Semi-mechanistic modelling of ionic liquid-based biomass fractionation

Suhaib Nisara, Agnieszka Brandt-Talbotb, Jason P. Halletta, Benoit Chachuata\*

aDepartment of Chemical Engineering, South Kensington Campus, Imperial College London, London SW7 2AZ, United Kingdom

bDepartment of Chemistry, Molecular Sciences Research Hub, Imperial College London, London W12 0BZ, United Kingdom

\*[b.chachuat@imperial.ac.uk](mailto:b.chachuat@imperial.ac.uk)

Abstract

Fractionation of lignocellulosic biomass is a crucial step to provide cellulose, lignin, and hemicellulose for further processing. This work focuses on modelling the fractionation of woody biomass using the ionoSolv process, which employs low-cost ionic liquid water mixtures. We model a simple reaction network to describe the solvent-extraction of three main biopolymers from solid lignocellulosic biomass. We estimate the corresponding kinetic parameters and their credibility intervals using Bayesian parameter estimation and then exploit the calibrated model for a multi-criterion analysis employing three process metrics: glucan (cellulose) recovery, hemicellulose removal, and lignin removal (delignification). Specifically, we construct a probabilistic design space by propagating the model parameter uncertainty, with a view to predicting a feasible operating window for key process variables (pretreatment time, temperature and solids loading) to meet certain thresholds for each metric. Overall, the development of semi-mechanistic models provides a novel framework for the analysis and optimisation of ionic liquid-based biomass pretreatment.

**Keywords**: biomass, ionic liquids, semi-mechanistic models, Bayesian inference

* 1. Background and Introduction

Climate change and resource scarcity are driving the demand for sustainable alternatives to fossil carbon energy sources and reduced carbon materials such as plastics. Biomass is any renewable organic material coming from plants and animals and has been proposed for both energy and material sources. Specifically, lignocellulosic biomass comprises woody plant parts, including agricultural and food industry residues (Ragauskas et al., 2006, Barbará et al. , 2023). It is mainly composed of three biopolymers: cellulose, hemicellulose, and lignin. The chemistry and spatial arrangement of the components of lignocellulosic biomass hinders the direct conversion to specific products, thereby requiring a chemical transformation to maximise the availability of the biopolymers for specific applications.

Recently, ionic liquids (ILs) have been proposed as effective solvents for biomass fractionation. The ionoSolv process utilises protic ionic liquid-water solutions, with the ILs synthesised by simple combination of an aqueous Bronsted acid with a Bronsted base, to fractionate lignocellulosic biomass into a cellulose-rich pulp and a lignin and hemicellulose rich liquid. Solid lignin can then be separated after increasing the water content in the ionic liquid solution. This is depicted in Figure 1 (Brandt-Talbot et al., 2017). The cellulose pulp is washed after the pretreatment to remove the IL and as much lignin as possible followed by further processing to cellulosic biofuels (after hydrolysis) or biomaterials. The dissolved lignin, recovered from the IL by the addition of an anti-solvent such as water, can be either combusted for process heat or converted to products. Distillation can then be used to regenerate the IL-water solution for recycling.

