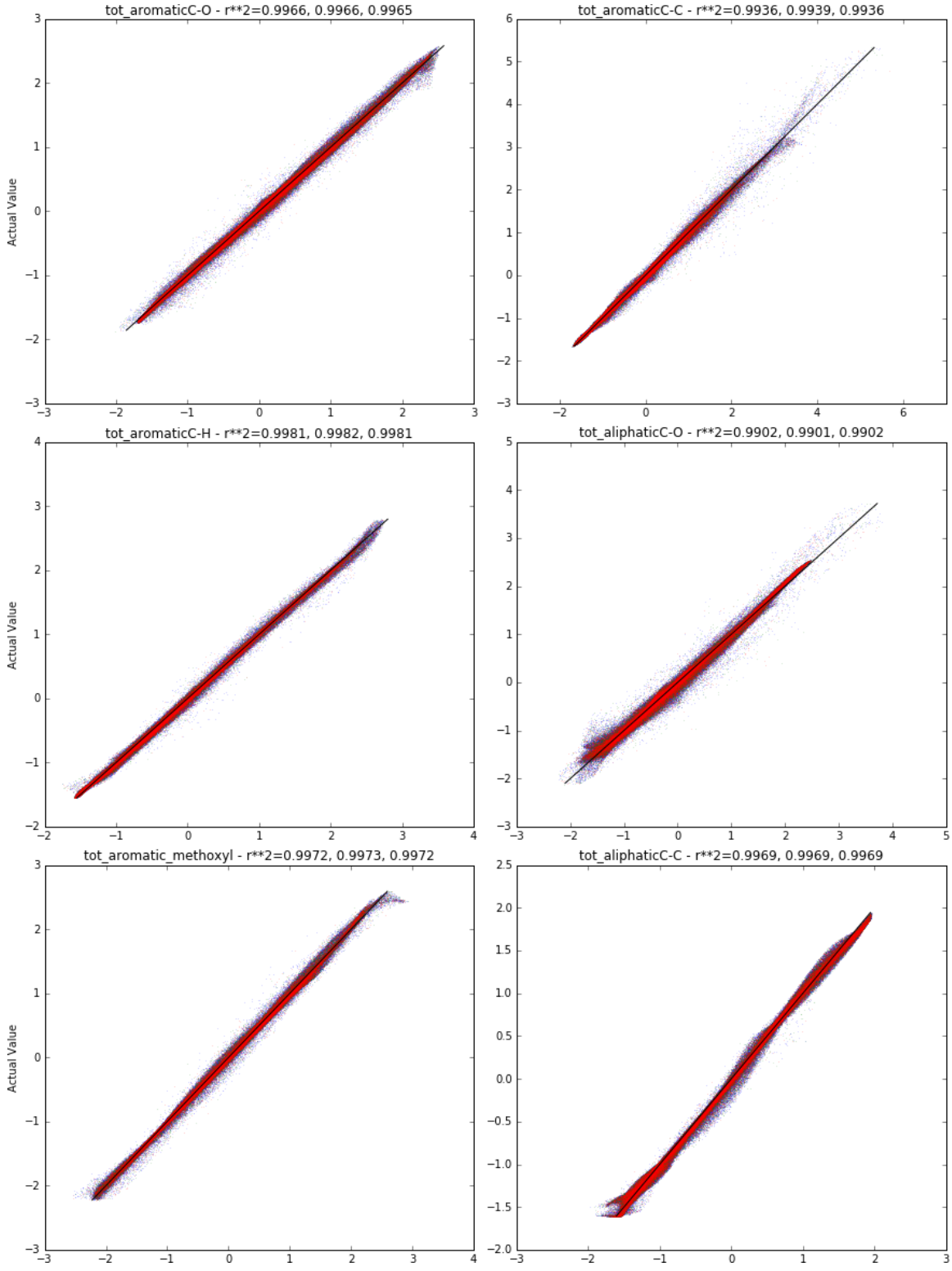

heavy_C--O - r**2=0.9964, 0.9966, 0.9965
heavy_aromatic_C-O - r**2=0.9996, 0.9996, 0.9996
heavy_aromatic_C-C - r**2=0.9987, 0.9987, 0.9987
heavy_aromatic_C-H - r**2=0.9997, 0.9997, 0.9997
heavy_aliphatic_C-O - r**2=0.9977, 0.9975, 0.9974
heavy_aromatic_methoxyl - r**2=0.9995, 0.9996, 0.9995
Actual Value
Actual Value
Actual Value

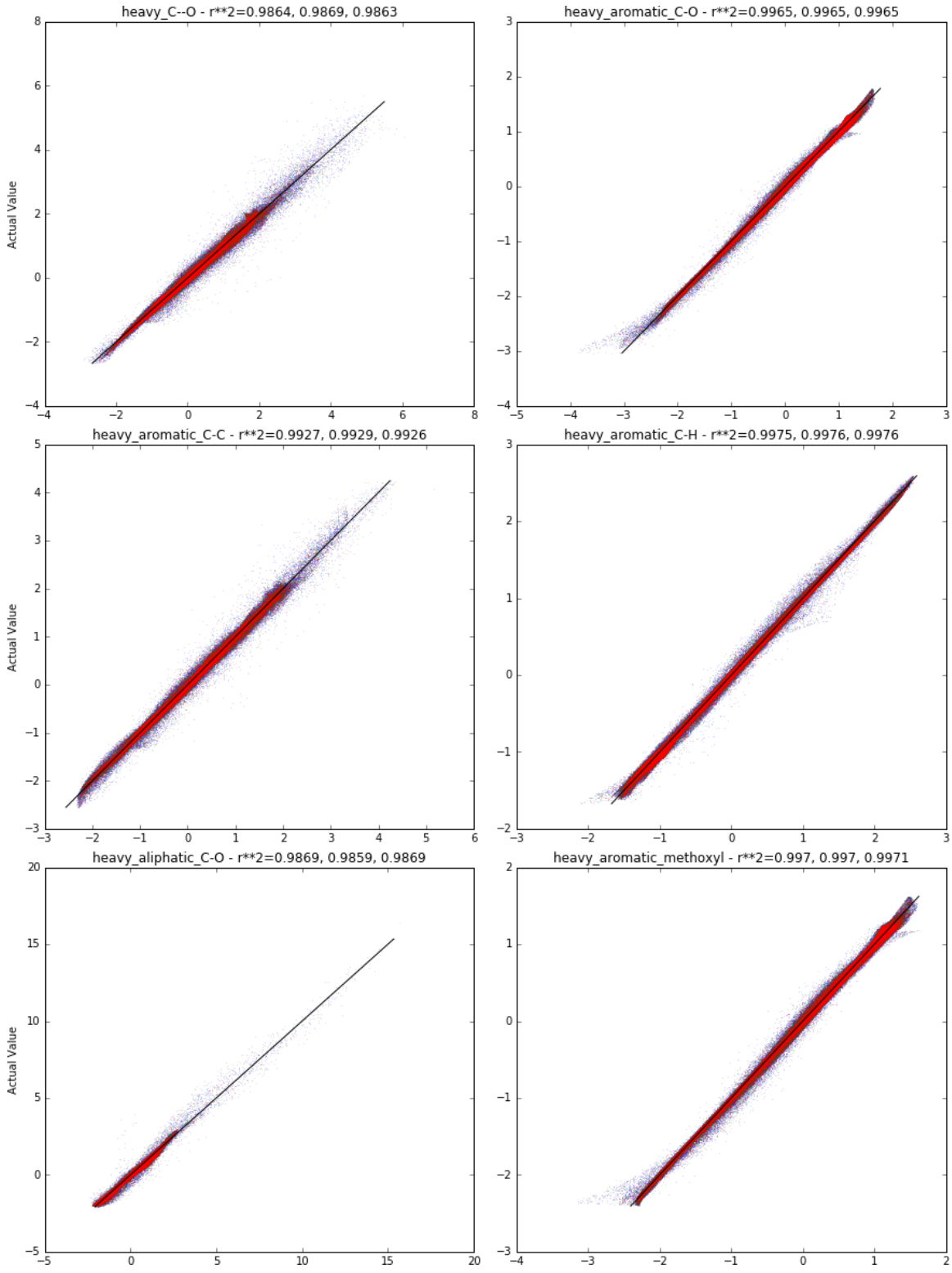

heavy_aliphatic_C-C - r**2=0.9978, 0.9975, 0.9967
light_C--O - r**2=0.9987, 0.9987, 0.9987
light_aromatic_C-O - r**2=0.9995, 0.9995, 0.9995
light_aromatic_C-H - r**2=0.9994, 0.9994, 0.9994
light_aliphatic_C-O - r**2=0.9979, 0.9979, 0.998
light_aliphatic_C-C - r**2=0.9997, 0.9997, 0.9997
Actual Value
Predicted Value

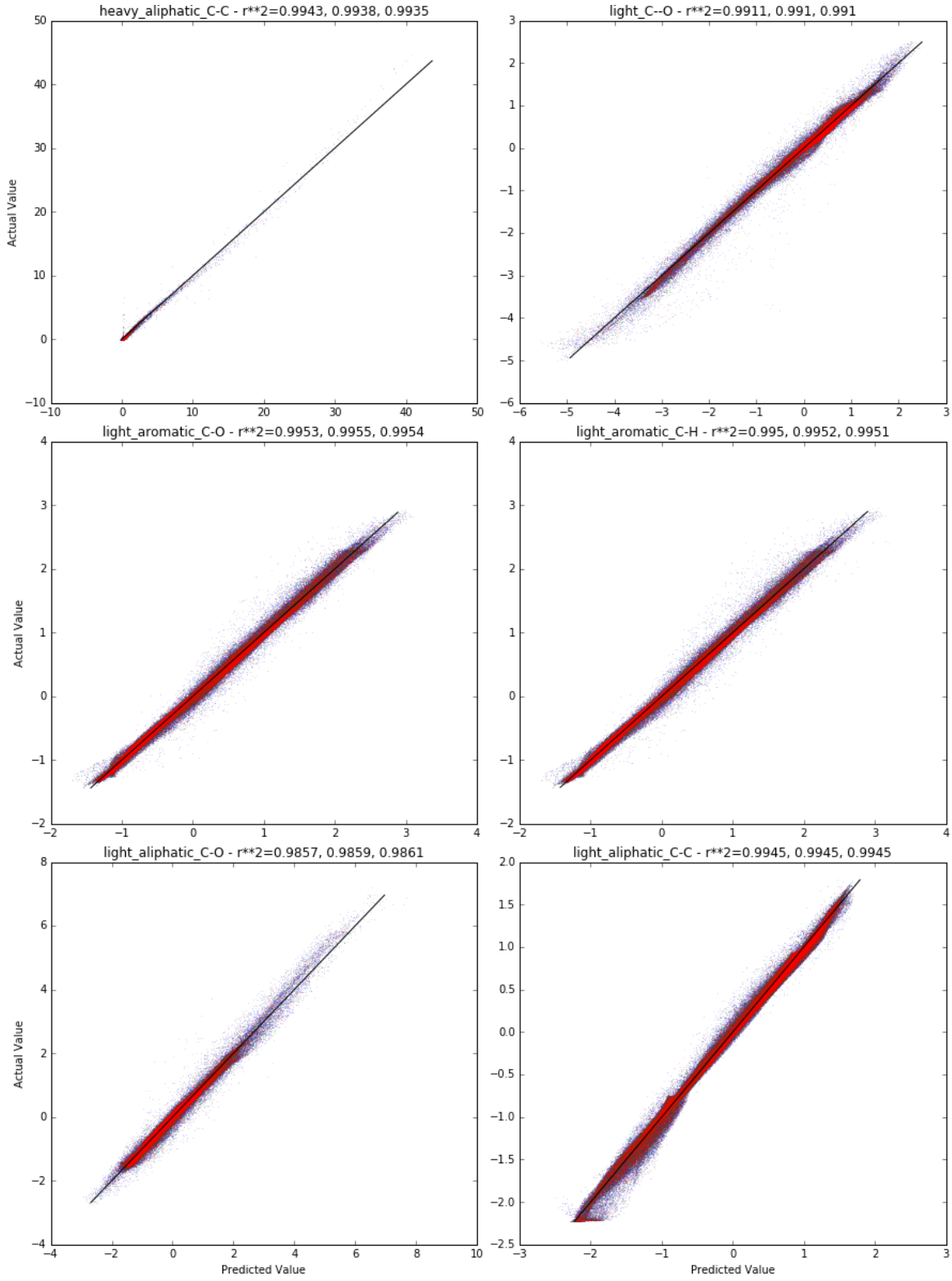
Supporting Figure C.2: Parity plots for the full net. Training data (blue), validation data (green), and test data (red) are shown. Plot titles indicate the output measure, followed by the R^2 values for the training, validation, and test data.



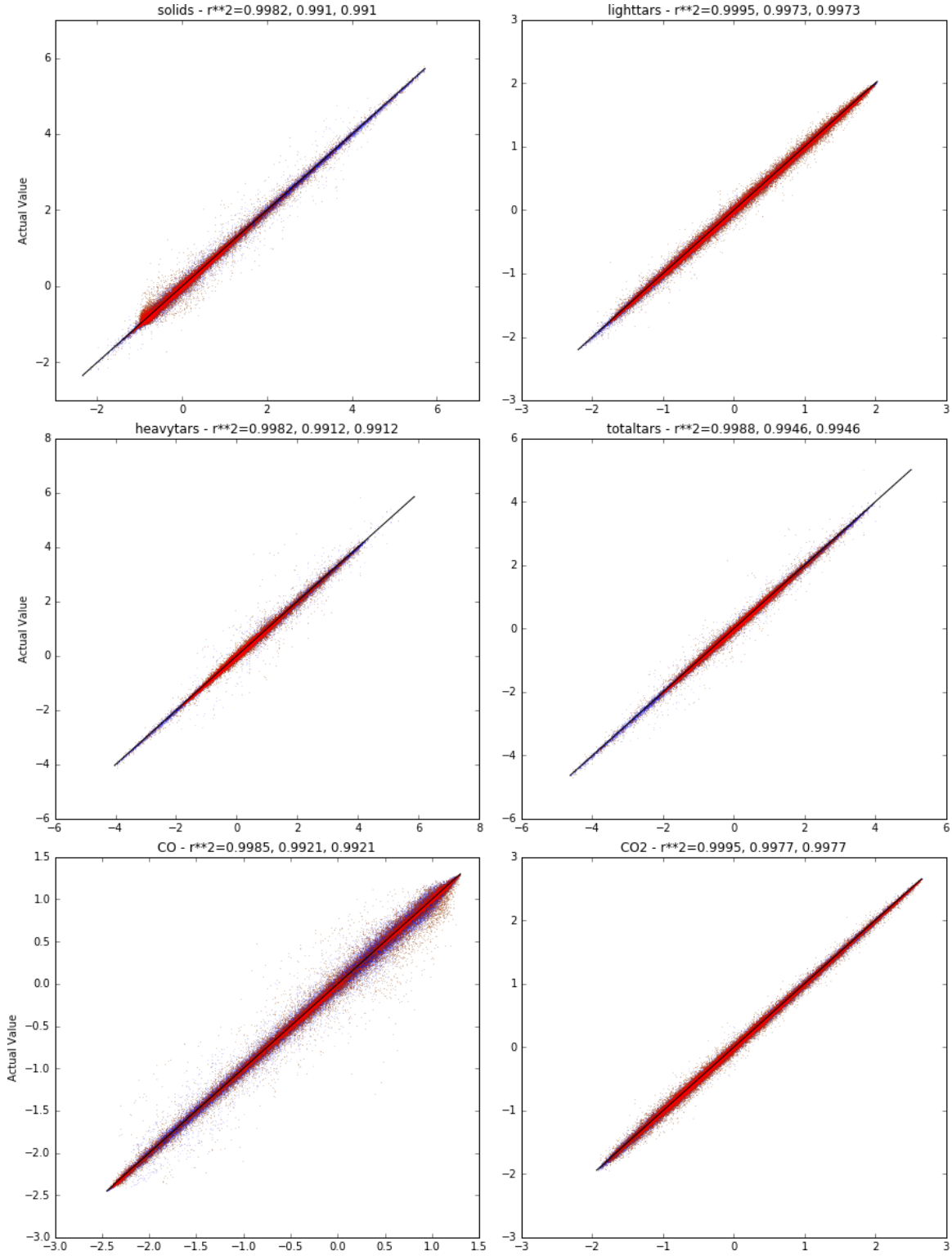

othergases - r**2=0.9913, 0.9912, 0.9911
totalgases - r**2=0.9945, 0.9945, 0.9945
H20 - r**2=0.9939, 0.9939, 0.9938
H-C-ratio - r**2=0.9933, 0.9933, 0.9931
Moisture - r**2=0.9933, 0.9932, 0.9931
tot_C--O - r**2=0.9946, 0.9945, 0.9946
Actual Value
Actual Value

