Advancing Biomass Fractionation with Real-Time Prediction of Lignin Content and MWd: A kMC-based Multiscale Model for Optimized Lignin Extraction

Juhyeon Kim[a,b,*], Silabrata Pahari[a,b,*], Jiae Ryu[c], Mairui Zhang[c], Qiang Yang[d], Chang Geun Yoo[c], Joseph Sang-Il Kwon[a,b,**]

[a] *Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77845 USA*
[b] *Texas A&M Energy Institute, Texas A&M University, College Station, TX 77845 USA*
[c] *Department of Chemical Engineering, State University of New York College of Environmental Science and Forestry, Syracuse, NY 13210, USA*
[d] *School of Packaging, Michigan State University, East Lansing, MI 48824, USA*

Abstract

Recently, lignin has garnered significant research attention due to its abundance in nature. However, lignin is viewed as a recalcitrance factor as it impedes the overall biomass fractionation. In this regard, harsh operating conditions have been applied for the effective separation of the biomass components but they may cause substantial lignin degradation. Another problem is that the overall kinetics of lignin reactions remain limited since current models primarily focus on the cellulose fiber. These pose a challenge when developing effective fractionation strategies for industrial lignin extraction. To this end, we propose a multiscale model and develop a controller to determine the optimal operation strategy. In terms of lignin, delignification and de/repolymerization happen simultaneously but in different length and time scales. We adopted a bilayer structure of the ODEs and kinetic Monte Carlo (kMC) algorithm, accounting for the multiscale reaction kinetics. Our model provides the key outputs including the lignin content in the bulk chip and lignin molecular weight distribution, which were validated with the experiments. Subsequently, we developed a reduced-order model (ROM) for soft sensor design

*The authors have contributed equally.
**Corresponding author
Email address: kwonx075@tamu.edu (Joseph Sang-Il Kwon)

and formulated a model predictive controller (MPC) to determine the optimal operation strategy and then maximize the profitability.

Keywords: Lignin valorization; pulp digester; layered simulation; layered-kMC; multiscale modeling

1. Introduction

The global community has been striving to address environmental issues, notably the pursuit of carbon neutrality and the replacement of petrochemical resources [1]. In this search for alternatives, lignin has garnered significant attention due to its high energy density and aromatic nature, making it a large source of alternative fuels and chemicals [2, 3]. Moreover, with its abundance in nature and substantial production in many biorefinery processes, lignin has been investigated as a potential alternative resource [4, 5] to petrochemical resources, which face uncertainties in the future [6]. However, lignin is still underutilized primarily attributed to its poor processability stemming from its structural complexity [7]. Current industrial practices treat biomass under harsh conditions, often resulting in irreversible degradation and condensation of lignin [8, 9]. Consequently, the quality of the fractionated lignin deteriorates, impeding its potential for further utilization. Additionally, the extreme operating conditions, including high pressure and temperature, make the overall process energy-intensive, limiting economic benefits [10]. This often leads to the wasteful incineration of lignin as a low-value fuel, diminishing the feasibility of bioresources compared to petrochemical sources. Consequently, efficient lignocellulosic fractionation strategies have recently gained significant research attention, aiming to harness lignin-derived chemicals and enhance overall profitability [11].

Significant progress has been made in the study of lignin valorization; however, a knowledge gap persists in understanding the intricate dynamics of multiscale reactions and achieving the

optimal control over fractionation processes. The primary process outputs of interest are the remaining lignin fraction within the wood chips and the molecular weight distribution (MWd) of the fractionated lignin. While numerous researchers have delved into the investigation of lignin MWd during the de/repolymerization process through experimental studies [12, 13, 14, 15] and *ex-situ* MWd measurement [16], there remains a strong need for process control strategies beyond the measurement, incorporating explicitly defined process inputs [17]. In this context, it becomes imperative to consider comprehensive mathematical models that focus on the fractionation process, with particular emphasis on lignin.

Various modeling approaches have been employed to investigate the kinetics of the de/repolymerization reactions of lignin. One such approach involved the development of a kinetic model to describe the reaction pathway networks [18]. Although this model provided valuable insights into the monoaromatic units attainable through lignin depolymerization, it could not accurately describe the distribution of lignin chain lengths. To address this limitation, a population balance equation (PBE)-based model was developed for electrochemical lignin depolymerization [19]. This model accounts for the kinetics of de/repolymerization for each length of the lignin chains, enabling the successful tracking of the lignin chain length distribution over time. It is worth noting that to achieve a good agreement with experimental results, the reaction rate constants for de/repolymerization had to be considered as time-varying during the operation, although such variations are unlikely to occur in practical scenarios. To overcome this limitation, stochastic methods can be applied to describe the complexity of the system, in contrast to the deterministic approaches previously mentioned. For example, a study on lignin depolymerization was conducted employing a kinetic Monte Carlo (kMC) algorithm [20]. This study listed all potential reactions and their corresponding rates, using a probabilistic simulation of

depolymerization processes that factored in the distribution of reaction rates.

Despite recent successes in modeling various fractionation processes, the valorization of lignin is still in its early stages when compared to well-established cellulose-centered pulping processes like Kraft pulping. The latter has highly developed mathematical models, such as the Purdue model [21], along with subsequent extended models [22, 23, 24, 25, 26]. However, the absence of comprehensive models for lignin hinders our ability to control its properties during full-scale processes. This limitation restricts the potential for the successful commercialization of this innovative biomass fractionation concept. In particular, many lignin valorization strategies start with the solubilized lignins during the pre-processing, like kraft pulping. Therefore, these lignins undergo depolymerization as well as repolymerization. It is well noted that the molecular weight of lignin plays an important role in its post-applications [27, 28, 29]. To overcome this challenge, a next-level comprehensive model that goes beyond just de/repolymerization kinetics is necessary. More specifically, in the actual fractionation process, delignification from the bulk biomass first occurs, and dissolved lignin chains undergode/repolymerization. Therefore, a comprehensive model is needed to account for the multiscale nature of the entire fractionation process.

Motivated by the challenges discussed earlier, we have developed a multiscale model that effectively describes both macroscopic (delignification) and microscopic phenomena (de/repolymerization of detached lignin). To achieve this, we employed a hierarchical framework [30, 31] that utilizes ordinary differential equations for describing mass and energy balances associated with the detachment of lignins from biomass (i.e., delignification). Furthermore, we used a kinetic Monte Carlo (kMC) algorithm [32] to simulate de/repolymerization reactions of detached lignins. Our multiscale model successfully captures the intricate dynamics of lignin reactions and closely aligns with experimental observations. In order to determine an optimal operating strategy,

we also developed a model-based controller [33, 34] that drives both the biomass lignin content and MWd towards desired set-points. This process involved the identification of a reduced-order model, which was then employed to design a soft sensor, specifically the Kalman filter. The Kalman filter allows for real-time estimation of process variables not readily available, including the lignin content in the solid phase and the MWd of lignin in the liquid phase. These estimations are derived from accessible measurements such as system temperature. The capability to simultaneously measure the lignin content in cellulose-rich solid fraction and MWd of the solubilized lignin in real-time is of immense importance. This capability ensures the maximization of biomass utilization, aligning with the ideal biorefinery concept that aims to valorize both carbohydrates and lignin. It is to be noted that the developed multiscale model is based on fractionation experiments using phenol-4-sulfonic acid (PSA). PSA has exhibited remarkable delignification capabilities under mild conditions while maintaining a high degree of catalyst reusability [35]. We are confident that this advancement will shed light on pathways for further industrialization and enhance the overall value of biorefinery processes.

This article is structured as follows: Section 2 provides a brief overview of the experimental settings. Section 3 provides detailed information about the mathematical formulation, including the mass/energy balances of the system. This section also explains the working principle of the developed model, followed by model validation. In section 4, a reduced-order model is developed to reduce the calculation burden during the closed-loop control. Additionally, a soft sensor is designed to estimate unmeasurable process variables during the operation. Then, in Section 5, we describe the design of the MPC and highlight the results of the closed-loop control using this framework.

2. Experimental

2.1. Chemicals

The PSA (85 wt.%) was purchased from TCI America, Inc. (United States). All the other chemicals including acetone, ethanol, 72% sulfuric acid, acetic anhydride, and pyridine, used in this study were purchased from VWR Internation LLC. (United States).

2.2. Preparation of Aspen wood chips

The Aspen wood chips used in this study were provided by SUNY ESF. These air-dried wood chips were cut into pre-determined thicknesses of 0.8, 3.0, and 5.0 mm for all directions. For the preparation of 0.8 mm biomass preparation, the wood chips underwent Wiley-milling and were then screened with a 20-mesh sieve. In the case of biomass with thicknesses of 3.0 and 5.0 mm, the wood chips were manually trimmed using a cutter. Subsequently, the prepared chips were soaked in deionized water for 48 hours under ambient conditions, and their moisture contents were measured prior to the experiment.