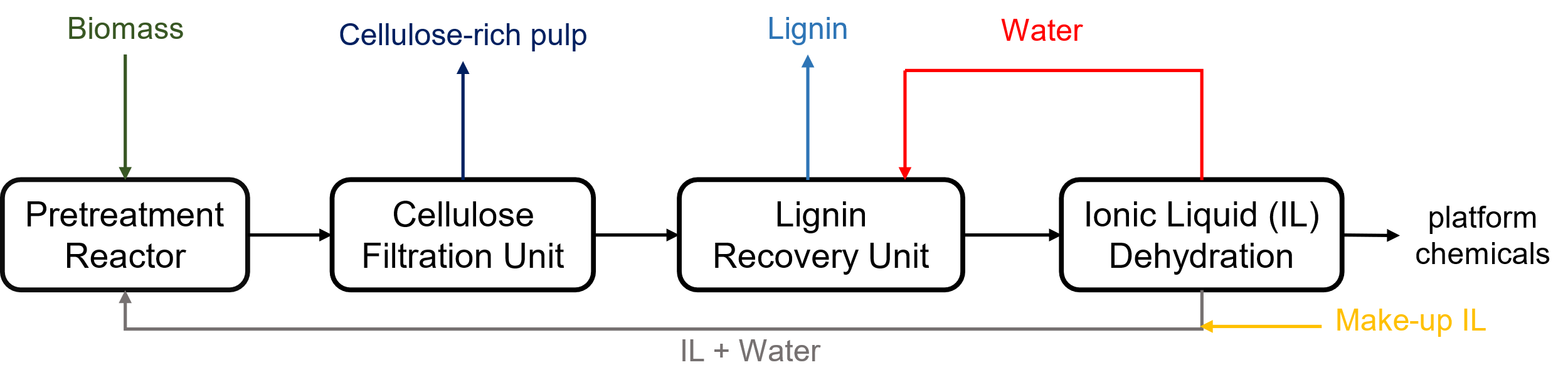


Figure 1: ionoSolv process for lignocellulosic biomass fractionation using protic ionic liquids

Modelling biomass fractionation to predict process behaviour helps design more intelligent experiments to optimise and study the process. Previous studies have considered quadratic response surface modelling of the ionoSolv process (Abouelela et al., 2023). Whilst simple to use, these quadratic models have poor extrapolative capability and require many experiments for wider applicability. This work focuses on the development of semi-empirical chemical models, with the aim of providing a feasible operating region to focus experiments for process optimisation. These models balance the required complexity to adequately model the system with the limited data available. The methodology is detailed in Section 2, followed by results and discussion in Section 3, and conclusions in Section 4.

* 1. Methodology
     1. Model Equations

Reaction scheme (R1) was developed to describe the main reactions occurring during lignocellulose fractionation. As the wet-chemical analysis method can only elucidate the compositional structure of the cellulose-rich pulp, the models were trained with the pulp compositions consisting of cellulose, residual hemicellulose, and residual lignin. The compositional analysis method cannot distinguish between glucose derived from cellulose or hemicellulose and as most glucose originates from cellulose; the term *glucan* refers to cellulose. For lignin, the compositional analysis method estimates native and condensed (re-precipitated) lignin as a lumped component as it measures the lignin content gravimetrically. Therefore, when calculating residuals for lignin, the components lignin and condensed lignin were combined.

|  |  |
| --- | --- |
|  | (R1) |

Reaction stoichiometries were used to enforce mass balances on the system. For each experiment, the initial composition was set as the raw biomass composition. An Arrhenius power kinetic expression was employed for reactions 2-4 (Eq. (1)), while a Haldane-type kinetic expression was used for reaction 1 (Eq. (2)). Temperature curves (Gschwend et al., 2018) were used to model temperature dependence through the parameters and . The solids loading and reaction order effects were described with the parameters and respectively in Eq. (1). The parameter in Eq. (2) primarily describes the inhibition regime in Haldane-style kinetics at high reactant concentrations, while represents the saturation regime, where the reaction rate is maximum at intermediate reactant concentrations.

|  |  |
| --- | --- |
|  | (1) |
|  | (2) |

Ordinary differential equations (ODEs) for each species involved in reaction scheme (R1) were expressed in terms of the reaction rates through the reaction stoichiometry and are shown in Eq. (3). The ODEs were integrated using CasADi (Andersson et al., 2019) in Python.

|  |  |
| --- | --- |
|  | (3) |

The main process metrics used to quantify the performance of lignocellulose fractionation (Eq. (4)) focus on cellulose retention, hemicellulose removal and delignification in the biomass.

|  |  |
| --- | --- |
|  | (4) |

* + 1. Bayesian Estimation and Design Space characterisation

Bayesian parameter inference views the model parameters as random variables and makes use of Bayes' rule to compute a posterior distribution of the parameters by accounting for experimental data through the likelihood function as well as prior knowledge on the parameters. A nested sampling-based method was used with the method outlined in Bernardi et al. (2019) implemented in gPROMS ModelBuilder 7.1.1 (Siemens Industry Software, 2023). The Bayesian estimation returns a set of parameter values with their posterior probabilities, expressing the parametric uncertainty in the estimated system. A flat prior was considered as there was no previous knowledge to suggest where the optimal parameter values may lie. Additionally, a constant variance model was used to represent measurement uncertainty in the likelihood function.