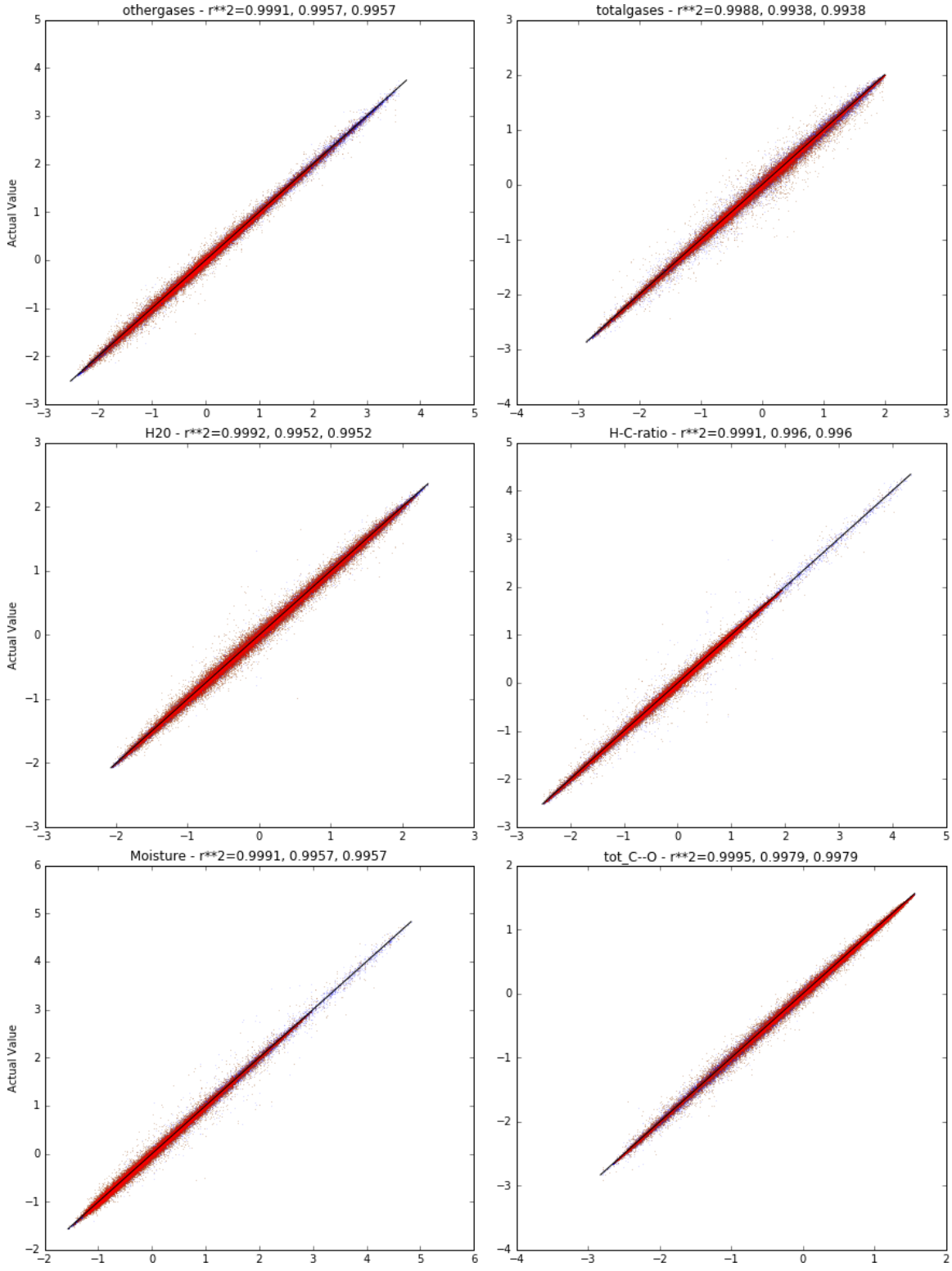

tot_aromaticC-O - r**2=0.9966, 0.9966, 0.9965
tot_aromaticC-C - r**2=0.9936, 0.9939, 0.9936
tot_aromaticC-H - r**2=0.9981, 0.9982, 0.9981
tot_aliphaticC-O - r**2=0.9902, 0.9901, 0.9902
tot_aromatic_methoxyl - r**2=0.9972, 0.9973, 0.9972
tot_aliphaticC-C - r**2=0.9969, 0.9969, 0.9969
Actual Value
Actual Value

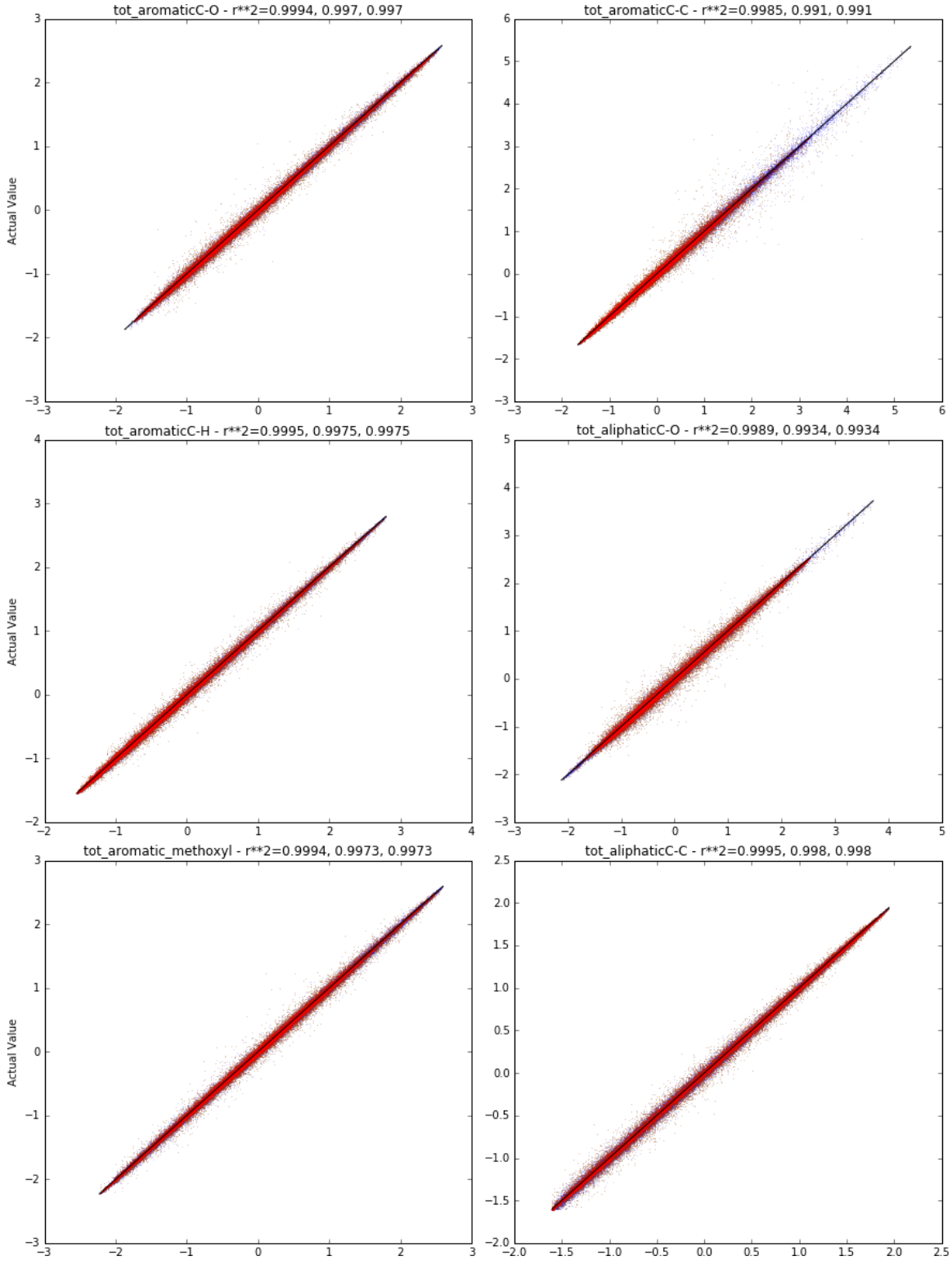

heavy_C--O - r**2=0.9864, 0.9869, 0.9863
heavy_aromatic_C-O - r**2=0.9965, 0.9965, 0.9965
heavy_aromatic_C-C - r**2=0.9927, 0.9929, 0.9926
heavy_aromatic_C-H - r**2=0.9975, 0.9976, 0.9976
heavy_aliphatic_C-O - r**2=0.9869, 0.9859, 0.9869
heavy_aromatic_methoxyl - r**2=0.997, 0.997, 0.9971
Actual Value
Actual Value
Actual Value

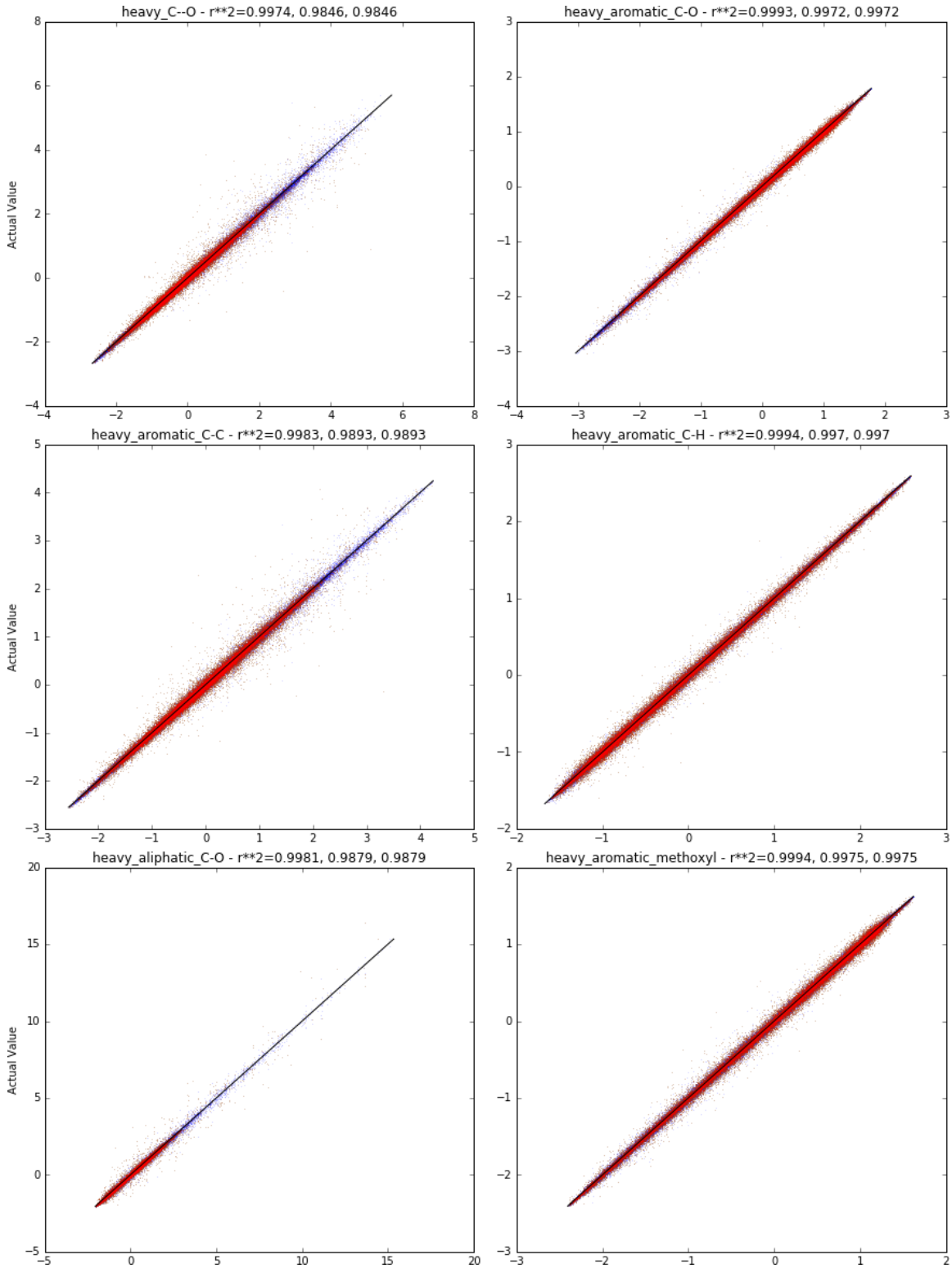

heavy_aliphatic_C-C - r**2=0.9943, 0.9938, 0.9935
light_C--O - r**2=0.9911, 0.991, 0.991
light_aromatic_C-O - r**2=0.9953, 0.9955, 0.9954
light_aromatic_C-H - r**2=0.995, 0.9952, 0.9951
light_aliphatic_C-O - r**2=0.9857, 0.9859, 0.9861
light_aliphatic_C-C - r**2=0.9945, 0.9945, 0.9945
Actual Value
Predicted Value

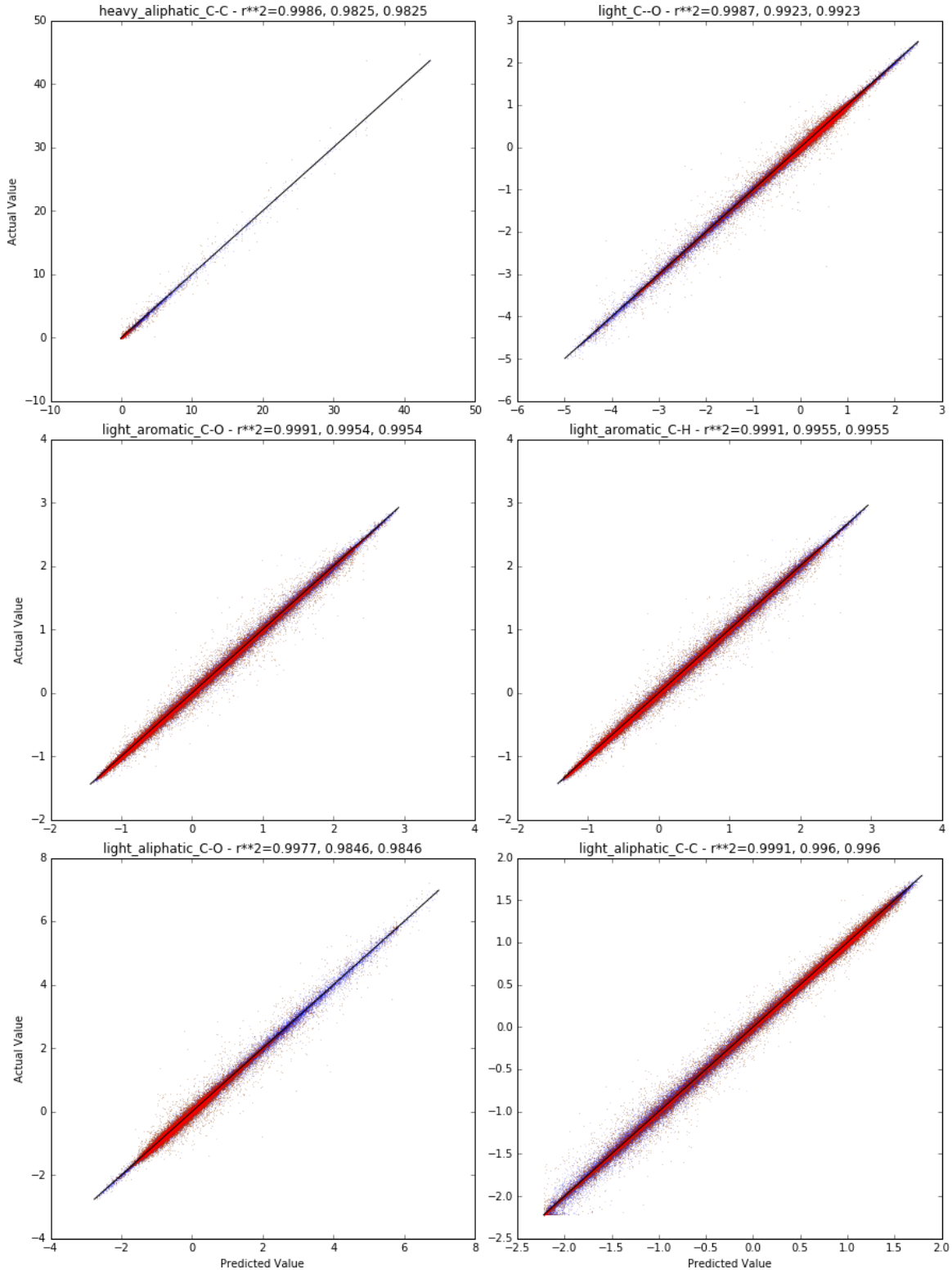
Supporting Figure C.3: Parity plots for the decision tree. Training data (blue) and test data (red) are shown. Plot titles indicate the output measure trained, followed by the R^2 values for the training and test data (test set value is repeated twice in these titles).