2.3. PSA pretreatment of Aspen wood chips

The Aspen chips that had been pre-soaked in water were loaded into a 40 mL glass vial equipped with a pressure relief cap. Considering the moisture content, the dry weight of wood chips was 2.0 $\pm$ 0.3 g. To achieve a concentration of 72 wt.% PSA solution with the water-presaturated Aspen samples, 85 wt.% PSA along with additional deionized water were introduced to the 20 ml vial. The reaction vial was placed in an oil bath maintained at temperatures of 70 $\sim$ 90°C for a duration of 10 $\sim$ 30 minutes. Following the reaction, the softened Aspen wood chips were manually disintegrated using glass rods until jelly-like pulps were obtained. Subsequently, the mixture was then immersed in acetone for 20 minutes and subjected to filtration. The

resulting solid residue was washed with acetone and water until the surface pH reached 5. Both the solid residue and filtrate were collected and retained for further analysis. Notably, for Aspen samples with a thickness of 0.8 mm, the mixture was directly soaked in an acetone-water mixture (1:1, v/v) after pretreatment, without the need for manual disintegration.

The lignin was recovered using an Amicon® stirred cell (UFSC20001, Amicon Corporation) equipped with a 1 kD regenerated cellulose membrane disc. The black liquor was diluted 5 times with ethanol/water (1:1, v/v), and pH was controlled at 3.0 to 3.5 using 1 wt.% sodium hydroxide in ethanol/water (2:1, v/v) solvent. The residues on the membrane disc were collected and dried in air condition.

2.4. Klason lignin analysis

To quantify the lignin content of the wood chips both before and after pretreatment, the NREL standard procedure was followed [36]. Before analysis, the untreated wood chips were ground to a 14-mesh size. The unpretreated and PSA-pretreated biomass were then hydrolyzed with 72% sulfuric acid at a temperature of 30 °C for 1 hour. The resulting mixture was then diluted to contain 4% sulfuric acid by the addition of deionized water. Following this, the biomass and hydrolysate in 4% sulfuric acid were autoclaved at 121 °C for 1 hour. After the two-stage hydrolysis process, the insoluble samples were filtered, followed by washing with deionized water. Then, the ash content was measured using a muffle furnace at a temperature of 575 °C for 24 hours. The content of acid-insoluble lignin, excluding ash content, was used as Klason lignin content in biomass samples.

2.5. Molecular weight analysis

The recovered lignin in Section 2.3 was acetylated for its molecular weight analysis. About 2 mg of the lignin was dissolved and acetylated in a mixture of pyridine and acetic anhydride (1:1

v/v) for 48 hours. After this process, the solvent was removed using a rotary evaporator, and the resulting acetylated lignin was dissolved in THF. The weight-average molecular weight (M_w), and number-average molecular weight (M_n) were measured using an Agilent Gel Permeation Chromatography (GPC) SECurity 1200 system equipped with Waters Styragel columns (Waters Corporation, Milford, MA) and a UV detector at 270 nm.

3. Model formulation

In this work, we developed a kinetic model based on the outcomes of the biomass fractionation experiment. This model was subsequently used to predict and delve deeper into the reaction kinetics. The biomass treatment comprises two distinct processes operating at different time and length scales.

The first process involves the dissolution of lignin from the bulk biomass, referred to as delignification (i.e., a macroscopic reaction). The second process involves the de/repolymerization of the dissolved lignin chains in the solution (i.e., microscopic reactions). These multiscale reactions yield several critical process variables, including the lignin content remaining in the bulk biomass and the MWd of the dissolved lignin chains.

Initially, a simple mass balance equation was employed to monitor the macroscopic changes and determine delignification kinetics, based on the experimental findings. Furthermore, microscopic kinetics were also determined using the kMC approach to obtain the MWd for each reaction condition, thereby establishing the relationship between these conditions and the MWd. Considering that these two events are occurring concurrently but at different scales, we coupled these models by layering, ensuring the accurate formulation of the entire process. The detailed working scheme will be elucidated in the following sections.

3.1. Macroscopic model

To model the macroscopic phenomena of delignification, we employed continuum-scale mass and energy balance equations. These were utilized to model and simulate the process, and their predictions were validated against the results of the delignification experiment.

3.1.1. Mass balance: delignification

The delignification reaction is often modeled as a first-order reaction; however, in practice, redeposition of the dissolved lignin has been observed during experiments. Therefore, the simplest reversible reaction scheme was chosen among the available kinetic models [37]. In summary, delignification is modeled as follows:

$$L \underset{k_2}{\overset{k_1}{\rightleftharpoons}} L_s \tag{1}$$

where L and L_s represent the lignin content in the wood chip and liquor phases (*i.e.*, dissolved lignin), respectively. The rate constants for delignification and lignin redeposition are denoted by k_1 and k_2, respectively. Therefore, the rate equations can be expressed as follows:

$$\begin{aligned} r_L &= -\frac{dL}{dt} = k_1 L - k_2 L_s \\ r_{L_s} &= -\frac{dL_s}{dt} = -k_1 L + k_2 L_s \end{aligned} \tag{2}$$

The analytical solutions for both mass balances are provided below:

$$\begin{aligned} L(t) &= L(0)\left(\frac{k_2 + k_1 e^{-(k_1+k_2)t}}{k_1 + k_2}\right) \\ L_s(t) &= L(0)\left(\frac{k_1 - k_1 e^{-(k_1+k_2)t}}{k_1 + k_2}\right) \end{aligned} \tag{3}$$

Both rate constants can be expressed using the Arrhenius-type equation as follows:

$$k_i = A_i exp\left(-\frac{E_{a_i}}{RT}\right) \tag{4}$$

where A and E_{a_i} are the pre-exponential factor and the activation energy, respectively, R is the universal gas constant, and T is the system temperature. Additionally, $i = 1, 2$ designates the delignification and redeposition reactions. Aspen wood is classified as a hardwood species, and its lignin mainly consists of the cinapyl monolignols with a MW of 0.210 kDa [38]. In this study, lignin is modeled as a polymerized chain of single monomers, where the MW of the dissolved lignin chains is observed as 13 kDa, which corresponds to the degree of polymerization of 62 monolignols/chain. The dissolved lignin chains are subject to de/repolymerization. Therefore, it is coupled with the microscopic layer of this simulation, which is described in detail in Section 3.3.

3.1.2. Energy balance

In addition to the continuum-scale mass balance, the energy balance is also considered in this study. The system is divided into two phases: the chip phase, where the macroscopic reactions occur, and the free-liquor phase, where the microscopic reactions occur. The energy balance for the chip phase is expressed as follows:

$$\frac{dT_c}{dt} C_{P_c} M_c = \Delta H_R r_L + U(T_f - T_c) \tag{5}$$

where T_c stands for the chip phase temperature, while C_{P_c} and M_c are the specific heat and the mass for the chip phase, respectively. ΔH_R is the heat of the reaction, U is the overall heat transfer coefficient, and T_f is the temperature of the free-liquor phase. The specific heat of the chip phase is a function of temperature and can be expressed as $C_{P_c} = 0.1031 + 0.003867 T_c$

[kJ/kg K] [39].

For the free-liquor phase, the energy balance can be represented as:

$$\frac{dT_f}{dt} C_{P_f} M_f = -U(T_f - T_c) + C_{P_{ext}} \dot{M}_{ext} (T_{ext} - T_f) \tag{6}$$

where C_{P_f} is the specific heat, and M_f is the mass for the free-liquor phase. The latter term of Eq. 6 is utilized for temperature control purposes, which is achieved through an external heat jacket. Here, $C_{P_{ext}}$, $\dot{M}_{ext}$, T_{ext} represent the specific heat, mass flow rate, and temperature of the external flow that exchanges heat with the free-liquor phase, respectively. The mixing rule [40] is utilized to calculate the specific heat of the free-liquor phase, which is influenced by the amount of solid mass dissolved into the free-liquor phase.

$$C_{P_f} = x_{fs} C_{P_c} + x_{fl} C_{P_l} \tag{7}$$

where x_f represents the mass fraction of solid and liquid, and C_{P_l} denotes the specific heat of the pure-liquid.

3.2. Microscopic model: de/repolymerization

As previously mentioned, the number of dissolving lignin chains at a given time is determined using Eq. 2. Our model takes into account the further microscopic interactions these chains undergo. Specifically, it incorporates three types of microscopic events: (i) a random scission of a selected chain, which represents depolymerization; (ii) an end-to-end polymerization between two selected chains, signifying repolymerization; and (iii) any other events that do not affect the MW of existing chains. In the free-liquor phase, lignin molecules can be solubilized without any associated length changes. Such interactions between lignin fragments and PSA solvent are

classified as null events. Within our model, de/repolymerization kinetics are assumed to follow first and second-order reactions, respectively, while null events are classified as zeroth-order reactions.