A Bayesian approach was also used to characterise the probabilistic design space using nested sampling with the Python package DEUS as outlined in Kusumo et al. (2020). The design space characterisation yields samples of optimal operating conditions (time, temperature, solids loading) that satisfy desired targets set for the metrics in Eq. (4) at a given feasibility. The design space characterisation considers parametric uncertainty by sampling parameter scenarios from the posterior parameter distribution, which are then considered as uncertainty scenarios for the design space characterisation. As with the parameter estimation, the design space estimation with nested sampling progressively samples in nested contours of increasing likelihood.

* 1. Results and Discussion
     1. Bayesian Estimation Results

Figure 2 illustrates the process metric profiles for the maximum a posteriori parameter estimates with the experimental data and actual measurement uncertainty plotted for comparison. Generally, all three components are modelled well. For the glucan recovery (left), the reduction is tracked well even at the highest modelled temperature of 170 °C, where experiments are challenged by the difficulties in washing the sticky pulp after pretreatment. For hemicellulose removal (centre), the models predict the decrease in hemicellulose content across the temperature range. For delignification (right), the combination of lignin extraction and re-deposition of condensed lignin onto the biomass pulp captured the delignification trend well, especially at 150 and 160°C, while anticipating the maximum at 170°C. It has previously been observed that the re-deposited “lignin” is a combination of condensed lignin fragments, and sugars-derived oligomers coined “pseudo-lignin” or humins (Shinde et al., 2018). This is compounded by the inability of the current analysis protocols to quantitatively distinguish water insoluble degradation products from unextracted lignin, complicating the understanding of the re-deposition reaction. As humins formation is expectedly more significant at higher severities, future model refinement should describe reactions involving dissolved sugars.

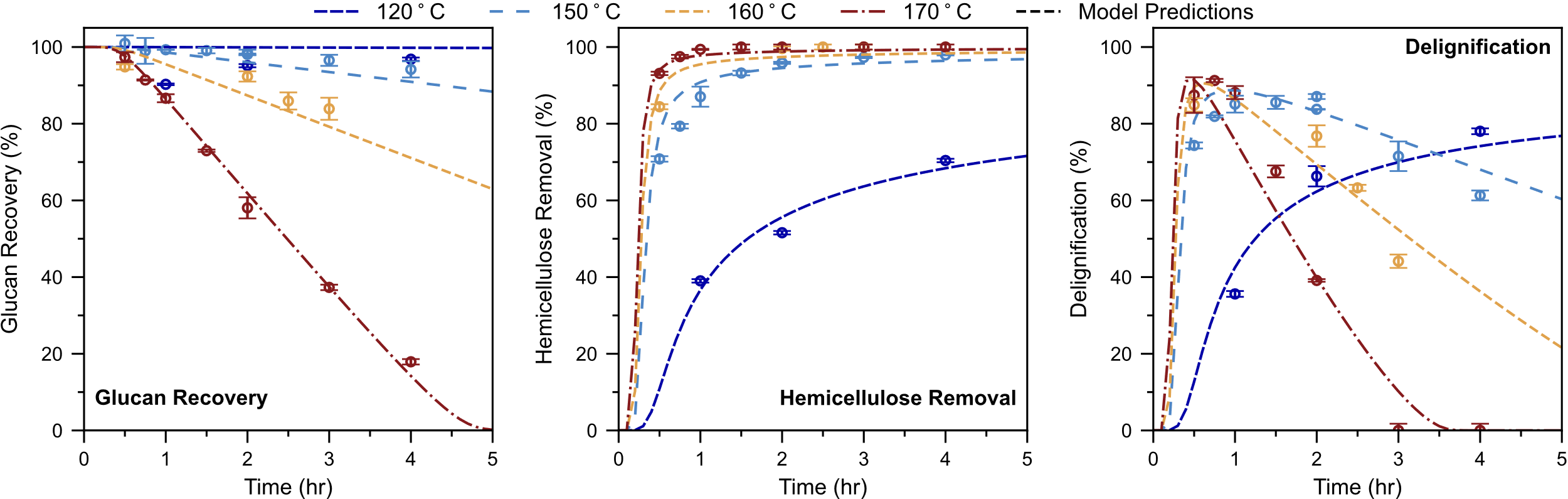


Figure 2: Maximum a posteriori prediction of ionoSolv experiments

Figure 3 illustrates the variability in the model predictions for all three metrics over the 95% credibility region of the estimated parameters at 150°C. In general, the predictive variability is contained within a narrow band through the experimental points. The overall certainty in model predictions provides confidence for using these estimates for further process optimisation and exploitation.