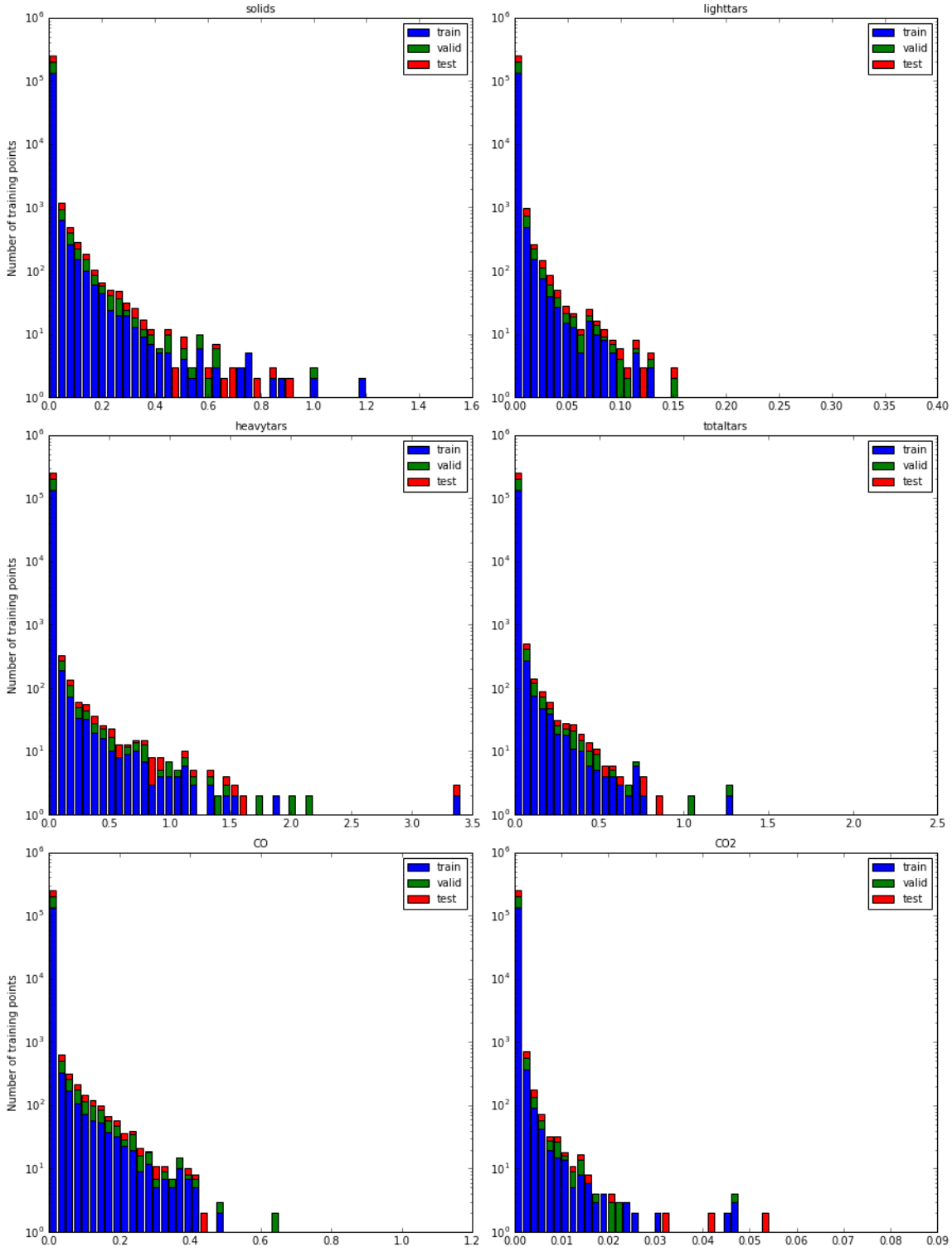

othergases - r**2=0.9991, 0.9957, 0.9957
totalgases - r**2=0.9988, 0.9938, 0.9938
H20 - r**2=0.9992, 0.9952, 0.9952
H-C-ratio - r**2=0.9991, 0.996, 0.996
Moisture - r**2=0.9991, 0.9957, 0.9957
tot_C--O - r**2=0.9995, 0.9979, 0.9979
Actual Value
Actual Value

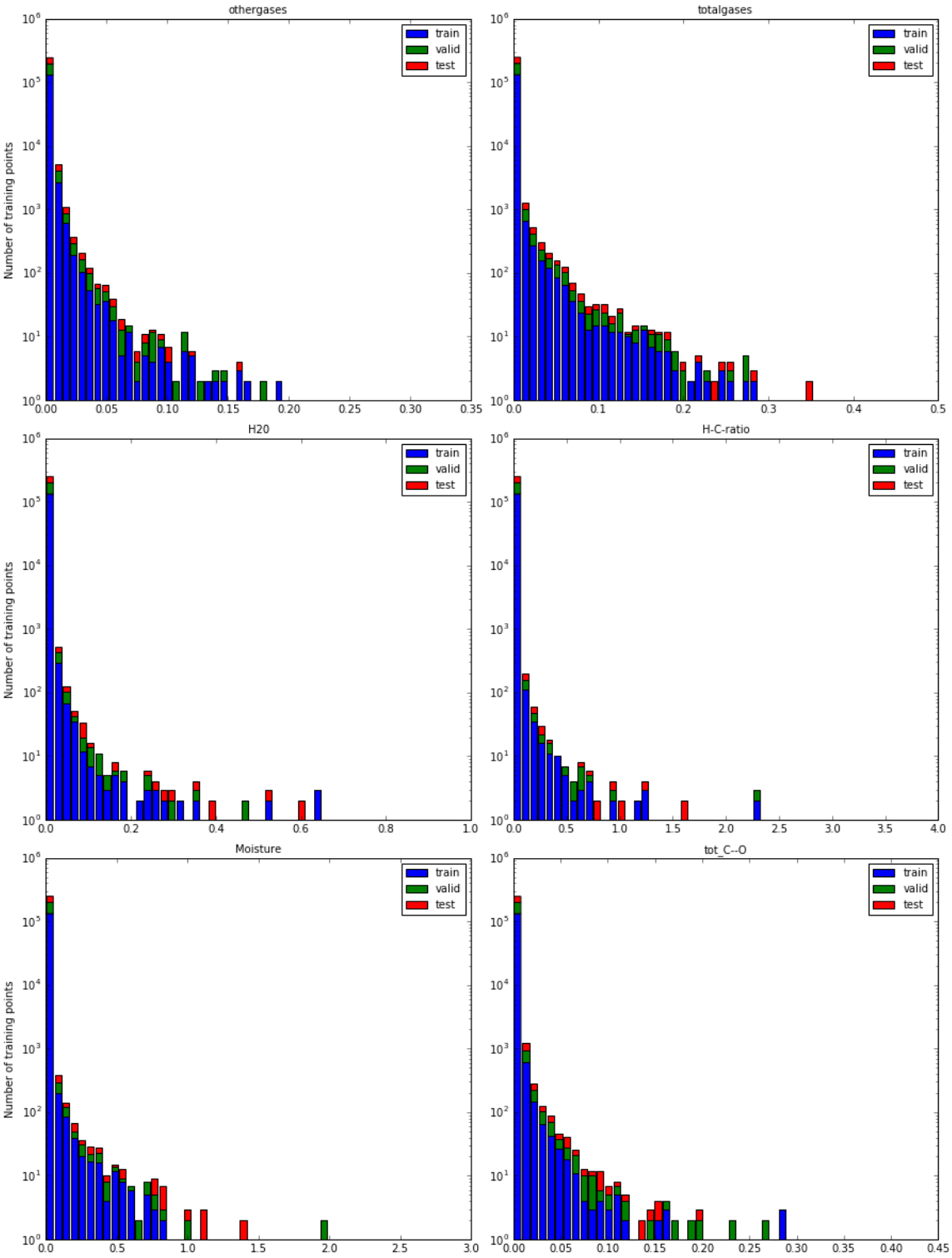

tot_aromaticC-O - r**2=0.9994, 0.997, 0.997
tot_aromaticC-C - r**2=0.9985, 0.991, 0.991
tot_aromaticC-H - r**2=0.9995, 0.9975, 0.9975
tot_aliphaticC-O - r**2=0.9989, 0.9934, 0.9934
tot_aromatic_methoxyl - r**2=0.9994, 0.9973, 0.9973
tot_aliphaticC-C - r**2=0.9995, 0.998, 0.998
Actual Value
Actual Value

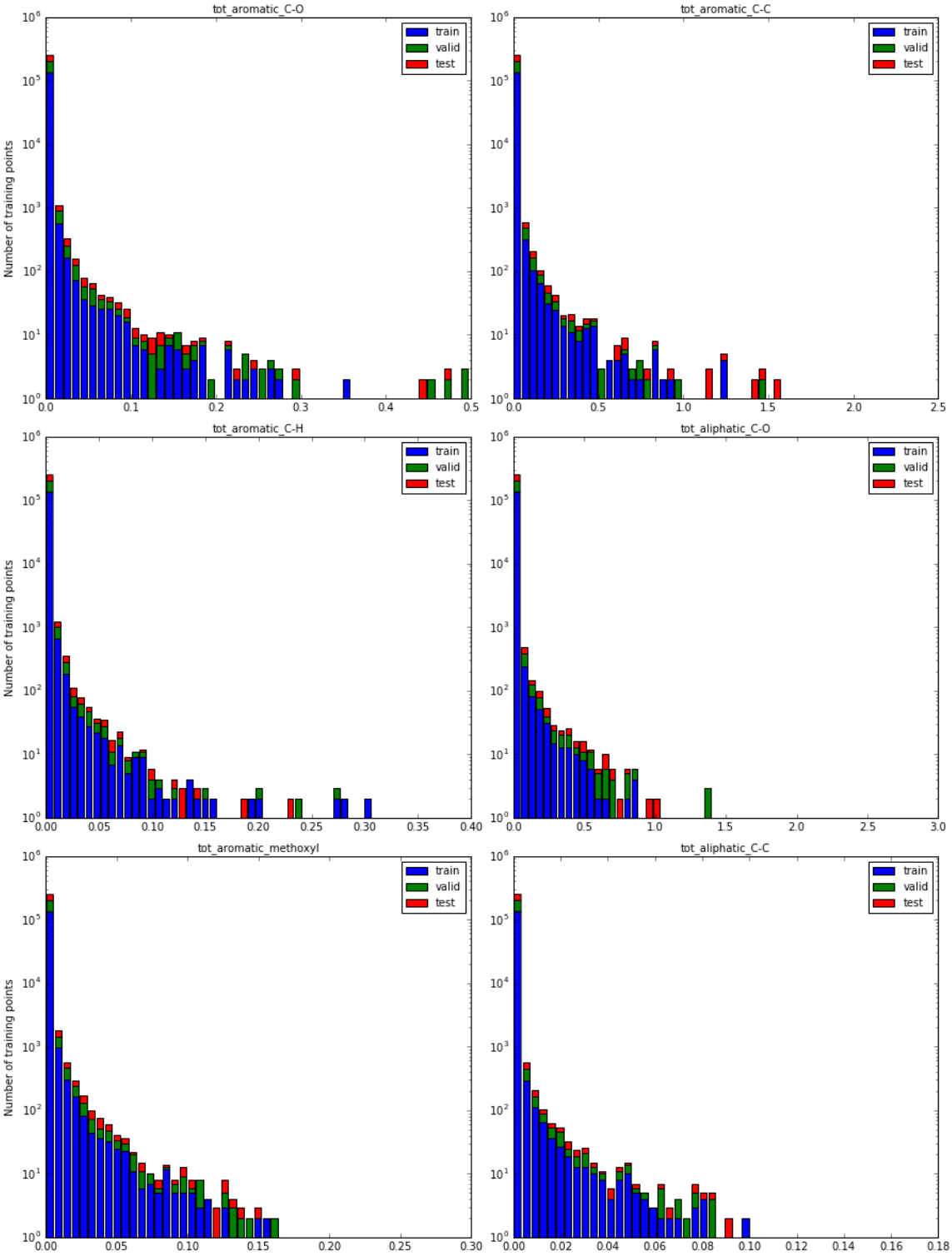

heavy_C--O - r**2=0.9974, 0.9846, 0.9846
heavy_aromatic_C-O - r**2=0.9993, 0.9972, 0.9972
heavy_aromatic_C-C - r**2=0.9983, 0.9893, 0.9893
heavy_aromatic_C-H - r**2=0.9994, 0.997, 0.997
heavy_aliphatic_C-O - r**2=0.9981, 0.9879, 0.9879
heavy_aromatic_methoxyl - r**2=0.9994, 0.9975, 0.9975
Actual Value
Actual Value
Actual Value

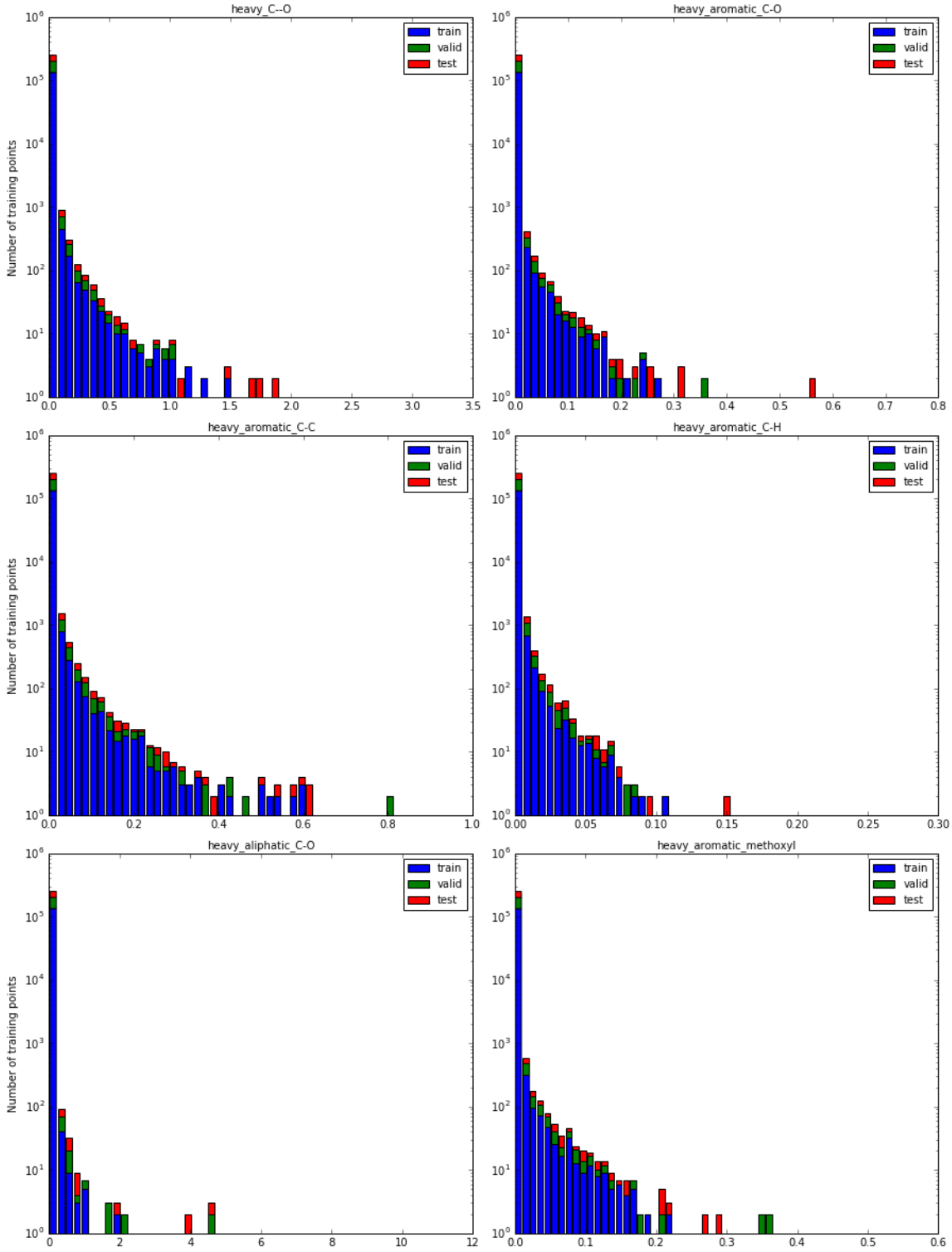

heavy_aliphatic_C-C - r**2=0.9986, 0.9825, 0.9825
light_C--O - r**2=0.9987, 0.9923, 0.9923
light_aromatic_C-O - r**2=0.9991, 0.9954, 0.9954
light_aromatic_C-H - r**2=0.9991, 0.9955, 0.9955
light_aliphatic_C-O - r**2=0.9977, 0.9846, 0.9846
light_aliphatic_C-C - r**2=0.9991, 0.996, 0.996
Actual Value
Predicted Value