At any given moment, each dissolved chain undergoes microscopic events at distinct reaction rates, subsequently updating the lignin population. Capturing these events mathematically is challenging due to the stochastic nature of the microscopic interactions. As a result, we employed the kinetic Monte Carlo (kMC) algorithm to simulate de/repolymerization. To simplify the process, we made the following assumptions: (i) Lignin fragments are considered linear polymer chains due to their less branched nature [5, 41]; (ii) All scission possibilities are identical for the constituent bonds in a given lignin chain [20]; and (iii) The MWd adheres to a specific distribution, such as log-normal, as reported in prior studies [12, 13, 14, 15]. For the calculations, zero, one, or two out of N chains can be chosen at any time, with the corresponding reaction rates defined as follows:

$$\begin{aligned} r_{dep}(N_i) &= k_{dep} C_L(N_i) \\ r_{rep}(N_i, N_j) &= k_{rep} C_L(N_i) C_L(N_j) \\ r_{oth} &= k_{oth} \end{aligned} \tag{8}$$

Here, N_i and N_j are the indices for the selected lignin chains, with the conditions $i \neq j$ and $1 \leq N_i, N_j \leq N$. Additionally, we define $C_L(N_i)$ as the concentration of all dissolved lignin chains that have a MW equal to that of the selected chain, N_i. Based on the actual experiments, this model aims to determine the rate constants for various processes: k_{dep} for depolymerization, k_{rep} for repolymerization, and k_{oth} for other events.

The parameters in the rate equations (i.e., pre-exponential factors and activation energies) for de/repolymerization, and other events are obtained by fitting the average molecular weights

Criteria	Event
$0 \leq \xi_1 \leq \frac{r_{dep}(N_i)}{r_{micro}}$	Depolymerization of N_i
$\frac{r_{dep}(N_i)}{r_{micro}} < \xi_1 \leq \frac{r_{dep}(N_i)+r_{rep}(N_i,N_j)}{r_{micro}}$	Repolymerization between N_i and N_j
$\frac{r_{dep}(N_i)+r_{rep}(N_i,Nj)}{r_{micro}} < \xi_1 \leq 1$	Null events

Table 1. The execution table of the microscopic events.

(i.e., M_n and M_w) predicted by the model with the experimental data. As seen in Eq. 4, the reaction temperature affects the reaction rates, while the chip size does not influence the microscopic kinetics. In this sense, the best-fit values of rate constants and the activation energy barriers were obtained by doing a grid search. With the rate constants determined, the kMC algorithm randomly selects a specific event based on the rate distribution, which can be described as follows. Since any lignin chain can be chosen, and the rates are provided in Eq. (8), the overall de/repolymerization rates can be expressed as:

$$\begin{aligned} r_{dep} &= \sum_{i=1}^{N} k_{dep} C_L(N_i) \\ r_{rep} &= \sum_{i=1}^{N} \sum_{j \neq i} k_{rep} C_L(N_i) C_L(N_j) \end{aligned} \tag{9}$$

The overall microscopic reaction rate ($r_{micro} = r_{dep} + r_{rep} + r_{oth}$) can be calculated, allowing for the determination of the microscopic reaction rate distribution. Utilizing this information, the kMC algorithm selects and executes an event based on Table 1, using a random number generated (ξ_1). Each event progresses the clock, and the advancement amount is calculated using another random number, $\xi_{\delta t}$, as follows:

$$\delta t = -\frac{\ln \xi_{\delta t}}{r_{micro}} \tag{10}$$

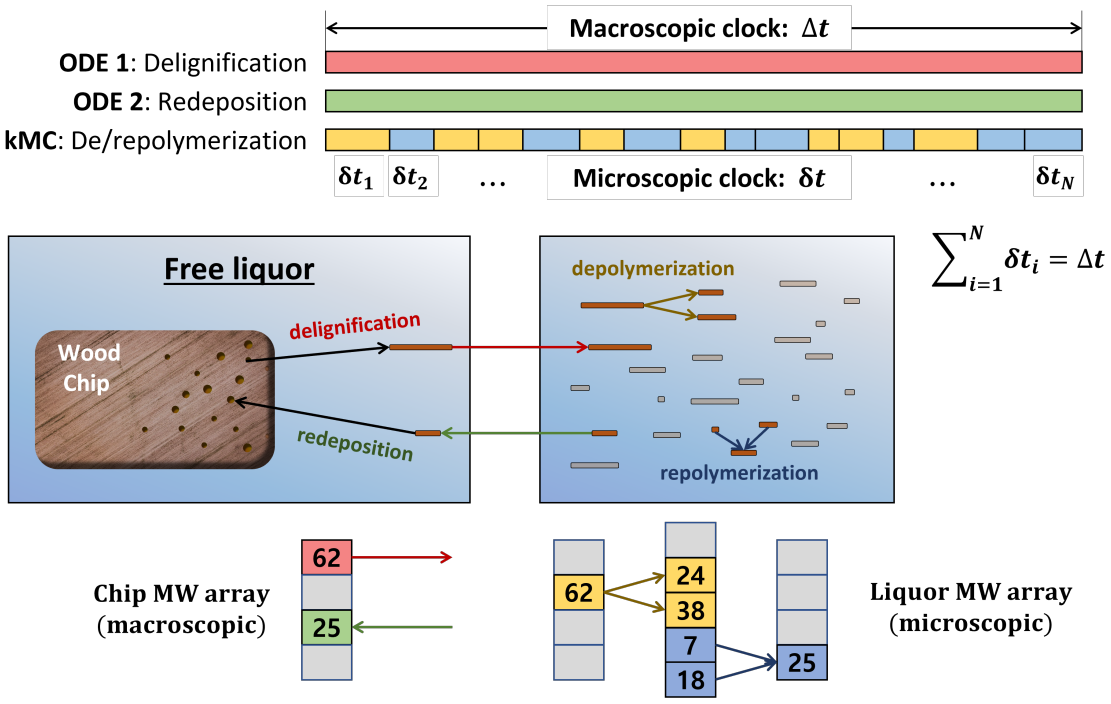

Fig. 1. A schematic illustration showing the lignin MW arrays and the reactions occurring at multiple scales.

3.3. Overall simulation scheme

As previously described, two layers of ODE (Section 3.1.1) and kMC (Section 3.2) are integrated and operate together to simulate the temporal evolution of lignin content in bulk biomass and the MWd of dissolved lignin chains.

Since dissolved lignin chains undergo de/repolymerization, the macroscopic mass balance is linked to the microscopic layer. To achieve this, specific arrays are implemented to store the MW information of lignin chains in both the wood chip and free-liquor phases, as illustrated in Fig. 1. When a lignin chain dissolves from the bulk wood chip, it is transported to the free-liquor phase (red arrow). Simultaneously, some dissolved chains reattach to the wood chip via redeposition (green arrow). The dissolved lignin chains then undergo de/repolymerization reactions (yellow and blue arrows), which are simulated in the microscopic kMC layer. The arrays are updated for each macro/microscopic time segment during the simulation.

For the macroscopic layer, k_1 and k_2 are used to fit the experimental data for 0.8/3.0/5.0 mm chips. With these values, the macroscopic mass balance is calculated discretely for each time step (Δt), which is set to 0.0005 min. The amount of lignin dissolved from the biomass at each Δt is determined to calculate the quantities of dissolved and condensed lignin. Consequently, lignin chains are either introduced to or removed from the liquor phase.

For the microscopic layer, three microscopic rate constants (k_{dep}, k_{rep}, and k_{oth}) are incorporated into the kMC algorithm. For each Δt, a microscopic event is selected and executed for δt, and this process is repeated until $\Sigma\delta t$ reaches Δt. De/repolymerization reactions break and recombine existing lignin chains in the solution, without altering the overall lignin masses in both phases. The entire cycle is then repeated for 30 min, as the experimental data are available for each condition up to 30 min.

After validating the model with the experiment data, the kinetic parameters are analyzed using traditional methods, such as plotting and examining them on $\ln k$ vs. $1/T$ plots. Subsequently, the kinetic parameter estimation is completed, which yields the activation energies and pre-exponential factors for delignification and de/repolymerization processes, which can be used for further applications in biomass fractionation.