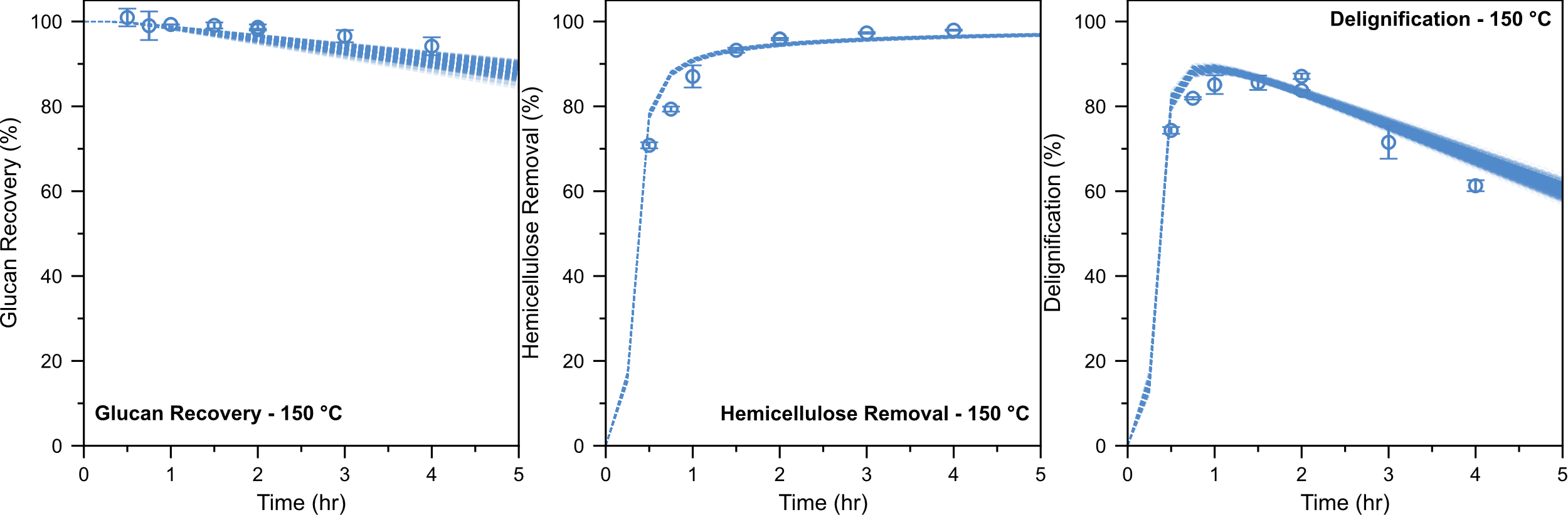


Figure 3: Model prediction variability over the 95% credibility region at 150°C

* + 1. Probabilistic Design Space Characterisation

The probabilistic design space characterisation can provide a region for feasible operation of the ionoSolv process. Two scenarios were investigated, one where glucan recovery and hemicellulose removal needed to be greater than 90% and delignification greater than 75%. In the second scenario, the thresholds were tightened to require glucan recovery and hemicellulose removal to be greater than 90% and delignification to be greater than 75%. The probabilistic design spaces were estimated over 1500 parameter uncertainty scenarios obtained from the Bayesian estimation to determine 1000 feasible operating points with at least 85% feasibility.