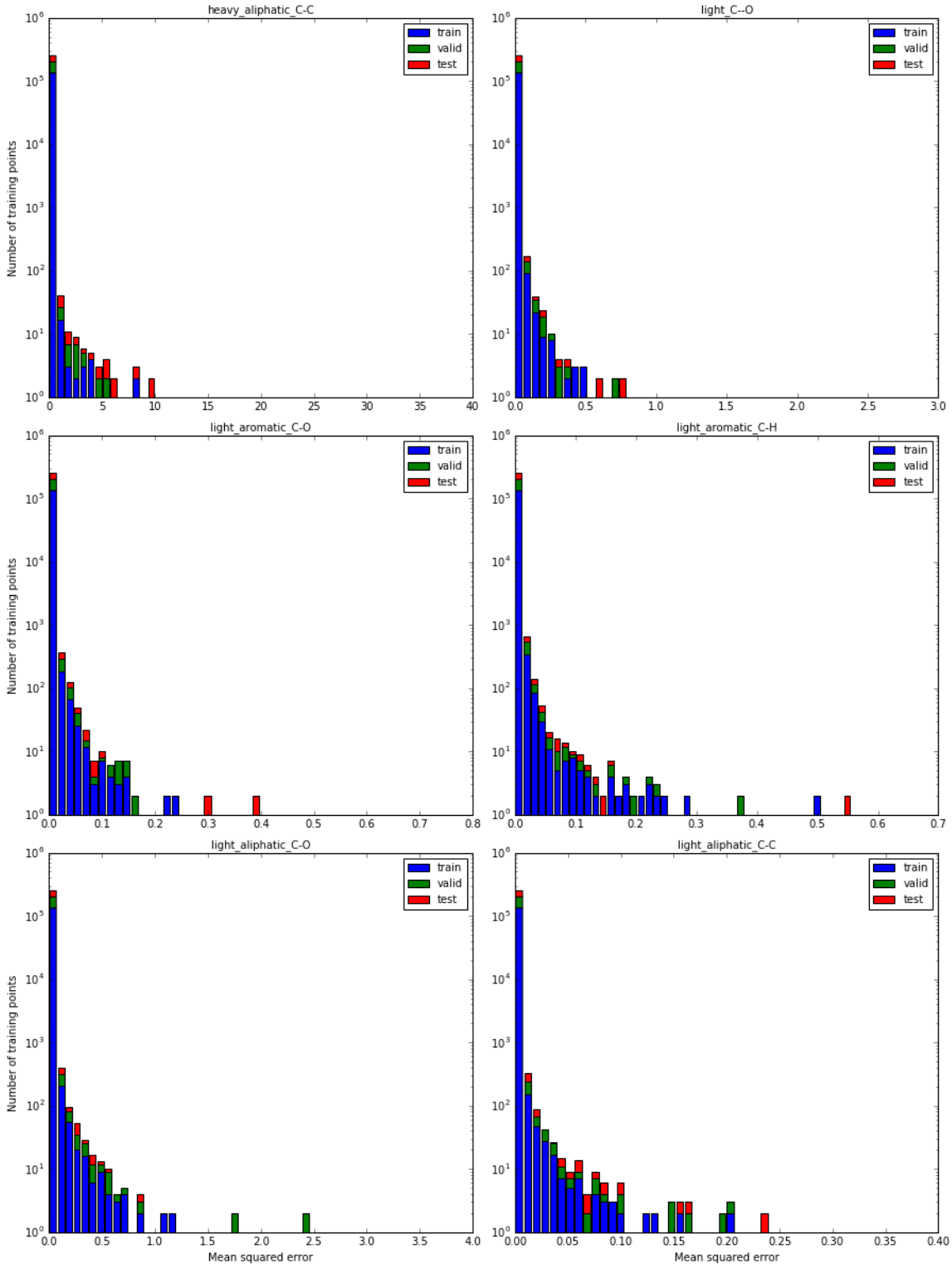
Supporting Figure C.4: Histogram of mean squared errors for single nets.



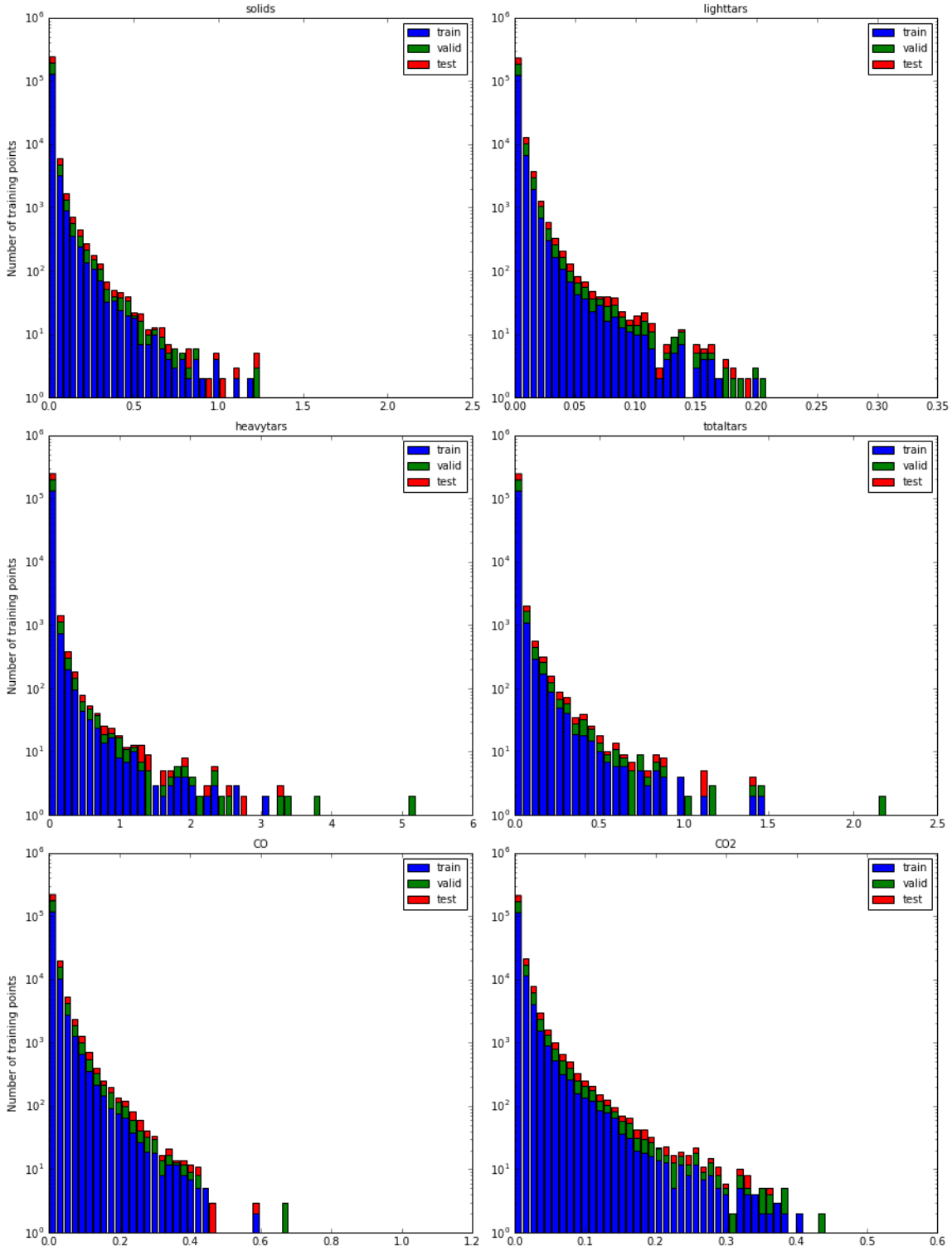

othergases
totalgases
H20
H-C-ratio
Moisture
tot_C--O
train
valid
test
Number of training points

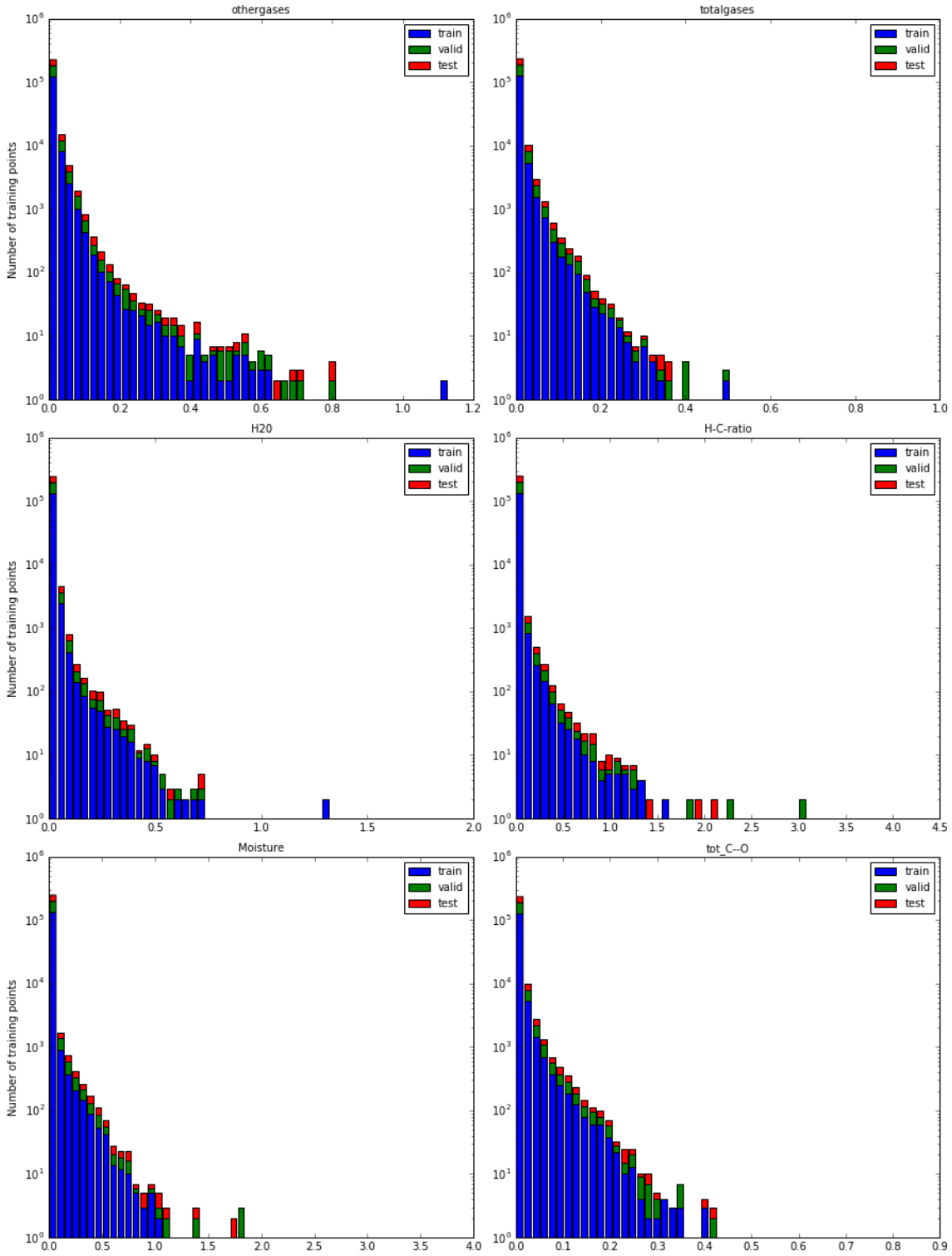

tot_aromatic_C-O
tot_aromatic_C-C
tot_aromatic_C-H
tot_aliphatic_C-O
tot_aromatic_methoxyl
tot_aliphatic_C-C
train
valid
test
Number of training points

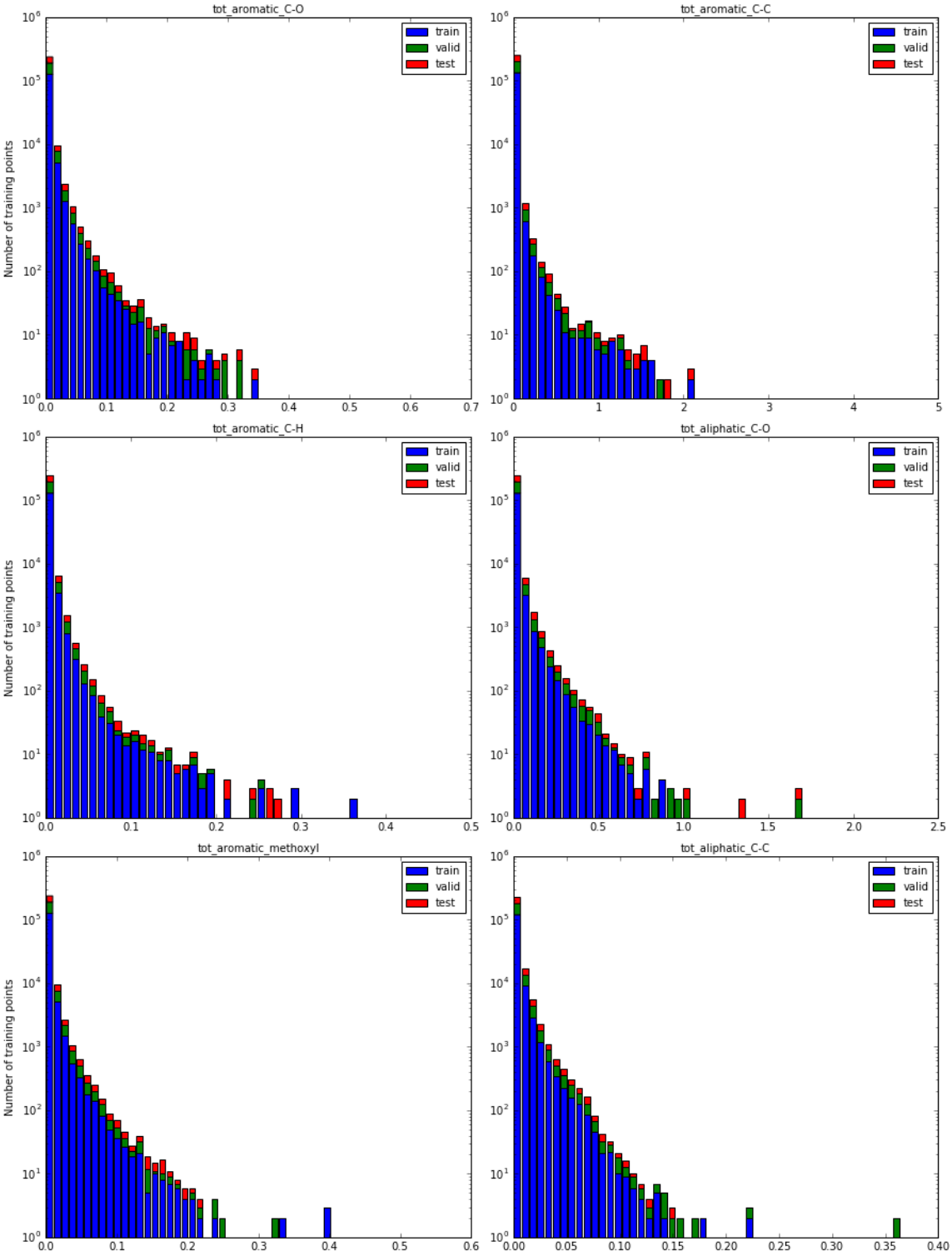

heavy_C--O
heavy_aromatic_C-O
heavy_aromatic_C-C
heavy_aromatic_C-H
heavy_aliphatic_C-O
heavy_aromatic_methoxyl
train
valid
test
Number of training points