3.4. Open-loop simulation results

3.4.1. Delignification kinetics

In this section, we present the results of the open-loop simulation for the delignification kinetics. To estimate the delignification/redeposition rate constants, we performed reactions under various conditions, and the results are displayed in Fig. 2. Based on these outcomes, we determined the delignification/redeposition rate constants, as shown in Fig. 3. With these results, we obtained the activation energies and pre-exponential factors, the values of which

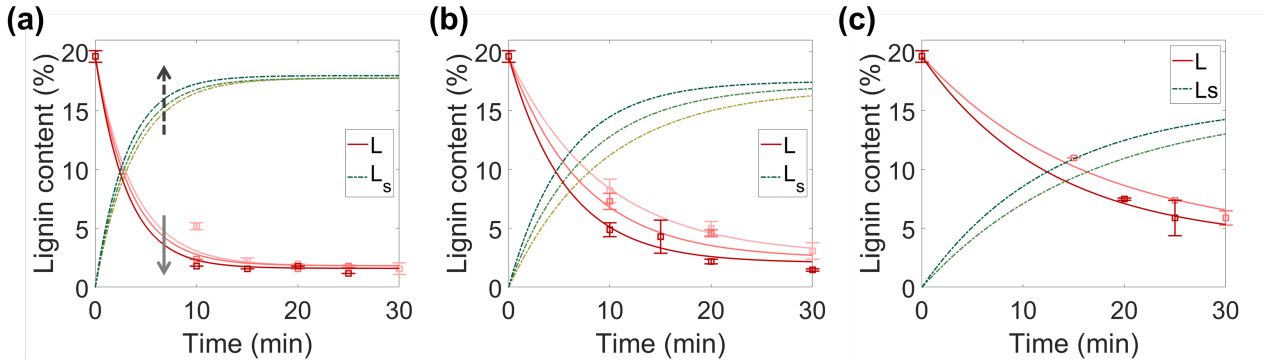

Fig. 2. The delignification results for different biomass sizes are presented, including (a) 0.8 mm, (b) 3.0 mm, and (c) 5.0 mm wood chips, under three different temperatures. The darker lines represent higher temperatures in the ascending order of 70, 80, and 90 °C. The solid and dashed arrows also indicate the increasing temperature for L, L_s, respectively.

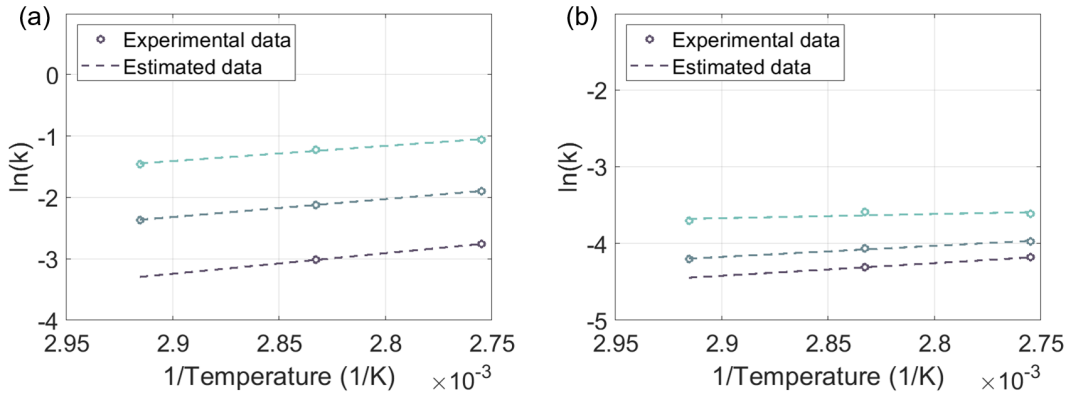

Fig. 3. The rate constants are plotted as a function of reaction temperature, with (a) representing delignification, and (b) redeposition. The data correspond to 0.8, 3.0, and 5.0 mm wood chips, displayed from top to bottom.

are presented in Table 2. It is important to note that while PSA generally shows superior fractionation performance, the 5 mm chips are considerably large, making it difficult to obtain meaningful delignification and de/repolymerization data at low temperatures, such as 70 °C.

In general, reaction rates increase with the temperature, and delignification is more favorable with smaller biomass sizes. This is because the solvent can more easily penetrate the biomass, and the dissolved lignin chains can be released into the free-liquor phase with less resistance. Consequently, the trends observed are reasonable, as delignification rate constants increase with higher temperatures and smaller biomass sizes.

Biomass size (mm)	E_1	E_2	A_1	A_2
0.8	20.50	4.482	311.7	0.1216
3.0	24.25	11.97	464.3	0.9989
5.0	27.84	13.66	644.6	1.410

Table 2. The macroscopic kinetic parameters.

It is also noteworthy that the activation energies we obtained were lower than those reported in previous studies, even under more extreme reaction conditions [42, 43]. This suggests that the PSA reagent can stabilize the dissolved lignin chains and play a critical role as a catalyst, allowing for effective delignification even under moderate conditions, such as lower temperatures and atmospheric pressure. This is in contrast to traditional methods like Kraft pulping.

3.4.2. De/repolymerization kinetics

In this section, we examine the kinetics of de/repolymerization of lignin, employing the kMC algorithm as our analysis tool. Our study is based on the macroscopic rate data we have gathered. To begin, we present the evolution of MWd as illustrated in Fig. 4.

As the reaction progresses, lignin chains continuously dissolve out from the biomass, resulting in an increase in the number of lignin chains. Interestingly, the emergence of a small peak at the onset of the reactions can be observed. This peak corresponds to chains of length 62 or

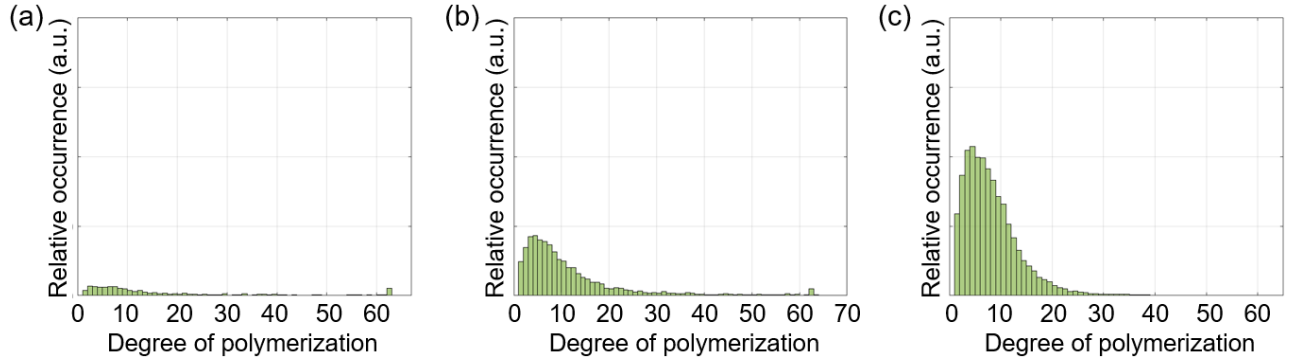

Fig. 4. The estimated DP distribution of the fractionated lignin in the liquor phase at selected time points (0.8 mm chip, 70 °C): (a) 1 min, (b) 5 min, and (c) 15 min following the initiation of the reaction.

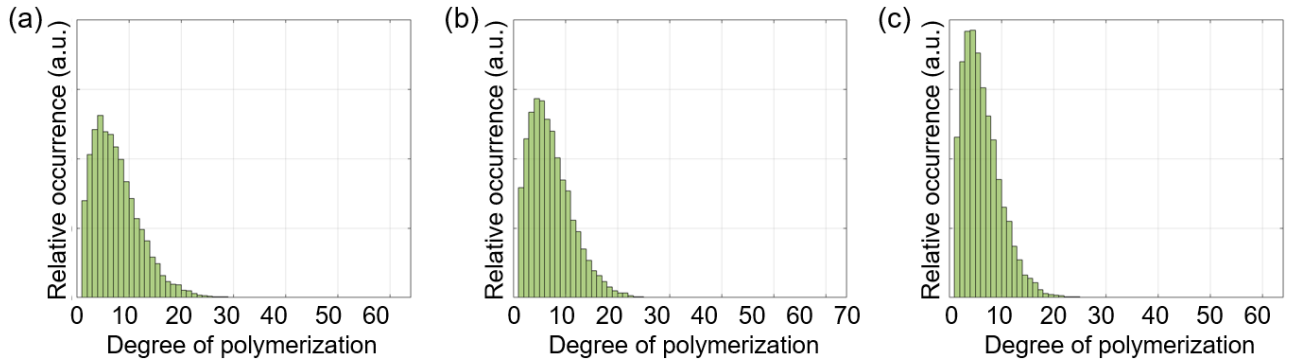

Fig. 5. The resulting DP distribution of the fractionated lignin in the liquor phase for the 0.8 mm chips after a reaction time of 30 minutes under varying temperatures, (a) 70 °C, (b) 80 °C, and (c) 90 °C.

a MW of 13 kDa. Such a phenomenon can be attributed to the significant reduction in lignin content within the bulk biomass.

As the reactions progress, the MWd tends towards specific log-normal distributions for each operating condition. This trend is presented in Fig. 5, and it aligns with the findings from prior studies [12, 13, 14, 15]. This observation suggests a higher dissolution rate of lignin chains at elevated temperatures.

With the MWd data, we computed the number-averaged MW (M_n) and the weight-averaged MW (M_w). Fig. 6 shows that the M_n and M_w values, when aligned with the estimated kinetic parameters, correspond well with the experimental results. This concurrence serves to validate our multiscale modeling framework. Furthermore, as depicted in Fig. 6, the average MW decreases as temperature increases. This trend can be attributed to the depolymerization reaction rate dominating over other microscopic reaction rates.