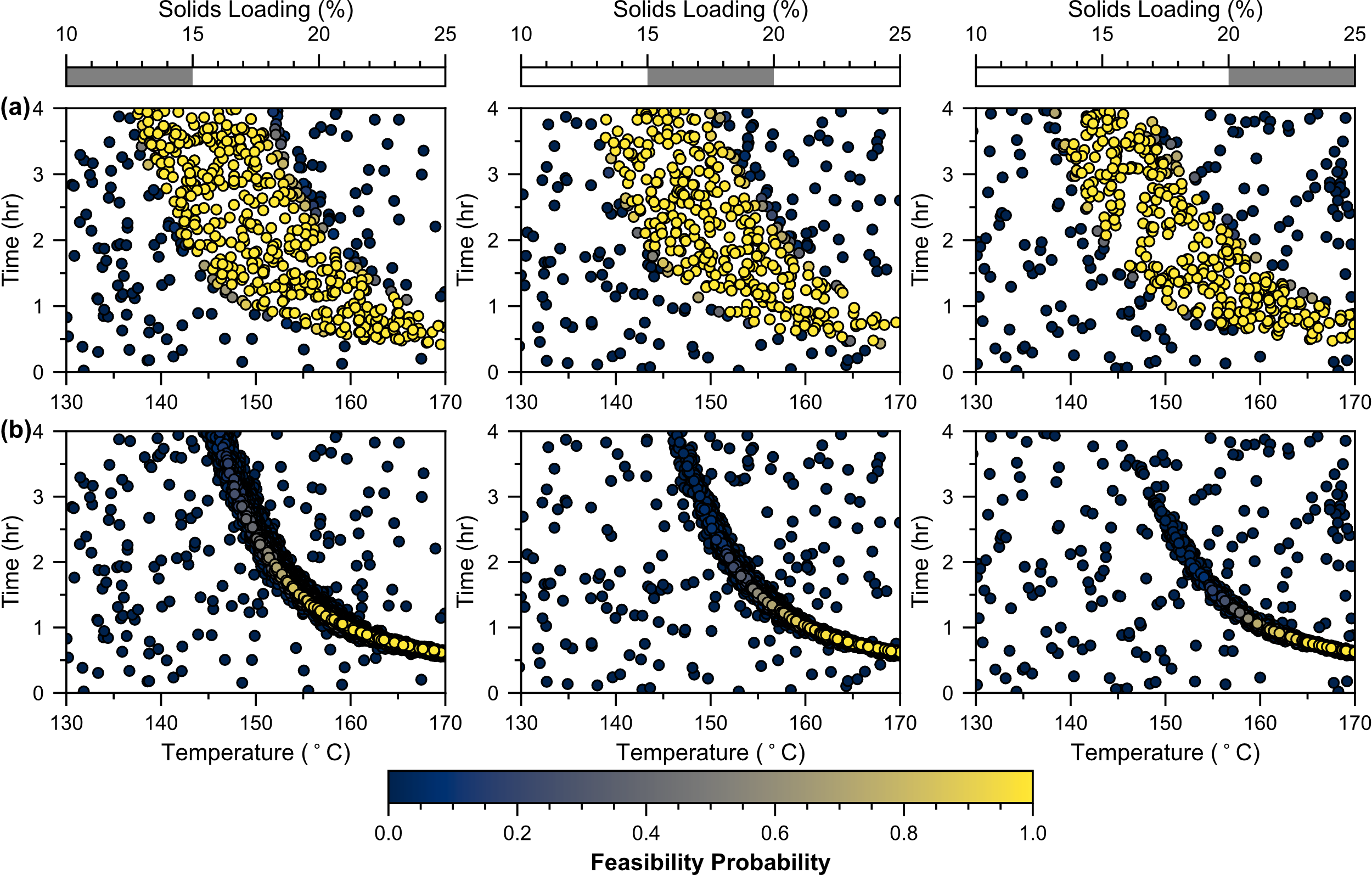


Figure 4: Probabilistic design space for feasible operation of the ionoSolv process for (a) lower and (b) higher quality thresholds

This probabilistic design space in Figure 4 (a) shows that the feasible operating region is sandwiched between regions of low feasibility at higher and lower temperatures of the feasible region. At lower temperatures, there is insufficient hemicellulose and lignin removal from the biomass. At higher temperatures and longer times, the glucan recovery decreases due to hydrolysis of cellulose to water-soluble glucose, cellobiose and short oligomers, while the lignin re-deposition and humin deposition also become dominant. Interestingly, the feasible operating region shrinks with increasing temperature, suggesting a reduced process flexibility and higher risk of error as the temperature increases. This may be traced back to the delignification, with the maximal region becoming narrower with increasing temperature, due to the more pronounced lignin re-deposition. These observations concur with the experimental data in Figure 2, which imply that the ionoSolv process must be operated at low severity for high glucan recoveries and high severity for high hemicellulose removals, with delignification essentially determining the feasible space in this overlapping region. The parametric uncertainty also has a limited impact on the feasibility probability, with most points having a probability of either 0 or 1. This matches the Bayesian estimation results, which implied a limited impact of parametric uncertainty, particularly in high feasibility regions.