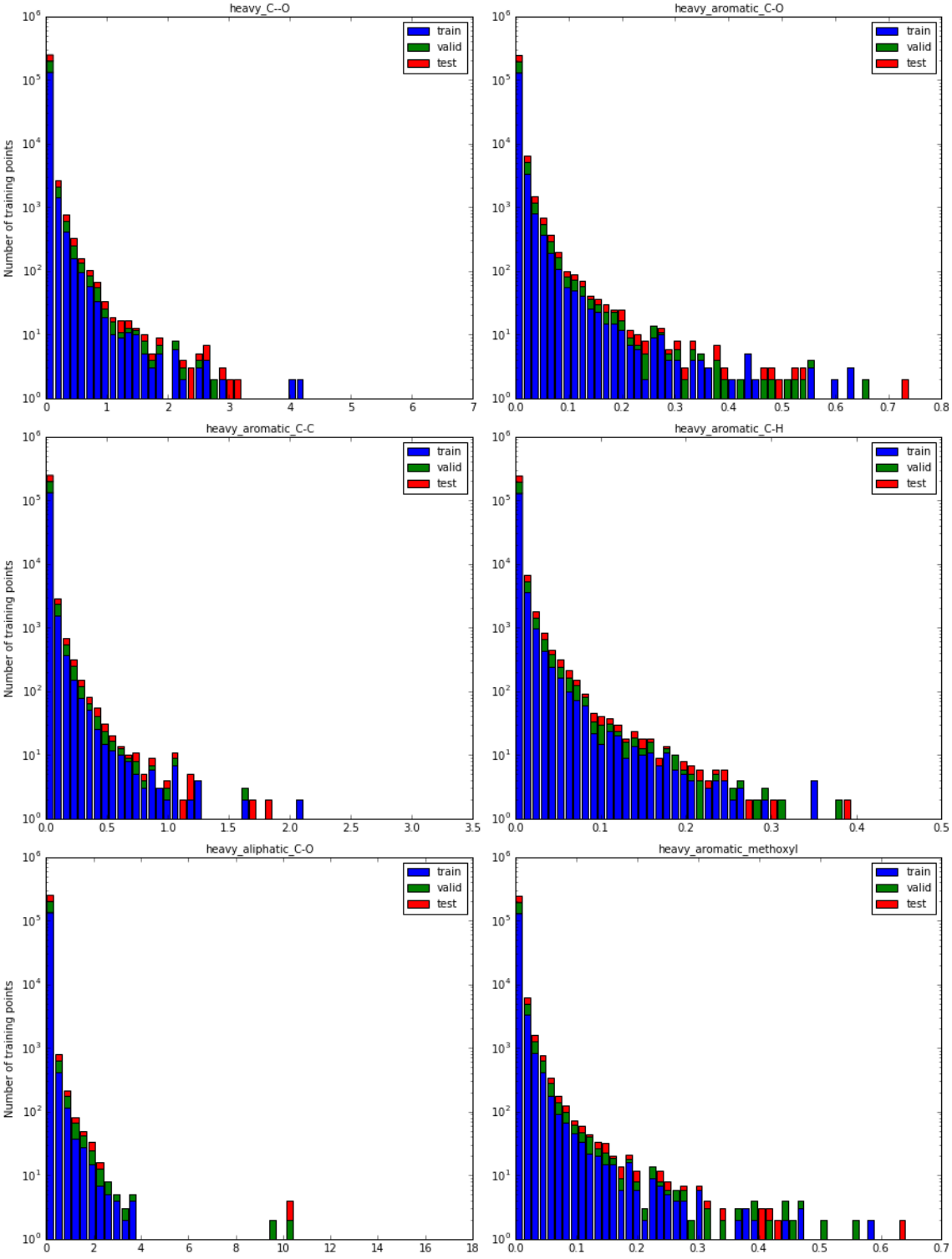

heavy_aliphatic_C-C
light_C--O
light_aromatic_C-O
light_aromatic_C-H
light_aliphatic_C-O
light_aliphatic_C-C
train
valid
test
Number of training points
Mean squared error

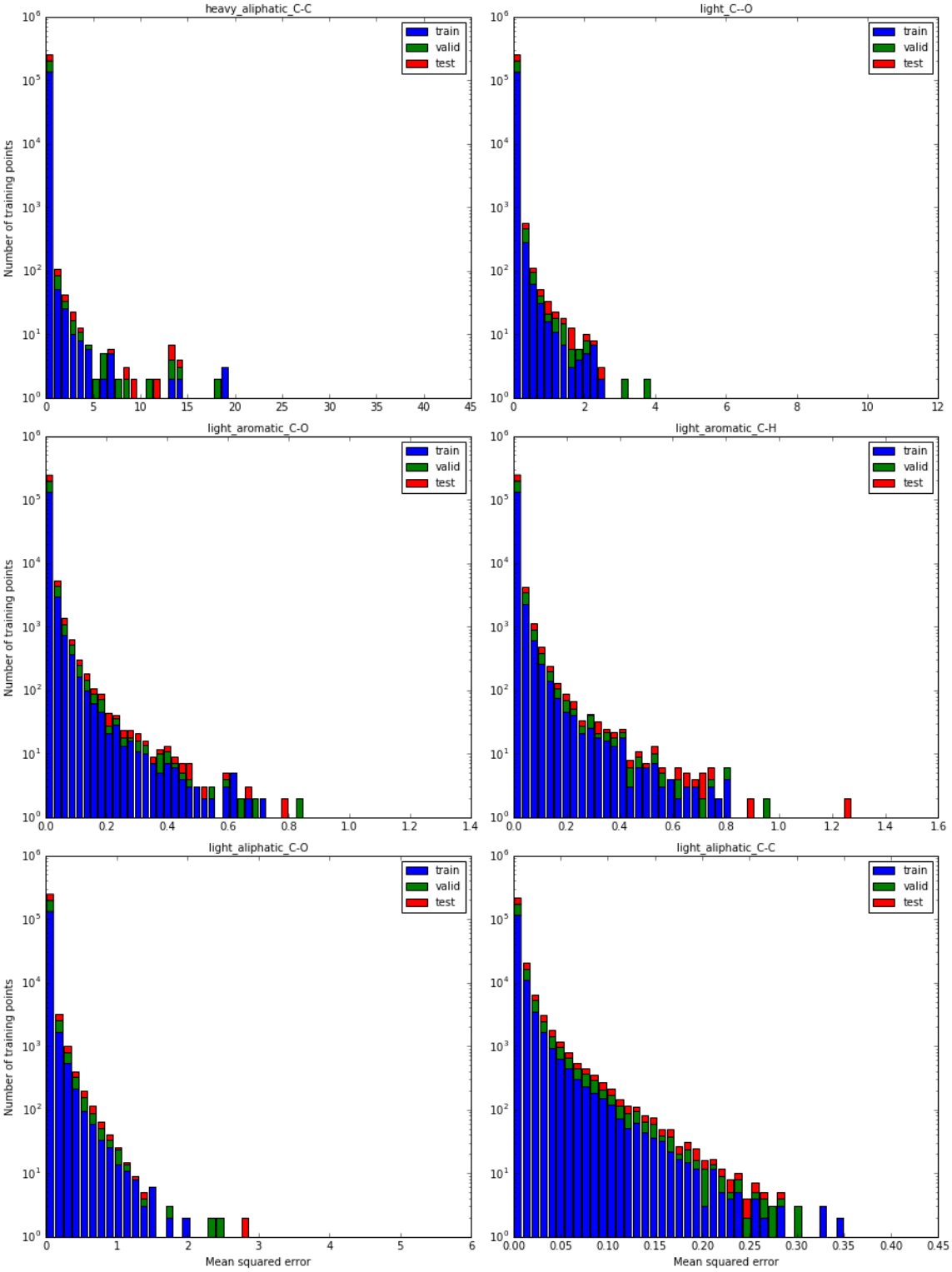
Supporting Figure C.5: Histogram of mean squared errors for full net.



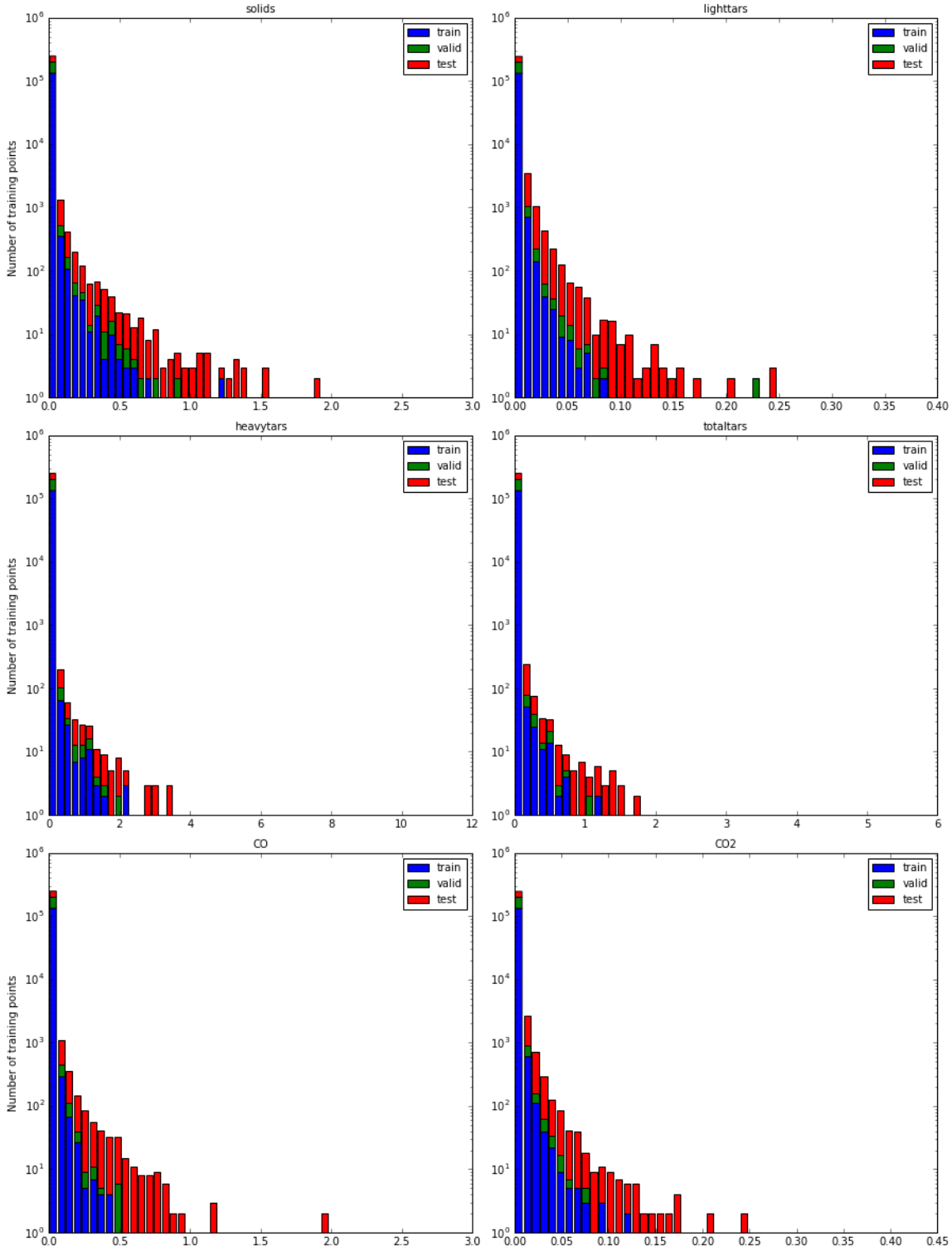

othergases
totalgases
H20
H-C-ratio
Moisture
tot_C--O
train
valid
test
Number of training points

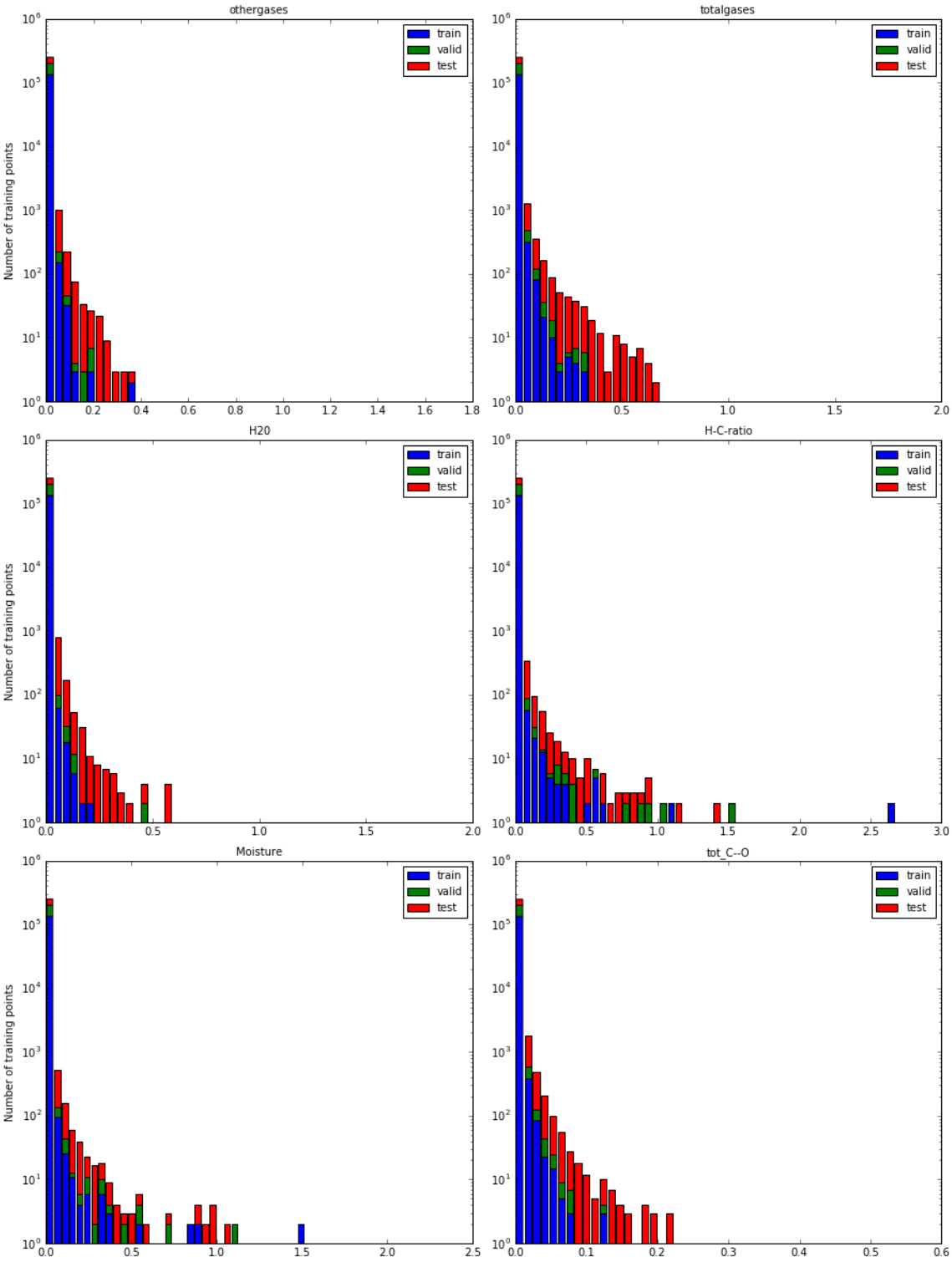

tot_aromatic_C-O
tot_aromatic_C-C
tot_aromatic_C-H
tot_aliphatic_C-O
tot_aromatic_methoxyl
tot_aliphatic_C-C
train
valid
test
Number of training points