It is noteworthy that the initial fluctuations in the M_n and M_w become more pronounced as larger wood chips are employed. This phenomenon can be attributed to the interplay between macroscopic reactions, such as delignification, and microscopic phenomena. During the early stages of the fractionation process, fewer lignin chains are released from larger wood chips into

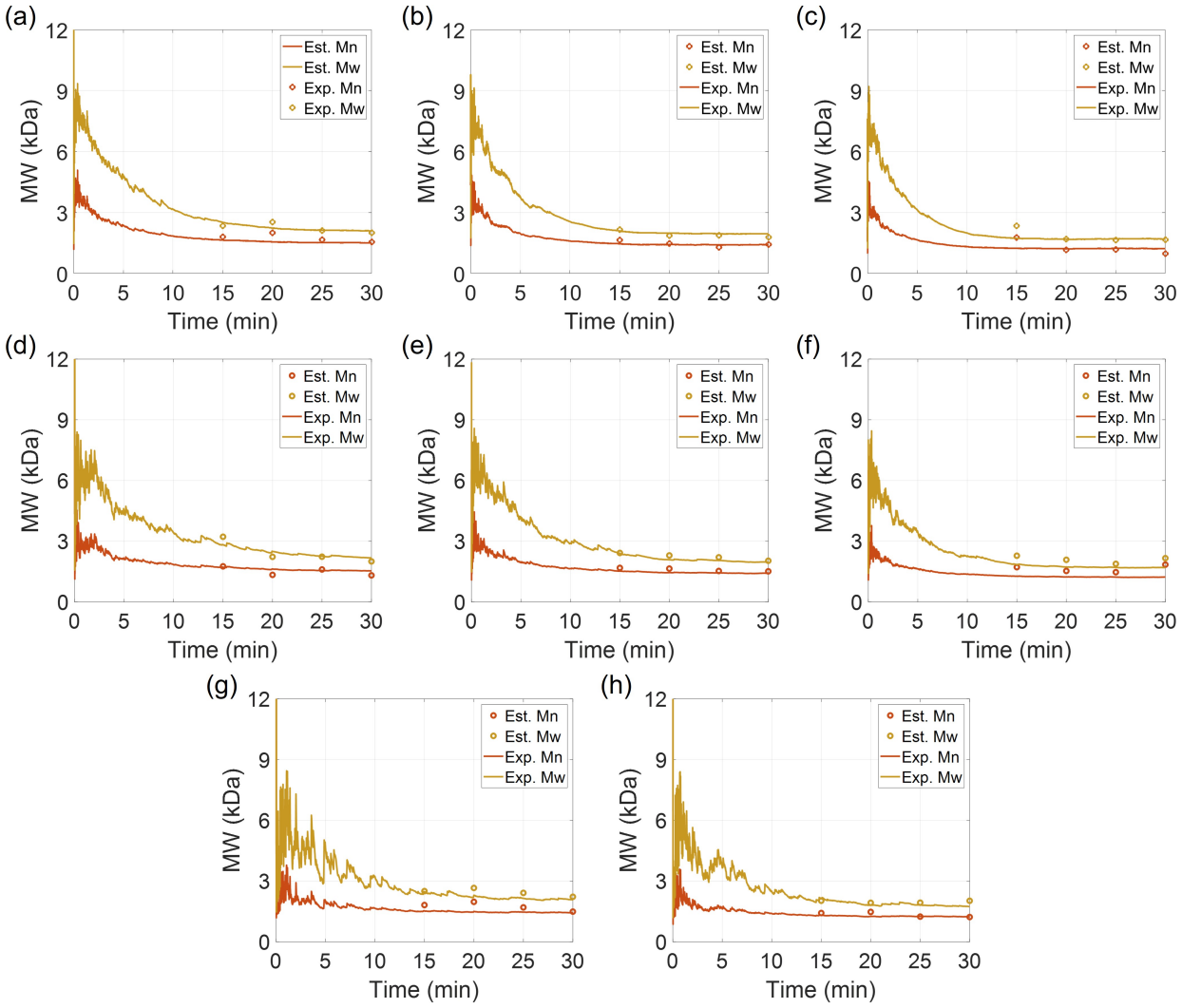

Fig. 6. The de/repolymerization results (MWs) under three different temperatures, each with varying biomass sizes: (a) 0.8 mm/70 °C, (b) 0.8 mm/80 °C, (c) 0.8 mm/90 °C, (d) 3.0 mm/70 °C, (e) 3.0 mm/80 °C, (f) 3.0 mm/90 °C, (g) 5.0 mm/80 °C, and (h) 5.0 mm/90 °C.

the liquor phase. Consequently, a small population of lignin chains undergoes rapid de/repolymerization reactions, leading to significant changes in the average MWs. Once the wood chips supply a sufficient number of additional chains, both the M_n and M_w stabilize, converging to their respective values.

The microscopic kinetic parameters, derived through the fitting model outputs to experimental results, are presented in Fig. 7. As predicted, the rate constants exhibit an increasing trend with escalating temperature. Importantly, it should be emphasized that the microscopic

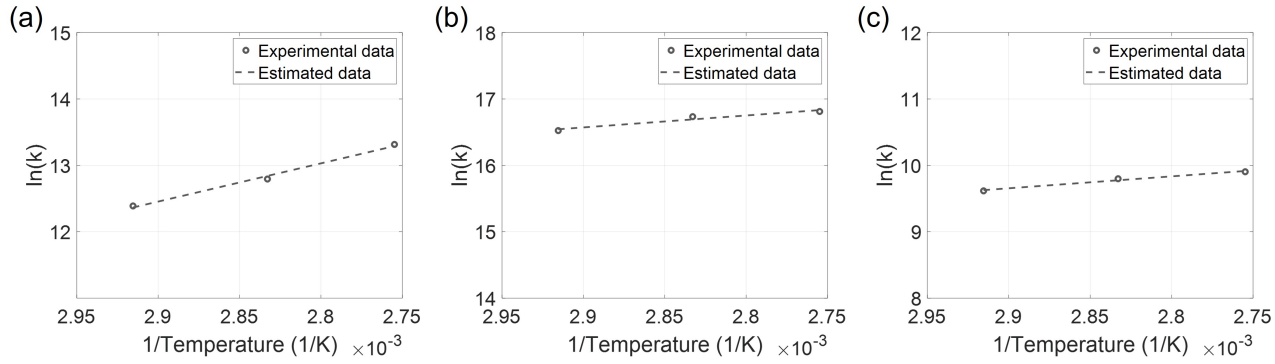

Fig. 7. Microscopic rate constants described as a function of reaction temperature for (a) depolymerization, (b) repolymerization, and (c) null events.

rate constants are solely influenced by temperature and remain unaffected by the size of the wood chips. This is attributed to the fact that the de/repolymerization reactions occur within the already-dissolved species situated in the liquor phase. In contrast, the kinetics of the delignification/redeposition reactions are directly impacted by the size of the wood chips. Consequently, the de/repolymerization rates are influenced by both temperature and the concentration of lignin chains (refer to Eqs. 8-9). The latter is determined by the quantity of lignin dissolved in the macroscopic layer. Ultimately, our high-fidelity model effectively encapsulates the multiscale nature of the biomass fractionation process through the integration of macroscopic and microscopic kinetics.

4. Model order reduction and soft-sensor design

4.1. Reduced-order model development

In the preceding section, we delved into the intricacies of the high-fidelity kinetic model. Our ultimate objective is to identify the optimal operational strategy for the system, one that allows us to achieve the desired lignin content and MWd. Nevertheless, the multiscale model, while comprehensive, poses a computational challenge for control purposes. Thus, we have

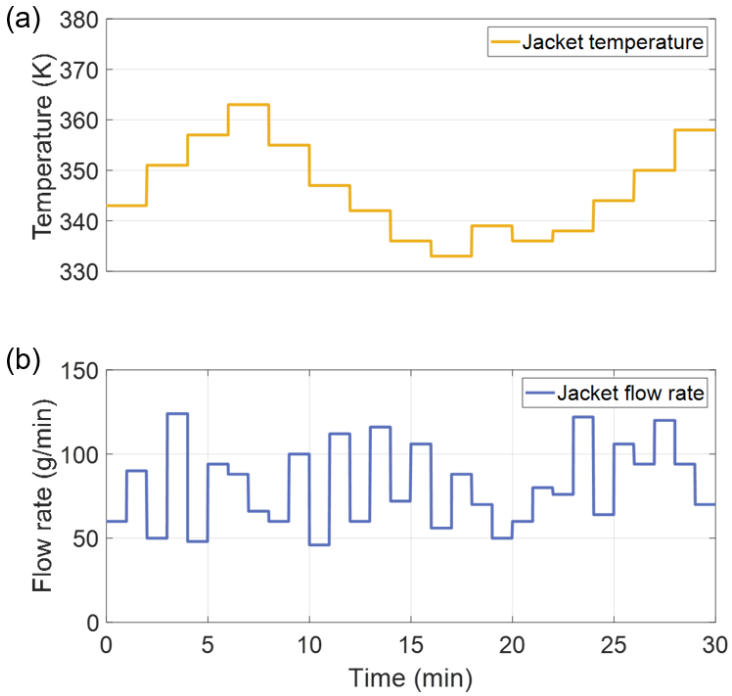

Fig. 8. Manipulated system inputs for ROM training: (a) the external jacket flow temperature and (b) the flow rate.

developed a reduced-order model (ROM) that is more computationally tractable. In this model, we propose the use of an external jacket to supply heat and control the system's temperature (Eq. 6). Our primary goal is to control the optimal operational conditions in order to achieve the desired outputs (i.e., lignin content and MWd). This is accomplished by adjusting the process inputs, namely, the external flow temperature and rate. We generate the ROM via the subspace state-space system identification algorithm, also known as N4SID.