For the higher quality thresholds in Figure 4 (b), the more stringent process requirements unsurprisingly limit the range of feasible operation, but the general trends remain unchanged. There is a noticeable shift in the feasible operating region with increasing solids loading, requiring higher temperatures to meet the process constraints. The more constrained design space also requires a balance between conservative and riskier validation experiments to meet process constraints.

* 1. Conclusions and Future Work

In this work, a semi-mechanistic framework modelling the extraction of the main biopolymers composing lignocellulosic biomass into an ionic liquid water mixture was proposed. Bayesian parameter estimation was employed to estimate the kinetic parameters and their credibility regions. The calibrated model was then exploited for the characterisation of a feasible operating set for different thresholds of the process metrics with high confidence. Further work will require the consideration of humin formation from dissolved sugars for the improvement of the re-deposition model and further experimental validation of the feasible operating set predictions.

Acknowledgements

Suhaib Nisar is grateful to the Department of Chemical Engineering at Imperial College London for a PhD scholarship and to the EPSRC Centre for Doctoral Training in Next Generation Synthesis & Reaction Technology for the PhD studentship under grant EP/S023232/1.

References

A.R. Abouelela, P.Y.S. Nakasu, and J.P. Hallett. 2023. ‘Influence of Pretreatment Severity Factor and Hammett Acidity on Softwood Fractionation by an Acidic Protic Ionic Liquid’. *ACS Sustainable Chem. Eng.* 11 (6): 2404–15.

J.A.E. Andersson, et al. 2019. ‘CasADi: A Software Framework for Nonlinear Optimization and Optimal Control’. *Math. Program. Comput.* 11 (1): 1–36.

P.V. Barbará, A.A. Rafat, J.P. Hallett, and A. Brandt-Talbot. 2023. ‘Purifying cellulose from major waste streams using ionic liquids and deep eutectic solvents’. *Curr. Opin. Green Sustain. Chem.* 41: 100783

A. Bernardi, L. Gomoescu, J. Wang, C.C. Pantelides, D. Chadwick, and B. Chachuat. 2019. ‘Kinetic Model Discrimination for Methanol and DME Synthesis Using Bayesian Estimation’. *IFAC-PapersOnLine*, 52 (1): 335–40.

A. Brandt-Talbot, F.J.V. Gschwend, P.S. Fennell, T.M. Lammens, B. Tan, J. Weale and J.P. Hallett, 2017, ‘An Economically Viable Ionic Liquid for the Fractionation of Lignocellulosic Biomass’, *Green Chem.* 19 (13): 3078–3102.

F.J.V. Gschwend, F. Malaret, S. Shinde, A. Brandt-Talbot, and J.P. Hallett. 2018. ‘Rapid Pretreatment of Miscanthus Using the Low-Cost Ionic Liquid Triethylammonium Hydrogen Sulfate at Elevated Temperatures’. *Green Chem.* 20 (15): 3486–98.

K.P. Kusumo, L. Gomoescu, R. Paulen, S.G. Muñoz, C.C. Pantelides, N. Shah, and B. Chachuat. 2020. ‘Bayesian Approach to Probabilistic Design Space Characterization: A Nested Sampling Strategy’. *Ind. Eng. Chem. Res.* 59 (6): 2396–2408.

A.J. Ragauskas, et al. 2006. ‘The Path Forward for Biofuels and Biomaterials’. *Science* 311 (5760): 484–89.

S.D. Shinde, X. Meng, R. Kumar, and A.J. Ragauskas. 2018. ‘Recent Advances in Understanding the Pseudo-Lignin Formation in a Lignocellulosic Biorefinery’. *Green Chem.* 20 (10): 2192–2205.

Siemens Industry Software, 1997–2023. gPROMS