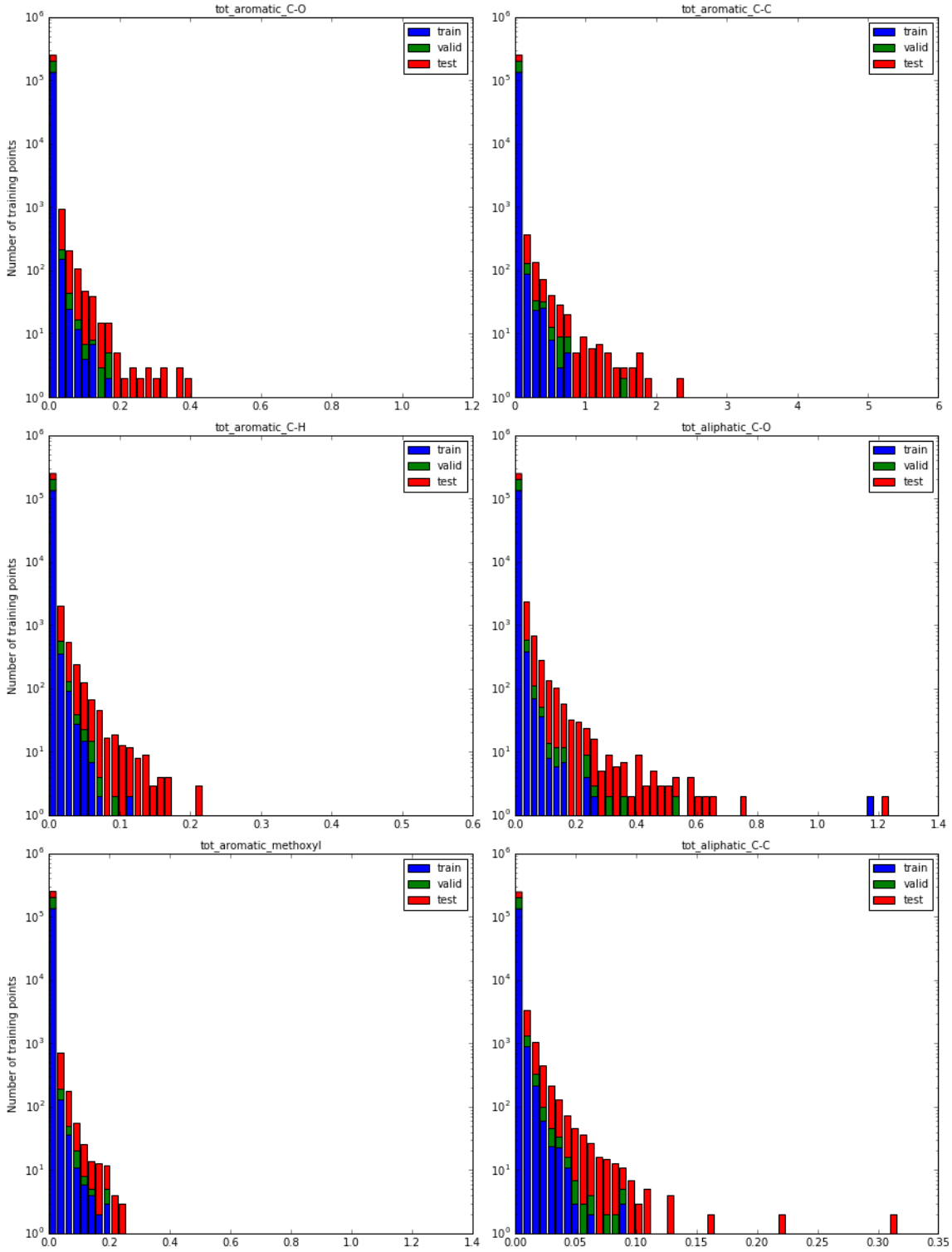

heavy_C--O
heavy_aromatic_C-O
heavy_aromatic_C-C
heavy_aromatic_C-H
heavy_aliphatic_C-O
heavy_aromatic_methoxyl
train
valid
test
Number of training points

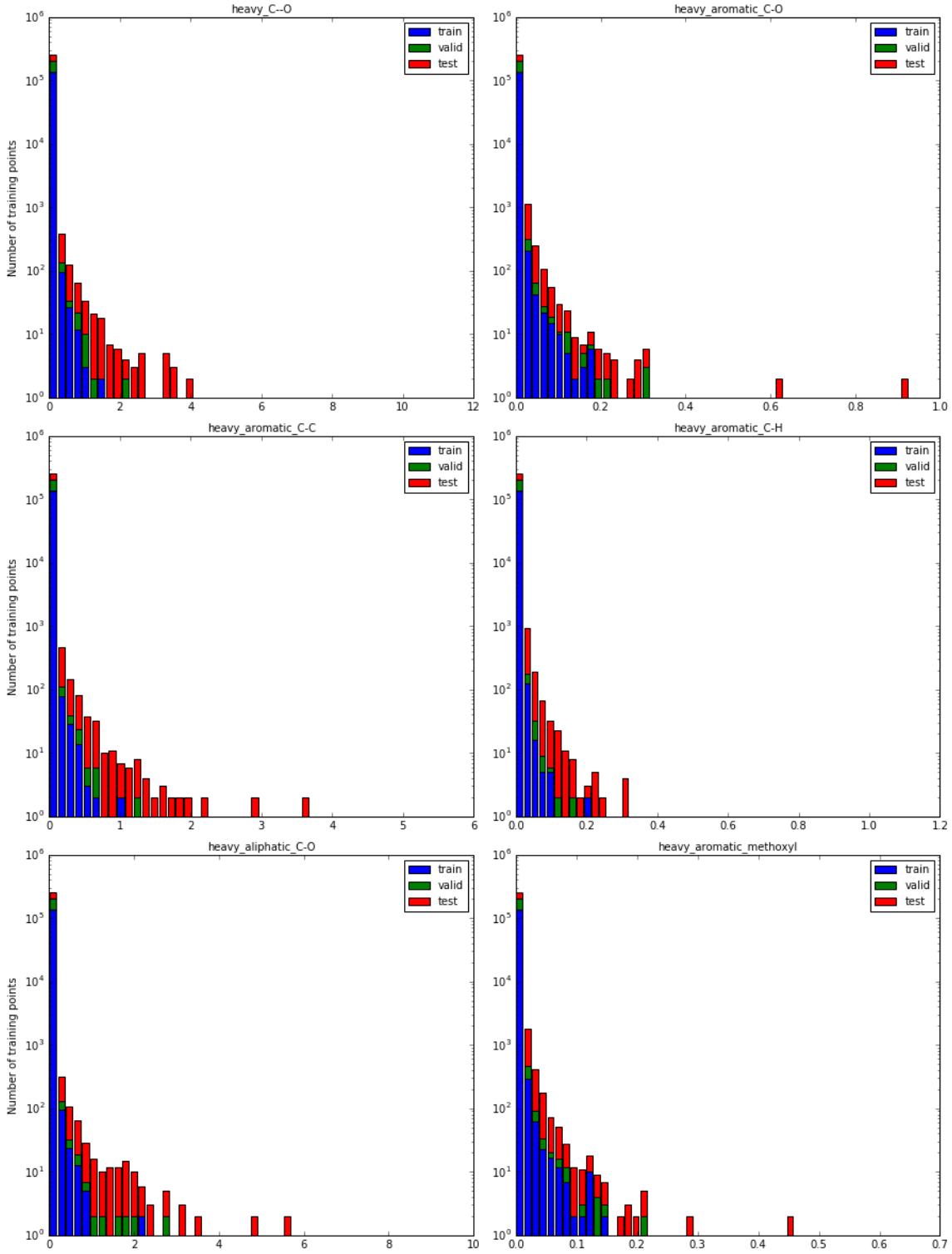

heavy_aliphatic_C-C
light_C--O
light_aromatic_C-O
light_aromatic_C-H
light_aliphatic_C-O
light_aliphatic_C-C
train
valid
test
Number of training points
Mean squared error

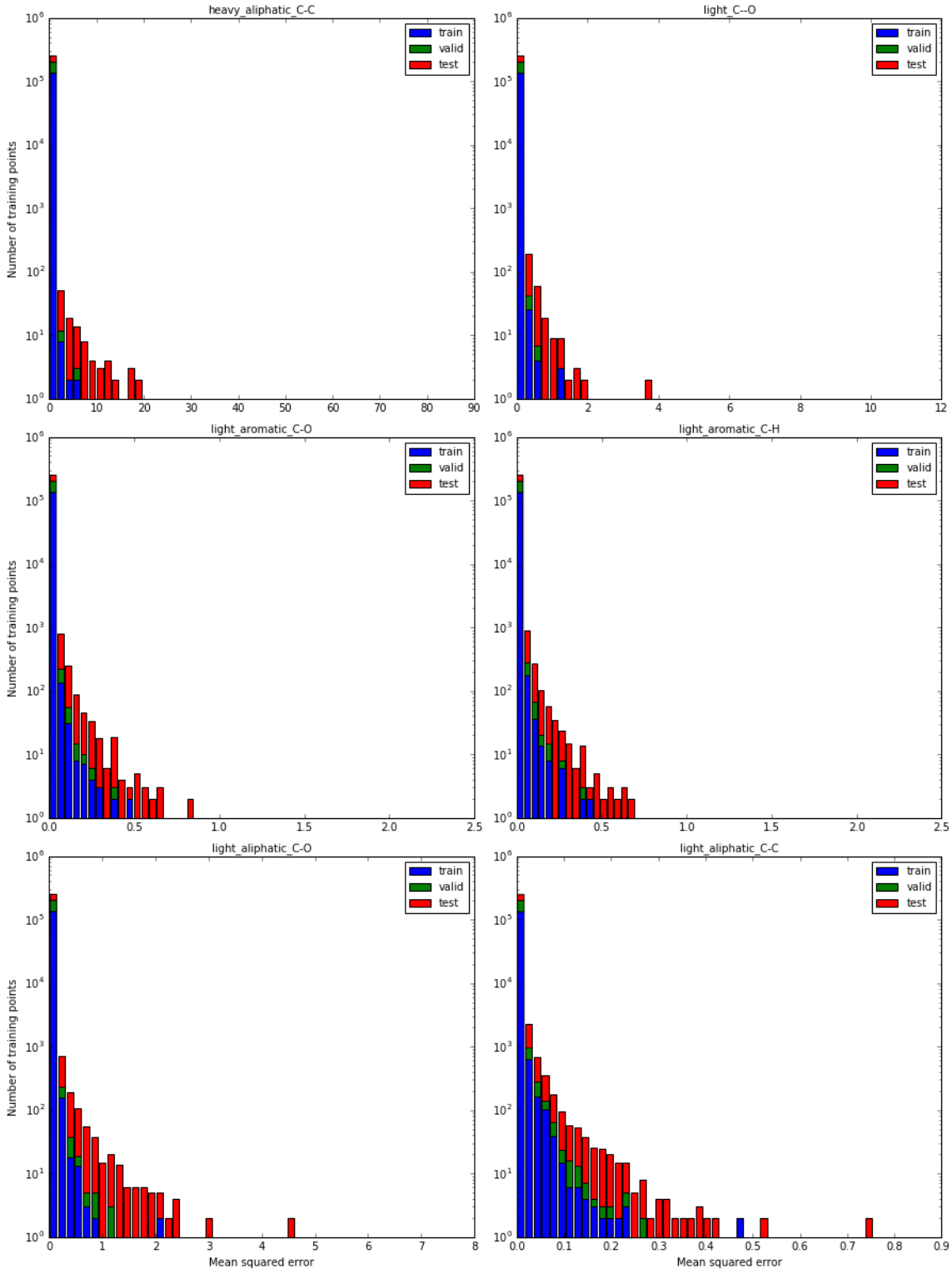
Supporting Figure C.6: Histogram of mean squared errors for the decision tree.




othergases
totalgases
H20
H-C-ratio
Moisture
tot_C--O
train
valid
test
Number of training points


tot_aromatic_C-O
tot_aromatic_C-C
tot_aromatic_C-H
tot_aliphatic_C-O
tot_aromatic_methoxyl
tot_aliphatic_C-C
train
valid
test
Number of training points


heavy_C--O
heavy_aromatic_C-O
heavy_aromatic_C-C
heavy_aromatic_C-H
heavy_aliphatic_C-O
heavy_aromatic_methoxyl
train
valid
test
Number of training points


heavy_aliphatic_C-C
light_C--O
light_aromatic_C-O
light_aromatic_C-H
light_aliphatic_C-O
light_aliphatic_C-C
train
valid
test
Number of training points
Mean squared error

APPENDIX D

IMPLICATIONS OF HOLISTIC MANAGEMENT ON BIOMASS RESIDUES FROM THE FORESTLANDS OF THE CONFEDERATED SALISH AND KOOTENAI TRIBES

Blake R. Hough[a] – designed and performed research, analyzed data, wrote the paper
Tom Richards[b]* – performed research
Jim Durglo[c] – designed research, contributed data
Laurel James[d] – designed research
Cody Sifford[e] – analyzed data
Daniel T. Schwartz[a] – designed research, edited the paper

* Deceased before submission of this paper
[a] Department of Chemical Engineering, University of Washington, Seattle, WA 98195
[b] Northwest Management, Inc., PO Box 9748, Moscow, ID 83843
[c] Confederated Salish & Kootenai Tribes Forestry Department, 104 Main St. SE, Ronan, MT 59864
[d] Interdisciplinary PhD Program (School of Environmental and Forest Resources and the Department of Anthropology), University of Washington, Seattle, WA 98195
[e] School of Environmental and Forest Resources, University of Washington, Seattle, WA 98195

Abstract

Ecologically-based active forest management is practiced in Confederated Salish and Kootenai Tribal (CSKT) forests to reduce the risk of wildfire and to transition forests toward pre-European conditions. Selling biomass residues, which are generally part of the waste stream, for biofuels and other products may help offset the costs of this active management. The supply and costs to process biomass residues on the CSKT reservation in Montana are estimated through a detailed landscape-level assessment based on high resolution continuous forest inventory data. We find that planned forest management activities on the reservation from 2013 to 2022 can produce about 11,900 Mg of residual biomass chips per year at an average delivered cost of $41/bone dry Mg (2014 dollars). Our estimates match regional estimates for conventional forestry practices when applied over large areas of the reservation, demonstrating that active management guided

by ecological and cultural goals can deliver the same volume of biomass residue as conventional treatment methods, while providing additional deliverables for the environment and community. At a stand level our landscape-based estimates offer a more detailed picture of the distribution of biomass residues than regional estimates, facilitating the development of more effective planning and management decisions.

Significance Statement

Forestry practices vary widely based on the management goals of the landholders. In the fire-prone dry forests of the Confederated Salish and Kootenai Tribes (CSKT) of western Montana, USA, managing for fire resilience and restoration to more historic structures are central goals, along with sustaining timber harvests, cultural preservation, and a wide array of ecosystem services. The extraordinary density of forest plot data available on the CSKT reservation enables a detailed landscape-scale analysis of the Tribes' holistic management practices and their impact for the biomass residues available for sale in emerging biofuel markets. The Tribes' holistic management practices result in the same yields of biomass as nearby forestlands managed for more conventional timber harvest.

Introduction

Fire exclusion policies, amplified by the effects of climate change, have led to significant changes in forest composition in the western United States. Many forests are transitioning from low density, fire-tolerant ecosystems to high density forests comprised of fire-sensitive species that burn easily and result in high severity fires.[125-129] In addition, rapid ongoing increases in tree mortality, likely driven by regional warming and consequent hydrologic changes, are further

stressing western forests and changing forest structure and composition.[130] One major consequence of these ecological changes is the dramatic increase in recent decades of high severity fires in the west, and with further warming these trends are expected to continue.[131-134]

Recovering a forest to maturity after high severity fire can take decades to generations, and studies suggest that declining forest resilience may lead to ecological regime shifts after such fires, with forests transitioning to semi-permanent shrub-/grass-dominated landscapes.[132, 135] This recovery period is especially hard on regional economies that depend heavily on their forests, as standing forests provide significant economic value through ecosystem services including erosion prevention, regulating local climates, carbon sequestration, timber production, recreation and eco-tourism, and the impact of scenic appreciation on real estate values.[136] In addition to the damage to local economies, the cost to fight wildfires in the United States is skyrocketing with 10-year average federal expenditures on fire suppression more than doubling over the last two decades to about $1.4 billion/year in the 2003-2012 period (2012 dollars).[137] We see this trend continuing; the 2015 fire season set new records for the number of hectares burned (4,097,502 ha in 2015 compared to an average of 2,600,350 ha/year between 2005-2014) and fire suppression costs, which exceeded $2.1 billion (2015 dollars). Shortfalls in wildfire suppression costs are often appropriated from other budgets, including the forest fuels management and restoration budgets that would help prevent fires in the first place.[138, 139]

The use of active forest management is one strategy for reducing the severity of future fires, and minimizing the damage to local people, economies, and forest structure.[140] About 27 million ha in the western United States are classified as burnable forests where active management should

be considered,[141] however the USDA Forest Service's over $300 million annual budget for fuel treatments only allows treatment of about 700,000 ha per year on average in recent years.[138] Active restoration and fuel reduction treatments are often prohibitively expensive, with estimates for mechanical treatments ranging from $620-6800/ha (2014 dollars),[142-144] thus addressing the restoration challenge at this cost necessitates extracting additional economic value if active management is to occur.