The inputs are discretized at intervals of 1 minute and utilized as training data for the multiscale model. The manipulated input data used for the training of the ROM is illustrated in Fig. 8. The developed ROM is presented below:

$$\begin{aligned} x(t_{k+1}) &= Ax(t_k) + Bu(t_k) \\ y(t_k) &= Cx(t_k) \end{aligned} \tag{11}$$

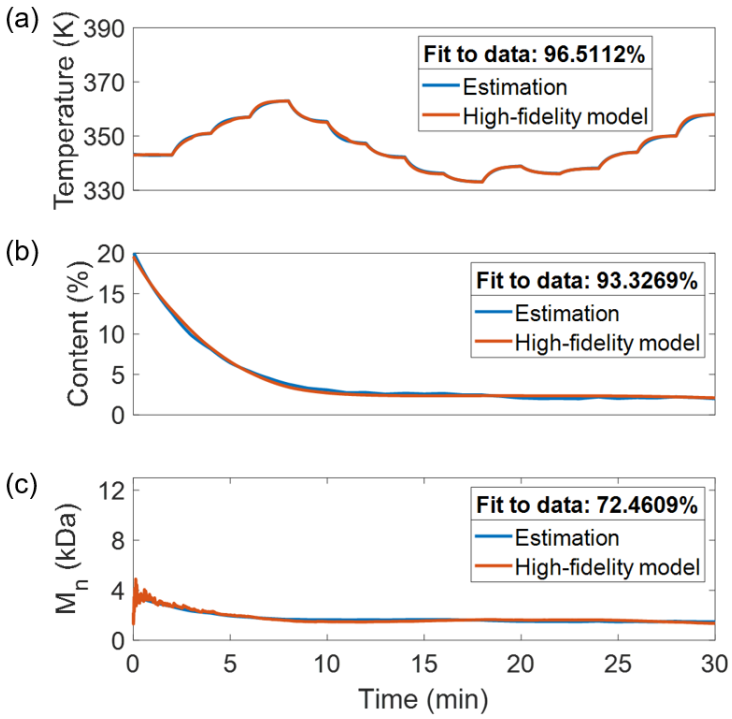

Fig. 9. System outputs derived from our ROM compared with results from the high-fidelity model: (a) free-liquor temperature, (b) lignin content, and (c) M_n.

where $x(t_k)$ is the vector of states, u denotes the system inputs, the external flow temperature and mass flow rate (*i.e.*, $u(t_k) = [T_{ext}\ \dot{M}_{ext}]^T$), y represents the system outputs, such that $y(t_k) = [T_f\ M_n\ L]^T$. The process inputs were modulated at a sampling rate of 1 minute for the extraction of the ROM. The above state-space model comprises three states and is fully controllable and observable. The matrices A, B, and C are of dimensions 3×3, 3×2, and 3×3, respectively.

Utilizing the trained ROM, we generate the results shown in Fig. 9, complete with an accompanying root-mean-square error (RMSE). Please note that both the free-liquor temperature and the lignin content align well with the data from the multiscale model. However, the M_n data exhibits less precision, primarily due to substantial fluctuations at the outset of the reaction. These fluctuations occur when a smaller number of lignin chains undergo de/repolymerization. Despite this, it is clear that the ROM predictions are largely in good agreement with the output

of the model.

4.2. Soft sensor development via Kalman filter

Real-time measurement of variables poses a significant challenge. Therefore, in this study, we estimate the lignin content and MWd from the readily available measurement, the free-liquor temperature (T_f). Given that the kinetic parameters across both scales are temperature-dependent as seen in Eq. 4, tracking the system temperature enables the computation of the reaction rate constants. This, in turn, allows for the simulation of multiscale reactions to estimate these values. It is important to note that M_n, which can be directly derived from MWd, is the variable we aim to control in this study. To estimate these variables, we implement a soft sensor, specifically a Kalman filter, for variable estimation. The design is as follows:

$$\begin{aligned} K(t_k) &= P(t_k)C^T(R(t_k) + CP(t_k)C^T)^{-1} \\ P(t_{k+1}) &= (I - K(t_k)C)P(t_k) \\ \hat{x}(t_{k+1}) &= A\hat{x}(t_k) + Bu(t_k) + K(t_k)(y_{meas}(t_k) - \hat{y}_{meas}(t_k)) \end{aligned} \tag{12}$$

where I is the identity matrix, $P(t_k)$ and $R(t_k)$ represent the state and error covariance matrices, respectively. These matrices correspond to the state estimation and measurement noise. The term $y_{meas(t_k)}$ refers to the measurable output - namely, $T_f(t_k)$ and $\hat{y}_{meas}(t_k)$ signify the inferred measurement at time $t = t_k$. The error covariance is updated at every instant with the filter gain, $K(t_k)$, and the filtered state, $\hat{x}(t_k)$. Subsequently, the updated state, $\hat{x}(t_{k+1})$, is employed to predict both the future state and the output variables, as shown in Eq. 11.

5. Closed-loop control of the fractionation process

5.1. MPC design

Given the challenges associated with real-time measurement and control of the fractionation process, we have designed a model predictive controller (MPC) using the ROM obtained from the previous section. The optimal control input profile is determined by solving the optimal control problem, as outlined in Eq. 13. The controller's performance is assessed using the fractionation process for 0.8 mm chips, with control actions executed at three-minute intervals.

$$
\begin{aligned}
\min_{T_{ext,k},\dot{M}_{ext,k}} \quad & \omega_1(M_n(t_N) - M_{n,sp})^2 + \omega_2(L(t_N) - L_{sp})^2 \\
\text{s.t.} \quad & \text{Reduced-order model, Eq. (11).} \\
& \text{Soft sensor, Eq. (12).} \\
& 340 \leq T_{ext,k}\,[K] \leq 370 \\
& 50 \leq \dot{M}_{ext,k}\,[g/min] \leq 150 \\
& |T_{ext,k+1} - T_{ext,k}| \leq 5 \quad \forall k \in [1,\ N-1]
\end{aligned}
\tag{13}
$$

where N represents the length of the prediction horizon, t_k is the reaction time, and T_k indicates the temperature at the sampling time of t_k. The cost function includes a control weight, which is set to $\omega = [\omega_1; \omega_2] = [1; 6000]$, and the set-points for M_n and L are defined as $[M_{n,sp} = 1.24\,[kDa], L_{sp} = 1.88\,[\%]]$.

5.2. Closed-loop operation results

The input profiles of the MPC, depicted in Fig. 10, are integrated into the multiscale model, serving as a virtual experiment. This model controls the temperature of the free-liquor phase, as illustrated in Fig. 11, with the aim of attaining the desired lignin content and M_n. Initially, the

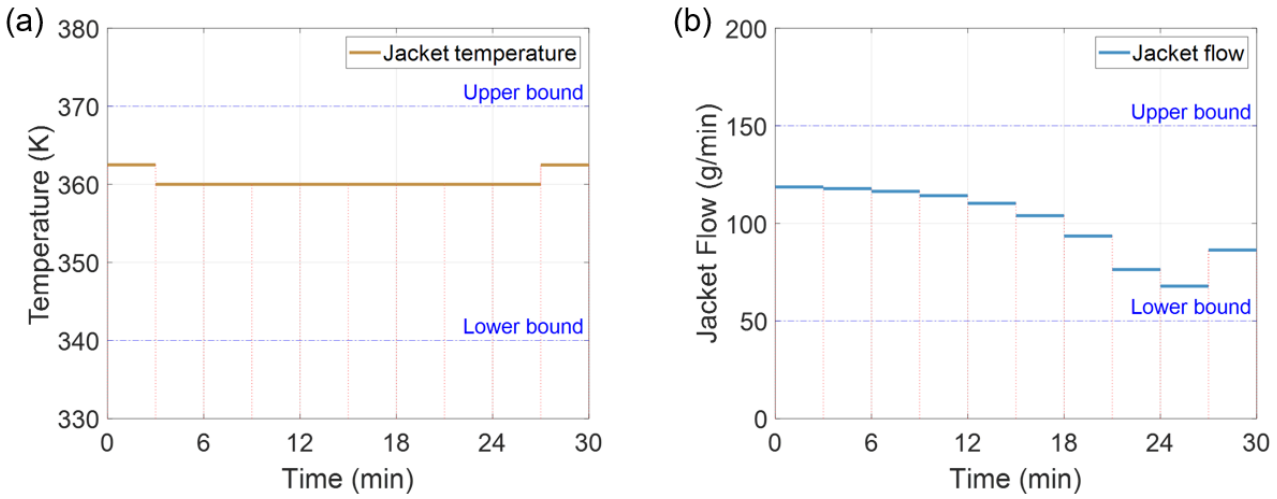

Fig. 10. The input trajectory, showing (a) the external jacket temperature, and (b) the flow rate, used for the process optimization.