In order to make active management viable, landholders need to generate revenue that can offset the cost of management. In addition, many landholders have recently adopted forest management strategies guided by emulating more natural or historical disturbance regimes as a means to restore forest ecosystems and biodiversity in contrast to general prescriptions to simply reduce fuels and thin dry forests.[128, 145, 146] Studies have demonstrated that such ecological restoration can offer significant economic gains, especially if the value of healthy forests is considered.[135, 136, 147] Selling the non-timber biomass residues left behind after active management as feedstock for bioenergy projects is of particular interest to many landholders, as these residues are often burned to reduce potential fuel for wildfires, so are currently an expense rather than a revenue-generating stream.

Assessing the volume and cost of available residues is increasingly important for determining the viability of bioenergy projects, understanding the fire-related characteristics of forests to help in wildfire prevention and decision making, and studying the effects of changing forest conditions on the carbon cycle and global climate change. We assess the potential supply of woody biomass residues and their collection costs over the period from 2013-2022 for the Confederated Salish

and Kootenai Tribes (CSKT), a large landholder in western Montana with a unique set of economic, social, and ecological management goals. The CSKT have a tradition of active forest management and today are interested in alternatives to burning biomass residues. Their Forest Management Plan (FMP) directs management in a holistic way that balances specific cultural, spiritual, economic, social, and environmental goals.[148] To meet these goals, Tribal foresters follow an ecosystem-centered approach that differs from typical industrial forestry practices in western Montana, so regional conversion factors that estimate biomass volumes may be inaccurate at this scale and assessment of available biomass requires more detailed knowledge of the distribution and composition of forest resources that determine stand-level silvicultural treatments.

The extraordinary density and quality of tribal continuous forest inventory data enables accurate projections for biomass residues generated by tribal active management in different treatment areas across the landscape. In this study, we assess whether the available volume of forest residues generated using the holistic management practices of the CSKT differs from regional landholders with conventional industrial goals and management practices. This question is critical to understanding whether bioenergy development from forestry residues is compatible or incompatible with ecological and cultural constraints that may be imposed by specific landholders.

Results and Discussion

Interest in using forest residues to make biofuels and other products is increasing in response to new and existing federal policies, such as the passage of the Renewable Fuel Standard (RFS) in

the US Energy Independence and Security Act of 2007 that mandated the use of 136 billion liters of renewable fuels per year by 2022. Forest residues are expected to provide the largest initial contribution to meeting the RFS goals from woody biomass,[149] and subsequent executive and legislative actions at the federal and state levels have continued to demonstrate growing support for the use of biomass for bioenergy.[150] Despite the growing interest and support for bioenergy projects, most methods to assess the availability of biomass residues rely on regional estimates that make planning decisions by individual landholders difficult, particularly landholders like the CSKT whose management practices diverge from regional norms.

The CSKT foresters oversee 185,916 ha of forested land on the Flathead Indian Reservation, and actively manage about 119,000 ha as commercial forestland. An annual allowable cut of 42,711 Mm^3 is set forth in the FMP. Management prescriptions are guided by two ecological descriptors: fire regimes and seral clusters. Fire regimes describe the type of fire behavior that occurred on the landscape during pre-European times and reveal basic information about traditional uses and how the ecosystem functioned before fire suppression. Five regimes have been defined based on fire frequency, fire intensity, and the pattern of vegetation that fires create. Seral clusters describe the present structure and composition of the forest – the size and age of trees, how close they are to each other, whether stands are single- or multi-layered, and whether species are shade tolerant or intolerant. They also provide information about fire risk and severity, cover for big game, habitat for insects and birds, and risk of disease. Twelve seral clusters, A-L, are defined by the FMP. Silvicultural treatments are ecologically determined by the seral cluster and fire regime at the harvest location (**Supporting Table D.1**), and are subject to modification and final approval by the Cultural Preservation Office and Tribal elders.

Supporting Table D.1: Silvicultural treatment matrix for CSKT-managed forests. TPA = trees per acre, BA = basal area, Q = quotient of diminution

Seral Cluster	Fire Regime				
	Encroachment	**Non-Lethal**	**Mixed Severity**	**Lethal**	**High Elevation**
A_0, A_1	No Harvest	No Harvest	No Harvest	No Harvest	No Harvest
A_2	No Harvest	No Harvest	No Harvest	No Harvest	No Harvest
B & C	No Harvest	Pre-Commercial Thin, 300 TPA	Pre-Commercial Thin, 300 TPA	Pre-Commercial Thin, 300 TPA	No Harvest
D	No Harvest	Commercial Thin, 300 TPA	Commercial Thin, 300 TPA	Even-Aged Clear cut, 20% of area	No Harvest
E & I	No Harvest	Even-aged Clear cut, 20% of area	Even-Aged Clear cut, 20% of area	Even-Aged Clear cut, 20% of area	No Harvest
F	No Harvest	Uneven-aged Q of 1.1, to 45 BA	Uneven-aged Q of 1.1, to 60 BA	Even-Aged Seed Tree, 20% of area	No Harvest
G	No Harvest	Uneven-aged Q of 1.1, to 45 BA	Even-Aged SW, 25% of area (30 BA); Uneven-aged, 50% of area (60 BA)	Even-Aged Seed Tree, 20% of area	Even-Aged Clear cut, 20% of area
H	No Harvest	No Harvest	No Harvest	Even-Aged Seed Tree, 20% of area	Even-Aged Clear cut, 20% of area
J	No Harvest	Thin from below to 70 BA	Uneven-aged Q of 1.1, to 70 BA	Even-Aged Clear cut, 20% of area	Even-Aged Clear cut, 20% of area
K	No Harvest	Thin from below to 80 BA	Uneven-aged Q of 1.1, to 80 BA	Even-Aged Clear cut, 20% of area	No Harvest
L	No Harvest	No Harvest	No Harvest	Even-Aged Clear cut, 20% of area	No Harvest

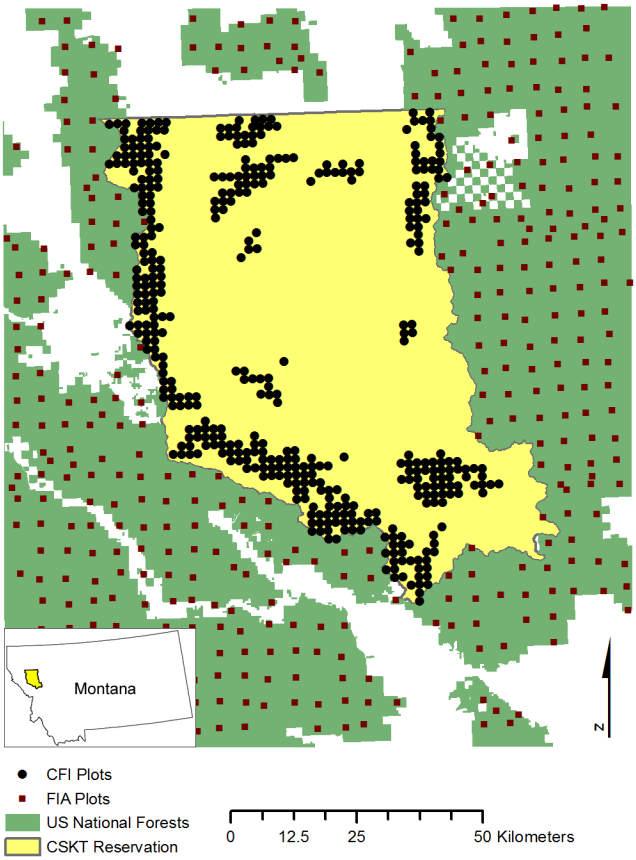
This oversight by the Cultural Preservation Office and the Tribal elders is one of the most significant differences between Tribal and conventional forest management. Proposed undertakings on the Reservation that might affect cultural resources (religious sites,

archaeological resources, burial sites, human skeletal remains, traditional cultural properties, historic resources, cultural items, food and medicinal plants, hunting areas, campsites, etc.) require a permit from the Cultural Preservation Office before any expenditure of funds or the issuance of any license.[148, 151] Permit applications are reviewed by the Tribes' Preservation Office, which investigates literature and oral history, and conducts field surveys when making permitting decisions. For forest management activities, Tribal elders often accompany Preservation Office staff on these field surveys of the management site. Elders provide feedback on what they remember about the area during site visits and how to best manage the location to meet specific lifestyle and cultural values. This feedback might be communicated to Tribal foresters as, *"this doesn't look like something you can walk through"* [implications for hunting/gathering], or, *"this used to be the site of a major encampment for people traveling from A to B"* [hence, a cultural site]. Guidance from Tribal elders generally leads to more active management and controlled burning than would occur in a similar, non-tribal forest in an effort to restore or maintain historical conditions that were changed by fire exclusion or former BIA management policies.

Management decisions are also impacted by other cultural values independent of the Preservation Office. For example, management activities are seasonally restricted in certain areas while culturally significant animals like elk or bear are present, then resume when those animals move out of the area for the season.[148] Thus, the biggest difference between Tribal and conventional foresters is not the silvicultural tools they use, but how they see the forest and value its cultural and ecological features. Rather than seeing the forest as a fuel model or for the value of its trees, Tribal managers at the CSKT see it in terms of how people and animals interact with

it and manage accordingly. In higher elevations where there is less traditional use, forest management decisions tend to be more focused on forest production instead of cultural uses. In light of their unconventional management approach, we expect that biomass assessments based on regional average data may not be accurate for CSKT-managed forests.

Because of these specific holistic management practices, we use detailed landscape-level information to assess the available supply of biomass from planned CSKT management activities over a ten year period, in place of regional estimates of biomass, and estimate the cost to transport it to a central site in Pablo, MT for introduction into a bioenergy supply chain. Our method estimates biomass from the individual tree species distribution and harvest at the stand level. We are able to achieve this resolution because of the density and quality of Tribal continuous forest inventory (CFI) plots (**Supporting Figure D.1**). CFI is a forest sampling system that monitors trends in forest growth, health, structure and composition by decennially revisiting 0.08 ha plots on Reservations across the country. CFI has about one plot per 160 ha, or 15 times denser than the forest inventory analysis (FIA) plots in the surrounding federal forests. Access to this rich information source means this study has access to a level of forest inventory data one might expect from an industrial forest, but which is treated following holistic management goals.



Supporting Figure D.1: CFI plot locations on the Flathead Indian Reservation shown with approximate locations of USDA Forest Service FIA plots in adjacent National Forests. CFI plots are located only in forested areas. The much greater density of CFI data compared to FIA data allows us to model forest resources at a sub-stand level to generate more accurate local biomass estimates than would be possible with Federal datasets.

The projected volumes of harvested timber and recovered residues, along with the cost of delivery to Pablo, are shown in **Supporting Table D.2** for each harvest area and year. We found that under their planned harvest schedule CSKT foresters can collect an average of 11,900 bone dry Mg of residual biomass chips per year at an average delivered cost of $41/bone dry Mg (2014 dollars). Important assumptions are described in the methods section.

Supporting Table D.2: Timber harvest and recoverable slash volumes in each harvest location for each year. Delivered cost includes chipping, loading, and transportation from landings to Pablo, MT.

Year	Harvest Location	Harvested Area (ha)	Timber Harvest (Mm^3)	Recovered Residues (bone dry Mg)	Delivered Cost ($/bone dry Mg)
2013	Hellroaring	1,765	20,583	5,667	27
	Jette	1,322	11,650	3,206	29
	Rattle Snake	443	4,819	1,319	35
	Sullivan	941	5,657	1,552	39
2014	Deep Draw	1,123	17,899	4,971	34
	Dry Fork	381	1,711	482	50
	North Buffer Zone	1,135	20,814	5,759	22
	Stevens	417	2,286	635	50
2015	Ferry Basin	1,909	17,089	4,719	34
	Skunk	948	10,122	2,775	53
	Yellow Bay	1,364	15,499	4,261	35
2016	Central Buffer Zone	923	14,767	4,091	28
	Delaware	846	9,034	2,490	36
	Eva Paul	452	3,358	929	40
	Revais	1,424	14,199	3,928	36
	Sheep Springs	183	1,353	382	39
2017	Irvine	1,038	7,242	2,047	41
	Moss Peak	1,173	20,272	5,631	27
	Pistol Creek	1,360	15,196	4,184	40
2018	Lamoose	623	5,697	1,554	36
	South Buffer Zone	669	11,364	3,167	35
	Sunny Slope	1,217	17,216	4,791	29
	Welcome Springs	662	8,433	2,371	40
2019	Boulder	2,109	25,151	6,911	30
	Meadow	1,169	13,256	3,717	39
	Saddle Mountain	559	4,303	1,227	47
2020	Charity Peak	2,421	30,067	8,453	40

	Schley	746	9,538	2,653	38
2021	Dog Lake	602	5,095	1,404	46
	Magpie	1,895	23,708	6,650	39
	Yellow Bay	1,364	13,907	3,838	32
2022	Seepay-Vanderburg	3,771	42,711	13,271	39

Decisions by the Cultural Preservation Office and feedback from Tribal elders determine exactly which stands in a harvest area are treated, but these decisions are not made years in advance as would be required to exclude off-limits stands from our landscape models. Projects that will adversely affect cultural, historical, or archaeological sites are only approved in extraordinary circumstances, and thus culturally sensitive sites are rarely treated mechanically. Natural processes like fire are allowed to pass through the sites, but people and machines are not allowed. For sensitive sites, forest managers routinely establish large, irregularly shaped buffers around the site to avoid identifying their locations.

These buffer zones are undisclosed, and thus not directly excluded from our estimates of the total available biomass. To address this uncertainty, we have looked in the harvest areas for each year to ensure that the available volume of harvestable timber exceeds the Tribes' annual allowable cut. We find that, on average, only 64% of the acreage in a named harvest area is necessary to meet the annual allowable cut limits. Based on this we are confident that our models have sufficient undisturbed areas within each harvest area to accommodate unknown cultural buffer zones which may be discovered during the tribally-mandated planning process.

Our estimates of the available biomass residues at the CSKT show good agreement with estimates made using a regional conversion factor of 421.3 kg residue/m^3 of wood delivered to a

mill in Montana[152] (**Supporting Table D.3**) when applied to annual totals, which cover multiple harvest locations. This is noteworthy because the CSKTs' forest deliverables and the goals that dictate their silvicultural practices and harvest strategies differ significantly from the regional norm, yet result in the same overall estimated volume of biomass residues. Within individual harvest areas, estimates using our method differ by up to ±12% from estimates based on the regional conversion factor. This suggests that when estimates for detailed distributions of biomass across a landscape are desired the use of regional conversion factors may be inadequate. Differences from regional averages are caused by a combination of stand-level variations in seral cluster, and the application of multiple treatment prescriptions within a given harvest area.

Supporting Table D.3: Comparison of annual totals from our calculated mass of available residue with estimates using a regional conversion factor.

Year	Residues predicted using our method (bone dry Mg)	Residues predicted using general MT factor (bone dry Mg)
2013	11,745	11,698
2014	11,846	11,698
2015	11,754	11,698
2016	11,819	11,698
2017	11,861	11,698
2018	11,883	11,698
2019	11,855	11,698
2020	11,106	10,847
2021	11,891	11,698
2022	13,271	11,698

A key question is whether emerging biofuels markets can reduce the cost of active management through the sale of biomass residues that otherwise cost money to gather and burn. The answer to this question depends on many factors, most important among them are the transportation costs, and the value of biomass as feedstock for biofuels or other products. In one detailed case study, collecting biomass residues for biomass energy projects cost less than conventional pile and burn

treatments when travel distances were less than 80 km in most treatment units, though for transport distances over 200 km conventional treatments cost less.[153] Unfortunately there is little published in this area, as functioning biofuel supply chains that utilize residual forest biomass are scarce at present.