free-liquor temperature is elevated by introducing a high volume of jacket flow at a high temperature, and it is then maintained around 360 K by gradually decreasing the flow rate. This approach is used because our aim is to attain a low lignin content in the bulk wood chip. Consequently, as illustrated in Fig. 12(a), the lignin content gradually converges to its predetermined set-point. Beyond the level of delignification, the real-time adjustment of lignin's MW presents a challenge, largely due to its susceptibility to both reversible macroscopic interactions and the ongoing processes of de/repolymerization. For example, the behavior of M_n, which appears to be on track towards the set-point, but manifests a slight increase around the 26-minute mark of operation. This trend is presented in Fig. 12(b). To counteract this, a slight elevation in system temperature was implemented towards the end of the operation. This was achieved by reintroducing a higher temperature to the external jacket flow, which successfully allowed M_n to finally meet its set-point. The outcome aligns with the prior observation that depolymerization governs the microscopic kinetics, leading to a decrease in M_n as the depolymerization rate accelerates in response to the increased temperature. The control objectives were thus successfully met:

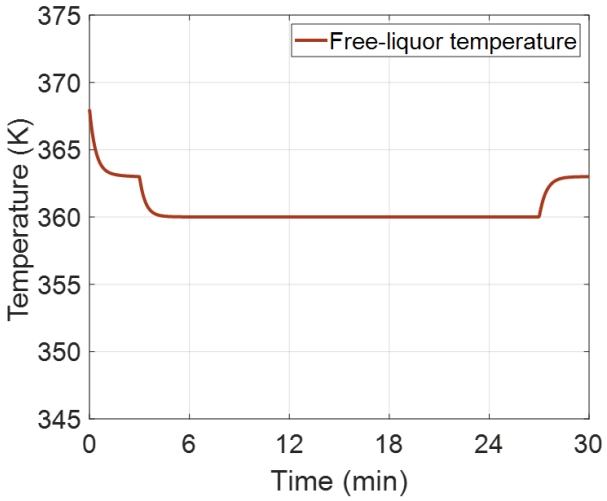

Fig. 11. The free-liquor temperature, controlled by an external jacket flow.

deviations from their respective set-points were only 0.28 % for lignin content and 0.17% for M_n. Moreover, these control actions were carried out in accordance with the control constraints provided in Eq. 13.

Despite the notable importance of Online measurement and control, these tasks have not been conducted in previous studies. In this study, it is noteworthy that the model, soft sensor, and controller were successfully integrated, enabling the effective estimation and control of critical state variables throughout the operation. Furthermore, we carried out the PSA fractionation under moderate temperatures, thereby avoiding any harsh operating conditions. This method suggests the potential for a significant increase in the profitability of biomass fractionation. By reducing energy demand and facilitating a more comprehensive utilization of lignin from bulk biomasses, this approach holds promise for enhancing the overall efficiency and sustainability of the process.

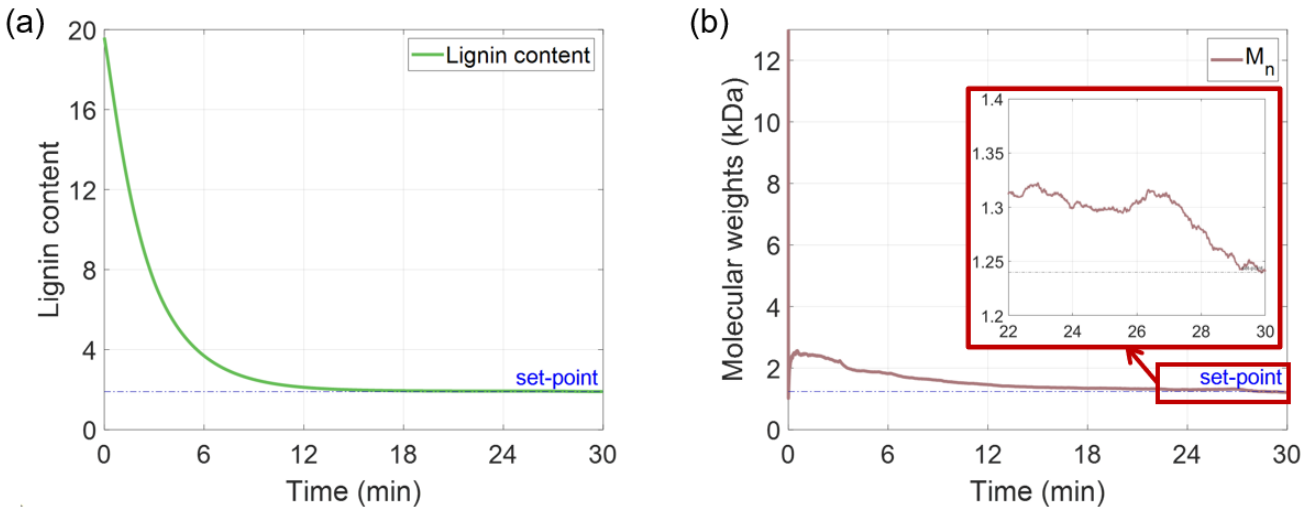

Fig. 12. The optimal operation output for the 0.8 mm wood chips: (a) lignin content, (b) M_n.

6. Conclusion

In this work, a bilayer simulation framework is employed to examine lignin behavior during the biomass fractionation process with the use of an innovative reagent, PSA. It is shown that our simple framework can successfully simulate the complicated fractionation process including delignification and de/repolymerization.

First, the macroscopic layer is developed to capture lignin dissolution and redeposition. The remaining lignin contents in the biomass could be calculated under various fractionation conditions by solving the simple ODEs. Second, the microscopic layer is constructed to simulate de/repolymerization reactions. It turned out that our stochastic and probabilistic kMC algorithm tracked the evolution of lignin MWd in the system and then helped with finding the microscopic kinetic parameters. Both layers worked in tandem within our simulation framework, and the simulation results closely aligned with the experimental results, validating our multiscale model.

Once the high-fidelity kinetic model is obtained, we implemented an MPC to regulate the

process and attain desired outcomes. A soft sensor was also incorporated to estimate parameters such as lignin content and MWd, which are not directly measurable within the actual system. The controller exhibited exceptional performance, optimizing the fractionation process by guiding both the lignin content and MWd to their designated set-points while adhering to process constraints.

In summary, this work provides a holistic framework for kinetic modeling and optimal control application for biomass fractionation processes. We believe that our comprehensive study can provide insights into lignin valorization and will pave the way for the successful development of alternative resources in the future.

Declaration of competing interest

The authors declare no competing interest.

Acknowledgments

Financial support from the Artie McFerrin Department of Chemical Engineering and the Texas A&M Energy Institute is gratefully acknowledged. This work was also supported by the National Science Foundation grant CBET-2027125.

Literature Cited

[1] F. Cherubini, A. H. Strømman, Chemicals from lignocellulosic biomass: opportunities, perspectives, and potential of biorefinery systems, Biofuel. Bioprod. Bior. 5 (5) (2011) 548–561.

[2] M. Garedew, D. Young-Farhat, J. E. Jackson, C. M. Saffron, Electrocatalytic upgrading of

phenolic compounds observed after lignin pyrolysis, ACS Sustain. Chem. Eng. 7 (9) (2019) 8375–8386. doi:10.1021/acssuschemeng.9b00019.

[3] C. Gao, M. Li, C. Zhu, Y. Hu, T. Shen, M. Li, X. Ji, G. Lyu, W. Zhuang, One-pot depolymerization, demethylation and phenolation of lignin catalyzed by hbr under microwave irradiation for phenolic foam preparation, Compos. B: Eng. 205 (2021) 108530.

[4] J. J. Bozell, Connecting biomass and petroleum processing with a chemical bridge, Science 329 (5991) (2010) 522–523.

[5] R. Rinaldi, R. Jastrzebski, M. T. Clough, J. Ralph, M. Kennema, P. C. A. Bruijnincx, B. M. Weckhuysen, Paving the way for lignin valorisation: Recent advances in bioengineering, biorefining and catalysis, Angew. Chem. Int. Ed. 55 (29) (2016) 8164–8215.