Avoided economic, social, and environmental costs of adopting active management strategies are also worth considering. The increasing severity of fires, as well as the burning of biomass in slash piles, directly impact human health. Wildfire smoke contains high concentrations of particulate matter, exposure to which has been shown to result in significantly increased hospital admissions during wildfires for acute respiratory illnesses[154-156] as well as increases in out-of-hospital cardiac arrests.[157] Using biomass residues for energy instead of burning them in slash piles almost entirely eliminates their emissions of particulates, and significantly reduces NO_X, nonmethane volatile organics, CO, and CO_2 emissions.[158] From an environmental perspective, the idea that fuel thinning treatments provide a carbon benefit by removing small diameter forest trees in order to protect the remaining large trees from burning is still debated in the literature, but regardless of the conclusion it is likely that using residues for bioenergy will improve the carbon benefits of thinning.[159, 160] In fact, fuels made of biomass from sustainably managed forests have been shown to have significantly lower carbon and environmental footprints than fossil fuels.[161, 162]

Conclusions

Active forest management is one strategy for reducing the severity of wildfires and transitioning forests to desired ecological and recreational conditions. Growing interest in offsetting the cost

of active management through the sale of biomass residues for biofuels or other products is prompting landholders to assess their potential supply of biomass residues. We found that planned active management treatments in the CSKT managed Tribal forests can produce an average of 11,900 bone dry Mg of residual biomass chips per year at an average delivered cost of $41/bone dry Mg (2014 dollars). Our estimates match those produced from regional conversion factors for conventional forestry practices when applied over large areas of the Reservation, demonstrating that active management guided by the Tribes' specific ecological and social goals can deliver the same volume of biomass residue as conventional treatment methods, while providing additional deliverables for the environment and community. At a stand level our landscape-based estimates offer a more detailed picture of the distribution of biomass residues than regional estimates, facilitating the development of more effective planning and management decisions.

Materials and Methods

Most biomass assessment strategies that quantify available forest residues rely on regional data from mills or other wood processing facilities, or from the USDA Forest Service Forest Inventory and Analysis program data (FIA, www.fia.fs.fed.us/tools-data). FIA data characterizes forests in the United States using aerial and satellite imagery, as well as field measurements at long term forest plots (about one plot per 2500 ha). Information from mills or FIA is generally used to determine the volume of timber and typical stand characteristics (forest structure, mix of species, tree sizes, etc.) within a county or state. A software program uses these stand characteristics to simulate forest growth and predict future conditions. Then the volume of residual biomass residues can be determined using regional conversion factors for volume of

slash per volume of delivered timber,[152, 163-165] or using allometric equations based on characteristics of each harvested tree,[166] or characteristics of the entire stand.[167] In rare cases biomass assessments use alternative strategies such as calculating the total volume of forest residue based on a percentage of the annual allowable cut,[168] or taking actual measurements of the forest, and slash piles, pre- and post-harvest across a harvest site.[169, 170]

When modeling Tribal forest resources we used data from 396 Continuous Forest Inventory (CFI) plots from across the managed areas in our study (**Supporting Figure D.1**). CFI is a forest sampling system that monitors trends in forest growth, health, structure and composition by decennially revisiting 0.08 ha plots on the Reservation. The characteristics of every tree within each CFI plot are recorded and used as input for simulating tree growth in harvest areas for future years. 1999 CFI data is used in this work because 2009 data was not yet available digitally.

Growth and projected harvest from each CFI plot for 10 years was simulated using the Forest Vegetation Simulator (FVS, www.fs.fed.us/fmsc/fvs/). Key input settings, defect and breakage ratios, and version information are described in the Supplementary Information. FVS output for CFI plots of the same seral cluster was combined to produce average values for each seral cluster, and harvest volumes were determined for possible silvicultural treatments (**Supporting Table D.1**). These simulations predict the volume of merchantable wood per hectare, broken down by tree species, delivered to a mill for all possible silvicultural practices from stands of each seral type over the 10 year harvest schedule.

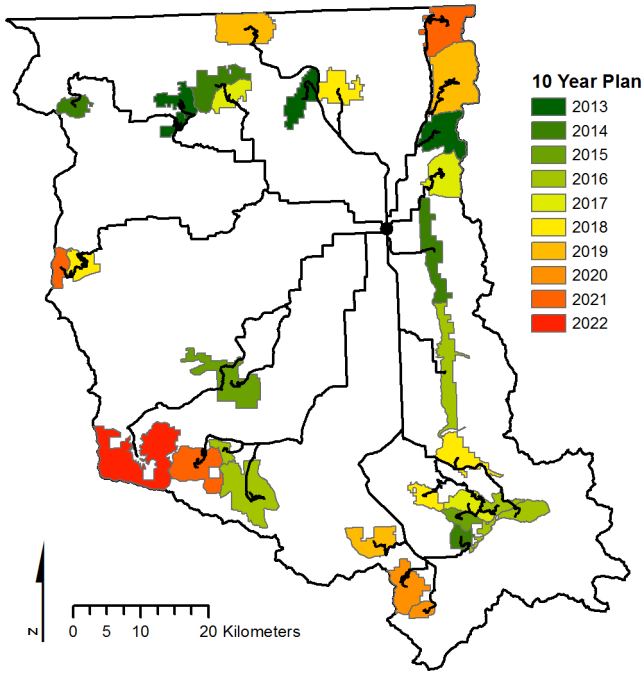
The volumes of forest residues left on the ground after a simulated treatment were estimated using the Resources Planning Act Timber Products Output database (RPA TPO, www.fia.fs.fed.us/program-features/tpo/). This database correlates the volume of merchantable wood of a particular species that is delivered to a mill with the volume of forest residue left on the ground at the harvest site (**Supporting Table D.4**). The residue volume estimated here includes tops, limbs, and any defect or breakage left on the site. Two important limitations to this definition must be noted: [1] pulp wood is considered a merchantable product and is not included in the residue, but in poor pulp markets it could add to the total volume, and [2] residues generated from thinning trees with a diameter at breast height smaller than 17.8 cm are not included in this estimate because they are not delivered to a mill. We limited the RPA TPO data to mills from western Montana to most accurately reflect regional characteristics of harvest sites.

Supporting Table D.4: Volume of residue remaining on the landscape per cubic meter of lumber delivered to a mill. Residue values are from the RPA TPO database and limited to information from mills in Western Montana.

Species	FIA Species Code	Residue (m^3)/m^3 lumber delivered to mill
Western Red Cedar	242	1.133
Ponderosa Pine	122	1.041
Whitebark Pine	101	0.887
Engelmann Spruce	093	0.834
Aspen	746	0.858
Douglas Fir	202	0.800
True Firs (Grand & Subalpine)	017, 019	1.009
Western Larch	073	0.779
Lodgepole Pine	108	0.840

The CSKT provided GIS layers of their scheduled forest management areas from 2013-2022 (**Supporting Figure D.2**), their logging road network, and the distribution of seral clusters and fire regimes. The road network layer contains information about the quality and safe speed of

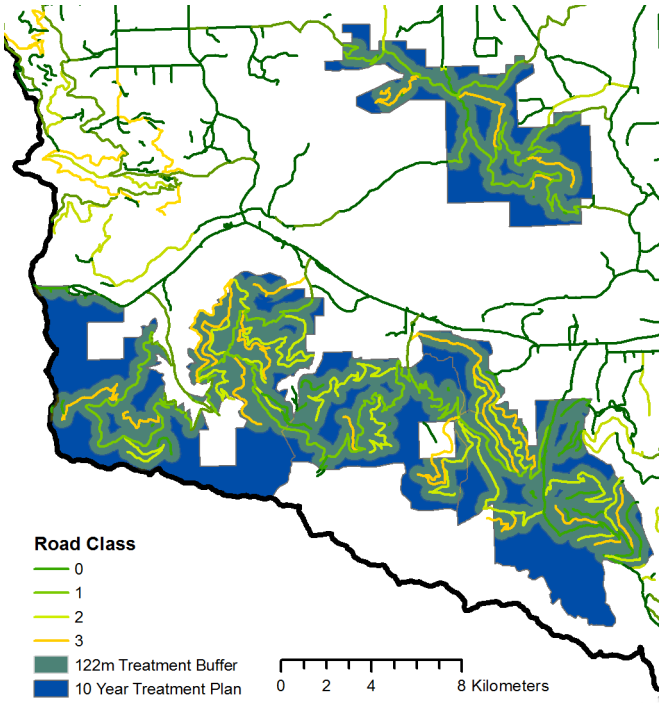
each road, allowing us to determine which roads are accessible to the chip vans best suited to removing chipped forest residues from harvest areas. A 122 meter harvest buffer around each accessible road defines the average distance from a road that harvest operations occur under current practices (**Supporting Figure D.3**). Computing intersections of the 13 seral clusters and five fire regimes produced 50 new combined seral cluster/fire regime designations (**Supporting Figure D.4**), and overlaying this layer with the harvest buffer allows extraction of the number of hectares of each seral cluster/fire regime designation within each harvest area by year. All GIS analysis was performed in ESRI's ArcMap version 10.2.



Supporting Figure D.2: Scheduled forests management areas in CSKT-managed forests from 2013-2022. Road networks connecting the centroid of each management area to a potential biomass processing site in Pablo, MT are shown and used to calculate transportation costs.

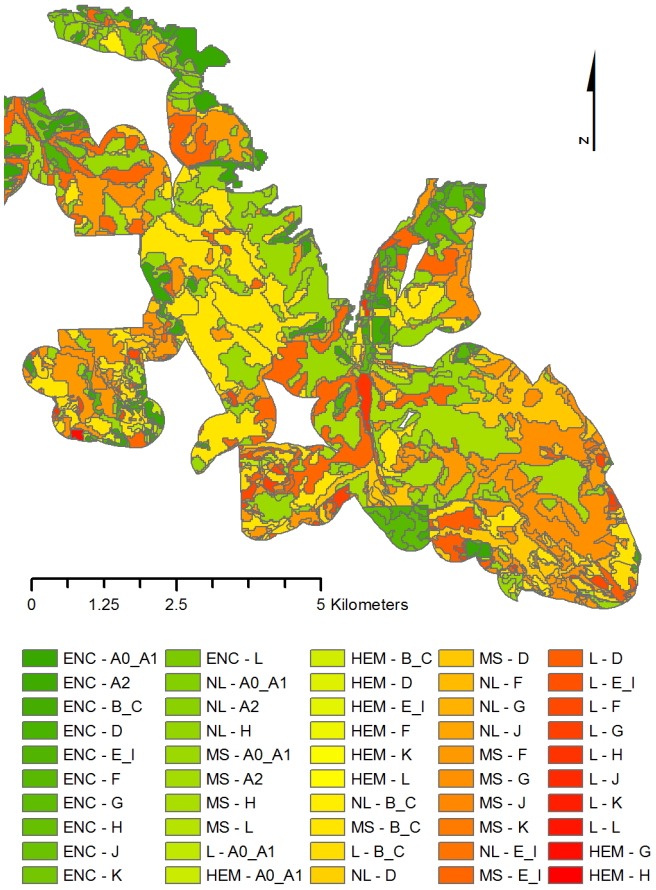
Each seral cluster/fire regime designation is associated with a particular estimate for the volume of forest residues/ha. To calculate the volume of residues available in a given harvest area for a

specific year we multiply the hectares of each seral cluster/fire regime designation within the harvest area by the appropriate residue volume/ha estimate, and by a recovery factor of 0.65. The recovery factor is an estimate for the percentage of total residue that is recoverable for chipping and removal from the forest. A value of 65% was chosen as a conservative estimate based on discussions with foresters throughout the region, however we expect that the actual recoverable fraction may be much higher because harvest operations in CSKT forests often use whole-tree to landing techniques. Residue volumes were converted to mass using a conversion of 480 kg/m^3 [150]. For most years the simulated available volume of harvestable timber exceeds



Supporting Figure D.3: Sample treatment area showing chip van-accessible roads (class 3 and below) and a 122 meter harvest buffer around these roads within planned treatment areas. Biomass resources within these treatment buffers were analyzed for this study.

the Tribes' annual allowable cut of 42,711 Mm^3, and for these years the treated fraction of each harvest area was reduced by the same fraction until the harvest volume restriction was met.



Supporting Figure D.4: Detailed map of the distribution of fire regime - seral cluster designations within a subset of the treatment buffer shown in Figure 3. Each of the 50 designations corresponds to a silvicultural treatment from Table 1. Fire regime – seral cluster colors indicate the "severity" of silvicultural treatments, ranging from green (minimal or no treatment) to red (even-aged clear cuts).

Cost estimates for chipping, loading, and transporting forest residues from harvest locations to a central processing site in Pablo, MT are based on the following assumptions provided by Tribal foresters:

- Landings are located every 0.4 km along logging roads
- Chip vans used have a capacity of 14.5 Mg
- The recoverable fraction of forest residues is deposited at the landing in the course of normal harvest operations
- The cost to chip and load slash at a landing into a chip van is $20.94/Mg
- The cost to transport chips from the forest to Pablo is a flat rate of $2.17/km
- The entire volume of slash from a given harvest area is evenly distributed over all landings in that harvest area (this results in an average of 36.3 Mg/landing)
- Chip vans pick up from only one landing before returning to Pablo, so they may not always be full

When calculating the cost of processing and transporting chips from a landing to Pablo the driving distance used was from the centroid of a harvest area to Pablo, MT USA.

Supplementary Information for Implications of holistic management on biomass residues from the forestlands of the Confederated Salish and Kootenai Tribes

FVS Settings

All simulations were run using FVS Version 0979 – Inland Empire Expanded variant. The starting year for each simulation is 1999 and the simulation end year is one year past treatment year. For example, for a treatment in 2014, the simulation would begin in 1999 and end in 2015.

Input Database – CSKT_Database

Total of 396 CFI Plots - 1/5th acre plots

- Seral Cluster "A" Plots – 140
- Seral Cluster "B" Plots - 23
- Seral Cluster "C" Plots - 46
- Seral Cluster "D" Plots - 7
- Seral Cluster "E" Plots - 29
- Seral Cluster "F" Plots - 57
- Seral Cluster "G" Plots - 27
- Seral Cluster "H" Plots - 16
- Seral Cluster "I" Plots - 1
- Seral Cluster "J" Plots - 37
- Seral Cluster "K" Plots - 8
- Seral Cluster "L" Plots – 5

Volume Settings

Defect – Cubic Foot – for all species:

- 5 inch trees = 0.05;
- 10 inch trees = 0.05;
- 15 inch trees = 0.07
- 20 inch trees = 0.07;
- 25 inch trees = 0.10;
- 30 inch trees = 0.11
- 35 inch trees = 0.12;
- 40 inch and larger trees = 0.13

Defect – Board Foot – for all species:

- 5 inch trees = 0.05;
- 10 inch trees = 0.05;
- 15 inch trees = 0.07
- 20 inch trees = 0.07;
- 25 inch trees = 0.10;
- 30 inch trees = 0.11
- 35 inch trees = 0.12;
- 40 inch and larger trees = 0.13

Board foot Volume Settings;

- ALL SPECIES (CODE = 0);
- MINIMUM DBH = 7.00;
- TOP DIAMETER = 4.00;
- STUMP HEIGHT = 1.00
- FORM CLASS = 80.00;
- METH OF VOL CALC = 6.

Prescription Settings

Prescriptions were applied based upon the Seral Cluster (forest type) and Fire-regime as described in Table 1 of the paper. Below are the general settings for each of the prescriptions.

Clearcut (CC)

- Reserve 5 TPA > 21”
- Applied to Seral Clusters D, E, G, H, J, K, L

PCT

- Low thin – thin to 300 trees per acre
- Applied to Seral Clusters B, C, D

Seed tree

- Residual of 10 TPA
- Applied to Seral Clusters F, G, H

Un-even aged (UA)

- Q quotient of 1.1 for all un-even aged scenarios
- Thinned to residual BA of 45, 60, 70. Depends on Seral Cluster
- Applied to Seral Clusters F, G, J, K

Commercial Thin (Thin)

- Applied to Seral Cluster J, K
- Thin from below to specified BA (70, 80). Depends on Seral Cluster

Shelterwood (SW)

- Residual of 30 TPA
- Applied to Seral Clusters G