[6] M. E. Himmel, S.-Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W. Brady, T. D. Foust, Biomass recalcitrance: Engineering plants and enzymes for biofuels production, Science 315 (5813) (2007) 804–807.

[7] C. G. Yoo, X. Meng, Y. Pu, A. J. Ragauskas, The critical role of lignin in lignocellulosic biomass conversion and recent pretreatment strategies: A comprehensive review, Bioresource Technol. 301 (2020) 122784.

[8] A. Kramarenko, D. Etit, G. Laudadio, F. N. D'Angelo, β-zeolite-assisted lignin-first fractionation in a flow-through reactor, ChemSusChem 14 (18) (2021) 3838–3849.

[9] S. Constant, H. L. Wienk, A. E. Frissen, P. de Peinder, R. Boelens, D. S. Van Es, R. J. Grisel, B. M. Weckhuysen, W. J. Huijgen, R. J. Gosselink, et al., New insights into the

structure and composition of technical lignins: a comparative characterisation study, Green Chem. 18 (9) (2016) 2651–2665.

[10] K. H. Kim, C. S. Kim, Recent efforts to prevent undesirable reactions from fractionation to depolymerization of lignin: Toward maximizing the value from lignin, Front. Energy Res. 6.

[11] T. Renders, S. Van den Bosch, S.-F. Koelewijn, W. Schutyser, B. F. Sels, Lignin-first biomass fractionation: the advent of active stabilisation strategies, Energy Environ. Sci. 10 (2017) 1551–1557.

[12] P. Dhar, R. Vinu, Understanding lignin depolymerization to phenols via microwave-assisted solvolysis process, J. Environ. Chem. Eng. 5 (5) (2017) 4759–4768.

[13] W. Xu, S. J. Miller, P. K. Agrawal, C. W. Jones, Depolymerization and hydrodeoxygenation of switchgrass lignin with formic acid, ChemSusChem 5 (4) (2012) 667–675.

[14] S. Nanayakkara, A. F. Patti, K. Saito, Lignin depolymerization with phenol via redistribution mechanism in ionic liquids, ACS Sustain. Chem. Eng. 2 (9) (2014) 2159–2164.

[15] S. Nanayakkara, A. F. Patti, K. Saito, Chemical depolymerization of lignin involving the redistribution mechanism with phenols and repolymerization of depolymerized products, Green Chem. 16 (4) (2014) 1897–1903.

[16] K. N. Khalili, P. de Peinder, J. Donkers, R. J. Gosselink, P. C. Bruijnincx, B. M. Weckhuysen, Monitoring molecular weight changes during technical lignin depolymerization by operando attenuated total reflectance infrared spectroscopy and chemometrics, ChemSusChem 14 (24) (2021) 5517–5524.

[17] S. Pahari, J. Moon, M. Akbulut, S. Hwang, J. S.-I. Kwon, Model predictive control for wormlike micelles (wlms): Application to a system of ctab and nacl, Chem. Eng. Res. Des. 174 (2021) 30–41.

[18] J. R. Gasson, D. Forchheim, T. Sutter, U. Hornung, A. Kruse, T. Barth, Modeling the lignin degradation kinetics in an ethanol/formic acid solvolysis approach. part 1. kinetic model development, Ind. Eng. Chem. Res. 51 (32) (2012) 10595–10606.

[19] B. Bawareth, D. Di Marino, T. A. Nijhuis, M. Wessling, Unravelling electrochemical lignin depolymerization, ACS Sustain. Chem. Eng. 6 (6) (2018) 7565–7573.

[20] J. McDermott, M. Klein, Chemical and probabilistic modelling of complex reactions: A lignin depolymerization example, Chem. Eng. Sci. 41 (4) (1986) 1053–1060.

[21] C. Smith, T. Williams, Studies of the mathematical modeling simulation and control of the operation of Kamyr continuous digester for the Kraft process, Ph.D. thesis, Purdue University, West Lafayette, IN (1974).

[22] P. A. Wisnewski, F. J. Doyle, F. Kayihan, Fundamental continuous pulp-digester model for simulation and control, AIChE J. 43 (12) (1997) 3175–3192.

[23] S. H. Son, H.-K. Choi, J. S.-I. Kwon, Multiscale modeling and control of pulp digester under fiber-to-fiber heterogeneity, Comput. Chem. Eng. 143 (2020) 107117.

[24] S. H. Son, H.-K. Choi, J. S.-I. Kwon, Application of offset-free koopman-based model predictive control to a batch pulp digester, AIChE J. 67 (9) (2021) e17301.

[25] O. Fearon, S. Kuitunen, K. Ruuttunen, V. Alopaeus, T. Vuorinen, Detailed modeling of kraft pulping chemistry. delignification, Ind. Eng. Chem. Res. 59 (29) (2020) 12977–12985.

[26] S. Pahari, J. Kim, H.-K. Choi, M. Zhang, A. Ji, C. G. Yoo, J. S.-I. Kwon, Multiscale kinetic modeling of biomass fractionation in an experiment: Understanding individual reaction mechanisms and cellulose degradation, Chem. Eng. J. (2023) 143021.

[27] Y.-Y. Wang, C. E. Wyman, C. M. Cai, A. J. Ragauskas, Lignin-based polyurethanes from unmodified kraft lignin fractionated by sequential precipitation, ACS Appl. Polym. Mater. 1 (7) (2019) 1672–1679.

[28] J. Xu, C. Li, L. Dai, C. Xu, Y. Zhong, F. Yu, C. Si, Biomass fractionation and lignin fractionation towards lignin valorization, ChemSusChem 13 (17) (2020) 4284–4295.

[29] T. Pang, G. Wang, H. Sun, W. Sui, C. Si, Lignin fractionation: Effective strategy to reduce molecule weight dependent heterogeneity for upgraded lignin valorization, Ind. Crops Prod. 165 (2021) 113442.

[30] S. Pahari, B. Bhadriraju, M. Akbulut, J. S.-I. Kwon, A slip-spring framework to study relaxation dynamics of entangled wormlike micelles with kinetic monte carlo algorithm, J. Colloid Interface Sci. 600 (2021) 550–560.

[31] P. Shah, H.-K. Choi, J. S.-I. Kwon, Achieving optimal paper properties: A layered multiscale kmc and lstm-ann-based control approach for kraft pulping, Processes 11 (3) (2023) 809.

[32] D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, J. Comput. Phys. 22 (4) (1976) 403–434.

[33] P. Shah, M. Z. Sheriff, M. S. F. Bangi, C. Kravaris, J. S.-I. Kwon, C. Botre, J. Hirota, Multi-rate observer design and optimal control to maximize productivity of an industry-scale fermentation process, AIChE J. 69 (2) (2023) e17946.

[34] S. Pahari, P. Bhandakkar, M. Akbulut, J. S.-I. Kwon, Optimal pumping schedule with high-viscosity gel for uniform distribution of proppant in unconventional reservoirs, Energy 216 (2021) 119231.

[35] D. He, Y. Wang, C. G. Yoo, Q.-J. Chen, Q. Yang, The fractionation of woody biomass under mild conditions using bifunctional phenol-4-sulfonic acid as a catalyst and lignin solvent, Green Chem. 22 (16) (2020) 5414–5422.

[36] A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter, D. Templeton, D. Crocker, Determination of structural carbohydrates and lignin in biomass, Tech. rep., National Renewable Energy Laboratory, Golden, CO (USA), NREL/TP-510-42618 (2008).

[37] M. A. Gilarranz, F. Rodríguez, A. Santos, M. Oliet, F. García-Ochoa, J. Tijero, Kinetics of eucalyptus globulus delignification in a methanol-water medium, Ind. Eng. Chem. Res. 38 (9) (1999) 3324–3332.

[38] T. Pang, G. Wang, H. Sun, W. Sui, C. Si, Lignin fractionation: Effective strategy to reduce molecule weight dependent heterogeneity for upgraded lignin valorization, Ind. Crops Prod. 165 (2021) 113442.

[39] A. TenWolde, J. D. McNatt, L. Krahn, Thermal properties of wood and wood panel products for use in buildings, Tech. rep., Forest Products Laboratory, Madison, WI (USA), DOE/USDA-21697-1 (1988).

[40] H.-K. Choi, J. S.-I. Kwon, Multiscale modeling and control of kappa number and porosity in a batch-type pulp digester, AIChE J. 65 (6) (2019) e16589.

[41] J. Ralph, C. Lapierre, W. Boerjan, Lignin structure and its engineering, Curr. Opin. Biotechnol. 56 (2019) 240–249.

[42] F. Beall, Thermogravimetric analysis of wood lignin and hemicelluloses, Wood Fiber Sci. (1969) 215–226.

[43] B. Zhang, H.-J. Huang, S. Ramaswamy, Reaction kinetics of the hydrothermal treatment of lignin, in: Biotechnology for Fuels and Chemicals, Humana Press, Totowa, NJ, 2008, pp. 487–499.

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: