PARAMETER ESTIMATION AND UNCERTAINTY QUANTIFICATION
FOR CHEMICAL PROCESS MODELS WITH UNCERTAIN INPUTS

by

Kaveh Abdi

A thesis submitted to the Department of Chemical Engineering
In conformity with the requirements for
the degree of Doctor of Philosophy

Queen’s University
Kingston, Ontario, Canada
(November, 2022)

Copyright ©Kaveh Abdi, 2022
Abstract

Statistical methods are developed and applied for two types of problems that modelers may encounter when developing fundamental models of chemical processes. First, I propose a mathematical model for photopolymerization of 1,6-hexanediol diacrylate with a bifunctional initiator. This model is complicated because it has 40 parameters that account for a variety of phenomena including crosslinking, backbiting, cyclization, and diffusion-dependent reactions. Parameter estimability and subset selection methods are used to determine the optimal subset of parameters that should be estimated using conversion vs. time data. The results reveal that crosslinking, backbiting, and cyclization are important reactions, and that propagation, branching, and termination reactions, along with initiation efficiency, are diffusion-dependent. The model fit to the data and predictions of validation data set are both excellent.

The second part of the thesis focuses on error-in-variables models (EVMs), which are used to account for uncertain inputs. Difficulties involving output-measurement-variance estimates for EVM parameter estimation are identified. Repeated experiments are categorized as either true replicates or pseudo replicates due to different assumptions about uncertain inputs. Proposed methods for obtaining variance estimates are illustrated using a n-butyl methacrylate/n-butyl acrylate copolymerization model. A parametric bootstrap method is extended for EVM situations to obtain uncertainties for the parameters. Results reveal that assumptions about the type of repeated experiments are important for measurement-variance estimation and determining parameter estimate uncertainties. Next, I provide a literature review about methods for parameter and prediction uncertainty quantification when input uncertainties are considered. Output measurement variance estimation and uncertainty quantification methods are then extended for use in a multi-response model describing hydroisomerization of n-hexane. The corresponding literature data contains data from pseudo-replicate experiments. Parameter estimation and uncertainty quantification are implemented for different assumed levels of input uncertainties. Results confirm that accounting for input uncertainties using EVM when input uncertainties are small (i.e., 1% of typical input
values) leads to results similar to weighted-least squares (WLS). Larger input uncertainties results in noticeable differences when EVM results are compared to WLS results.
Co-Authorship

The material in chapter 2 has been published in Chemical Engineering Science and the material in Chapter 3 has been published in AIChE journal. The material in Chapter 4 has been submitted to Industrial and Engineering Chemistry Research and material in Chapter 5 has been submitted to AIChE journal. The materials in all the chapters and the corresponding calculations, simulations, figures, equations were implemented by me under the supervision of Dr. Kim McAuley. Dr. Kim McAuley also edited and revised this thesis and provided technical suggestions. Coauthors for the manuscripts are listed in the footnotes at the beginning of Chapters 2 to 5 and Appendix B.
Acknowledgements

I would like to deeply thank my supervisor professor Kim B. McAuley for her endless helps and support during my PhD. I was lucky to work with such a knowledgeable supervisor who cares a lot about her students and tries to make their graduate studies enjoyable.

I would like to thank all my friends in Kingston and Toronto for making me happy with their emotional support during my PhD.

I would like to acknowledge the Department of Chemical Engineering at Queen’s University, MITACS, EUROKIN and Engineering and Research Council of Canada for their financial support. I also would like to thank Canon Production Printing for providing data for Chapter 2 of this thesis. My special thanks to Professor Piet Iedema from University of Amsterdam for helping me during my internships. I would like to thank graduate students at University of Amsterdam for their support during my stay in the Netherlands. I would like to thank Dr. Benoit Celse from EUROKIN for the guidance and providing the opportunity to work with EUROKIN. I would like to thank professor Joris W. Thybaut for his technical recommendations for the study in Chapter 5.

I want to thank all the faculty and staff at the Department of Chemical Engineering as well as my fellow graduate students for their assistance.

I want to thank my parents, Zahra and Ghasem, and my brother, Babak, for their kindness, supports and encouragement during my PhD away from home. Last but not least, I want to thank my love, Neda, for her love, emotional support, help and encouragement during my PhD.
Table of Contents

Abstract ........................................................................................................................................... ii
Co-Authorship ................................................................................................................................. iv
Acknowledgements ......................................................................................................................... v
Table of Contents ............................................................................................................................ vi
List of Figures ................................................................................................................................... x
List of Tables .................................................................................................................................... xix
Nomenclature .................................................................................................................................... xxii

Chapter 1 Introduction .................................................................................................................... 1
  1.1 Thesis Objectives ...................................................................................................................... 2
    1.1.1 Mathematical Modeling and Parameter Estimation for HDDA Photopolymerization with
      Bifunctional Initiator ..................................................................................................................... 2
    1.1.2 Output Measurement Variance Estimation, Parameter Estimation and Uncertainty Quantification
      in EVM situations ............................................................................................................................ 3
  1.2 Thesis Outline ............................................................................................................................. 5

Chapter 2 Mathematical Modeling and Parameter Estimation for 1,6-Hexanediol Diacrylate
  Photopolymerization with Bifunctional Initiator ............................................................................. 8
  2.1 Abstract ..................................................................................................................................... 8
  2.2 Introduction ............................................................................................................................... 9
  2.3 Experimental Data ...................................................................................................................... 14
  2.4 Model Development................................................................................................................... 16
    2.4.1 Reaction Scheme .................................................................................................................. 16
    2.4.2 Rate Coefficients ................................................................................................................. 26
    2.4.3 Balance Equations ............................................................................................................... 31
  2.5 Parameter Estimation and Selection of Important Phenomena ............................................... 34
  2.6 Final Simulation Results and Model Validation ....................................................................... 43
  2.7 Conclusions ............................................................................................................................... 46
  2.8 References for Chapter 2 .......................................................................................................... 48

Chapter 3 Estimation of Output Measurement Variances for EVM Parameter Estimation ............... 51
  3.1 Abstract ..................................................................................................................................... 51
  3.2 Introduction ............................................................................................................................... 52
  3.3 Proposed Classification of Replicate Experiments in EVM ....................................................... 60
3.4 Proposed Output Measurement-Variance Estimation in Situations Involving Replicate Experiments ..........................................................................................................................63
3.5 Case Study: EVM Reactivity-ratio Estimation Using Mayo-Lewis Equation ..............................................................65
3.6 Conclusions .................................................................................................................................................................75
3.7 References for Chapter 3 ......................................................................................................................................77

Chapter 4 Propagating Input Uncertainties into Parameter Uncertainties and Model Prediction Uncertainties- A Review ........................................................................................81
4.1 Abstract .......................................................................................................................................................................81
4.2 Introduction ..............................................................................................................................................................82
4.3 Propagating Input Uncertainties into Prediction Uncertainties ...............................................................................82
4.4 Propagating Input Uncertainties into Parameter Uncertainties ..............................................................................91
4.5 Simultaneous Propagation of Input and Parameter Uncertainties into Prediction Uncertainties ................................99
4.6 Conclusion and Discussion ......................................................................................................................................104
4.7 References for Chapter 4 ......................................................................................................................................106

Chapter 5 Parameter Estimation and Prediction Uncertainties for Multi-response Kinetic Models with Uncertain Inputs ........................................................................................................114
5.1 Abstract .......................................................................................................................................................................114
5.2 Introduction ..............................................................................................................................................................115
5.3 Background ..............................................................................................................................................................117
5.3.1 Using Pseudo-Replicate Experiments to Estimate Parameters and Obtain Output Measurement Variances ..............................................................................................................117
5.3.2 Parameter Uncertainty Quantification Using Bootstrapping for a Single-response Model with a Single Uncertain Input ........................................................................................................120
5.3.3 Quantifying Prediction Uncertainties based on Uncertainties in Inputs and Parameter Estimates ..................120
5.4 Proposed Methodology ...........................................................................................................................................121
5.4.1 Bootstrap Method for Quantifying Parameter Uncertainties in Multi-response EVM ..........................121
5.4.2 Bootstrap Method for Quantifying Prediction Uncertainties from Error-in-Variables Models 122
5.5 Case Study: Catalytic n-Hexane Hydroisomerization in a Well-mixed Reactor ........................................123
5.5.1 Model Equations and Experimental System ......................................................................................................123
5.5.2 EVM Parameter and Variance Estimation ..........................................................................................................126
5.5.3 Parameter and Prediction Uncertainty Quantification ..........................................................................................131
5.5.4 Uncertainties in Predicted Yield and Selectivities ..............................................................................................136
5.6 Conclusions ..............................................................................................................................................................138
List of Figures

Figure 2-1 Experimental data corresponding to a) constant-light-exposure experiments with the wavelength of 410 nm and different light intensities and b) short-time exposure photopolymerization with wavelength of 405 nm, intensity of 20 W/m² and a variety of exposure times. In part b), data from three replicate runs are shown for each exposure time..........................................................15

Figure 2-2 Decomposition of a) BAPO initiator I and b) the remaining carbon-phosphorous bond in Ī.........................................................................................17

Figure 2-3 Initiation reaction involving a phosphorous-centered initiator radical $I^*$ and monomer, resulting in a growing polymer chain containing a Ī end group (within the ellipse) that may later decompose ...... 17

Figure 2-4 Formation of different macroradicals including: a) regular radical ends $R^\equiv\bullet$, b) cyclized radical ends $C^\bullet$, c) branch-point radicals $B^\bullet$, and d) tertiary radicals $T^\bullet .......................................................................................19

Figure 2-5 The vinyl group that is part of a functional group $R^\equiv\bullet$ (in the ellipse on the left-hand side) is treated as a pendant vinyl group $V_p$ (in the circle on the right-hand side) when $R^\equiv\bullet$ is consumed by reaction........................................................................................................20

Figure 2-6 Experimental data ( ×, x, ×, x, and x ) and corresponding model predictions shown using curves with the same colors obtained using final parameter estimates for a) short-time-exposure experiments with light intensity $I_{t,0} = 20 \, W/m²$ and time exposures $t_{\text{exp}} = 0.5, 1$, and 2 s, and for constant-light-exposure experiments with light intensities b) $I_{t,0} = 105$ and 310 W/m², c) $I_{t,0} = 115$ and 760 W/m², and d) 195 and 1350 W/m². ........................................................................................................45

Figure 2-7 Comparison of model predictions and final monomer conversion data for three short-time-exposure experiments conducted using $I_{t,0} = 20 \, W/m²$. Straight lines between predicted conversions are shown to guide the eye........................................................................................................45

Figure 2-8 Model predictions (curved lines) obtained from final estimated parameters and the corresponding FTIR measured values (× and x ) for validation data obtained using light intensities: a) $I_{t,0} = 240 \, W/m²$ and b) $I_{t,0} = 320 \, W/m²$ ........................................................................................................45

Figure 3-1 JC information for reactivity ratios obtained using data from Table 3-7 and different uncertainty assumptions. Different reactivity-ratio estimate, indicated by • are obtained for situations where a) the output measurement variance $\sigma_{\text{Ln}F_1}^2 = 2.5 \times 10^{-3}$ is assumed known, b) the output measurement variance $\delta_{\text{Ln}F_1}^2$ is estimated from true replicate runs and c) the output measurement variance $\delta_{\text{Ln}F_1}^2$ is estimated assuming pseudo replicate runs. The red ellipse shown in all three plots is the linearization-based JCR obtained by assuming that $\delta_{\text{Ln}F_1}^2$ is perfectly known. This ellipse is repeated in b) and c) as a helpful...
reference for comparison purposes. Green dots in a), b) and c) are plausible alternative reactivity-ratio estimates obtained using bootstrapping. .......................................................... 75

**Figure 4-1** Latin hypercube sampling for two uncertain inputs where the first input is normally distributed \( u_1 \sim N(2,1) \) and the second uncertain input belongs to a uniform distribution \( u_2 \sim \text{Uniform}(2,3) \) ........ 87

**Figure 5-1** Contribution of input uncertainty and output uncertainty to the overall pooled estimates of variance \( \sigma_{EVM}^2 \) from equation (5.13) when different levels of input uncertainties are assumed during EVM parameter estimation. ............................................................ 131

**Figure 5-2** Parameter uncertainty boxplots for a) \( k_{ref1} \) and b) \( \Delta H^0 \) constructed using 200 bootstrap parameter estimates when different levels of input uncertainties are considered for both \( F_{H_2}^0 m \) and \( F_{nC_6}^0 m \). The ends of the whiskers correspond to approximate 95% confidence intervals. ........................................... 132

**Figure 5-3** Prediction uncertainty boxplots for a) mean prediction of \( F_{2MP} \), b) a single future measurement of \( F_{2MP} \), c) mean prediction of \( F_{C_3} \) and d) a single future measurement of \( F_{C_3} \) constructed based on 200 MC simulations. Predictions correspond to the experimental condition at the 15th run in Table E.1 (\( T = 513 \, K, P = 1 \, MPa, \) and \( F_{H_2}^0 m = 1.3 \times 10^{-3} \, \frac{mol}{s} \) and \( F_{nC_6}^0 m = 2.56 \times 10^{-5} \, \frac{mol}{s} \)). The whisker ends correspond to approximate 95% confidence intervals. .................................................. 134

**Figure 5-4** Prediction uncertainty boxplots for a) mean prediction of \( F_{2MP} \), b) a single future measurement of \( F_{2MP} \), c) mean prediction of \( F_{C_3} \) and d) a single future measurement of \( F_{C_3} \) constructed based on 200 MC simulations. The model predictions are obtained for a new experimental condition (\( T = 580 \, K, P = 1.0 \, MPa, \) \( F_{H_2}^0 m = 1.3 \times 10^{-3} \, \frac{mol}{s} \) and \( F_{nC_6}^0 m = 2.56 \times 10^{-5} \, \frac{mol}{s} \)). The whisker ends correspond to approximate 95% confidence intervals. .................................................. 135

**Figure 5-5** Boxplots for yield predictions constructed based on 200 MC simulations: a) for mean yield predictions at the target experimental condition with \( T = 513 \, K, P = 1 \, MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \, \frac{mol}{s} \) and \( F_{nC_6}^0 m = 2.56 \times 10^{-5} \, \frac{mol}{s} \); b) for mean yield predictions at a new experimental condition with higher temperature where \( T = 580 \, K, P = 1.0 \, MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \, \frac{mol}{s} \) and \( F_{nC_6}^0 m = 2.56 \times 10^{-5} \, \frac{mol}{s} \). The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals). ................................................................. 137

**Figure 5-6** Selectivity predictions for 3MP constructed based on 200 MC simulations: a) boxplot for mean selectivity predictions at the experimental condition with \( T = 513 \, K, P = 1 \, MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \, \frac{mol}{s} \) and \( F_{nC_6}^0 m = 2.56 \times 10^{-5} \, \frac{mol}{s} \) for two levels of input uncertainties; b) the corresponding scatter plot for computed values of \( F_{3MP}^0 \) vs. \( F_{nC_6}^0 - F_{nC_6} \) to investigate the correlation between the numerator and
denominator in the selectivity expression in equation 5-12. Ends of the whiskers correspond to the approximate 95 % confidence intervals.................................................................137

**Figure A. 1** Influence of the number of estimated parameters from the ranked list on the value of objective function in Case I parameter estimation .................................................................150

**Figure A. 2** Simulation results corresponding to Case I for a) Initiator molecules $I$, b) initiator fragments attached to polymer chains $\tilde{I}$, c) Monomer, and d) pendant vinyl groups $V_p$ .................................................................152

**Figure A. 3** Simulation results obtained from parameter estimates in Case I for concentration of a) carbon-centered initiator radicals $I^*$, b) small phosphorous-centered radicals $\tilde{I}^*$, and c) large molecules with phosphorous-centered radicals $I_p^*$ .................................................................153

**Figure A. 4** Numerical simulation corresponding to Case I for a) Regular radical ends $R^{\text{red}}$, and b) Radicals at the branch points $B^*$ ..................................................................................................................154

**Figure A. 5** Comparison between model predictions (curves) from Case I and their corresponding measurements (shown with $\times$,$\times$,$\times$,$\times$, and $\times$) for different light intensities. The same colors used for the prediction curves and the corresponding measurement data........................................................................154

**Figure A. 6** Comparison between final monomer conversion predictions and their corresponding measured values obtained using estimated parameters from Case I for short-time-exposure experiments with light intensity of 20 W/m$^2$. Straight lines between the model predictions are shown to guide the eye........155

**Figure A. 7** Influence of the number of estimated parameters from the parameter ranks list on the value of objective function in Case II parameter estimation..................................................................................................156

**Figure A. 8** Simulation results corresponding to parameter estimation in Case II for concentration of a) unreacted initiators $I$, b) Initiator fragments attached to other molecules $\tilde{I}$, c) monomer $M$, and d) pendant vinyl groups $V_p$ ..................................................................................................................158

**Figure A. 9** Simulation results corresponding to the Case II parameter estimation for concentrations of a) carbon-centered initiator radicals $I^*$, b) phosphorous-centered initiator radicals $\tilde{I}^*$, and c) large molecules with phosphorous-centered radicals $I_p^*$ ..................................................................................................159

**Figure A. 10** Simulation results associated with the Case II parameter estimation for concentration of a) Regular radical ends $R^{\text{red}}$, and b) radical at the branch points $B^*$, and c) tertiary radicals $T^*$ ............160

**Figure A. 11** Comparison between model predictions and their corresponding measurement values (shown with $\times$,$\times$,$\times$,$\times$, and $\times$) obtained using parameter estimates in Case II for some of the conversion vs. time data at some of the light intensities. Similar colors used for the prediction curves and the corresponding measurement data..................................................................................................161
Figure A. 12  Comparison between model predictions and the corresponding measurement data obtained using the estimated parameters in Case II for different pulse experiments at $I_{t,0}=20$ W/m$^2$. Straight lines between the model predictions are shown to guide the eye.

Figure A. 13  Influence of the number of estimated parameters from the parameter ranks list on the value of objective function in Case III parameter estimation.

Figure A. 14  Simulation results corresponding to parameter estimation in Case III for concentration of a) unreacted initiators $I$, b) Initiator fragments attached to other molecules $\tilde{I}$, c) monomer $M$, d) pendant vinyl groups $V_p$.

Figure A. 15  Simulation results corresponding to the Case III parameter estimation for concentrations of a) carbon-centered initiator radicals $I^*$, b) small phosphorous-centered initiator radicals $\tilde{I}^*$, and c) large phosphorous-centered radical $I_p^*$.

Figure A. 16  Simulation results corresponding to the Case III parameter estimation for concentrations of a) regular radical ends $R^{\text{reg}}$, b) radicals at the branch points $B^*$, c) cyclized radical ends $C^*$, and d) tertiary radicals $T^*$.

Figure A. 17  Experimental data (shown with $\times, \cdot, \cdot, \cdot, \cdot, \cdot$) and corresponding model predictions obtained using parameter estimates obtained in Case III for some of the conversion data at some of the light intensities. Similar colors used for the prediction curves and the corresponding measurement data.

Figure A. 18  Comparison between final monomer conversion predictions and their corresponding measured values obtained using estimated parameters from Case III for short-time-exposure experiments with light intensity of 20 W/m$^2$. Straight lines between the model predictions are shown to guide the eye.

Figure A. 19  Influence of the number of estimated parameters from the ranked list on the value of objective function in Case IV.

Figure A. 20  Simulation results corresponding to parameter estimation in Case IV for concentration of a) unreacted initiators $I$, b) Initiator fragments attached to other molecules $\tilde{I}$, c) monomer $M$, and d) pendant vinyl groups $V_p$.

Figure A. 21  Simulation results corresponding to the Case IV parameter estimation for concentrations of a) carbon-centered initiator radicals $\tilde{I}^*$, and c) large molecules with phosphorous-centered radicals $I_p^*$. 

Figure A. 22  Simulation results corresponding to the Case IV parameter estimation for concentrations of a) regular radical ends $R^{\text{reg}}$, b) radicals at the branch points $B^*$, c) cyclized radical ends $C^*$, and d) tertiary radicals $T^*$. 
Figure A. 23 Experimental data (shown with ×××× and ×) and corresponding model predictions obtained using parameter estimates obtained in Case IV for some of the conversion data at some of the light intensities. Similar colors used for the prediction curves and the corresponding measurement data. 

Figure A. 24 Comparison between final monomer conversion predictions and their corresponding measured values obtained using estimated parameters from Case IV for short-time-exposure experiments with light intensity of 20 W/m². Straight lines between the model predictions are shown to guide the eye.

Figure A. 25 Influence of number of estimated parameters from the ranking list on the optimal value of objective function in the final parameter estimation step.

Figure A. 26 Simulation results corresponding to the final parameter estimation step for concentration of a) unreacted initiators $I$, b) Initiator fragments attached to other molecules $\tilde{I}$, c) monomer $M$, d) pendant vinyl groups $V_p$.

Figure A. 27 Simulation results corresponding to the final parameter estimation step for concentrations of a) carbon-centered initiator radicals $I^*$, b) phosphorous-centered initiator radicals $\tilde{I}^*$, and c) phosphorous-centered radicals attached to large molecules $\tilde{I}^p$.

Figure A. 28 Simulation results corresponding to the final parameter estimation step for concentrations of a) regular radical ends $R^{\equiv\cdot}$, b) radicals at the branch points $B^{\equiv\cdot}$, c) cyclized radical ends $C^{\equiv\cdot}$, and d) tertiary radicals $T^{\equiv\cdot}$.

Figure B. 1 Decomposition of a) initiator $I$ (BAPO) and b) the remaining carbon-phosphorous bond in $\tilde{I}$.

Figure B. 2 Formation of different radical end groups including: a) regular radical ends $R^{\equiv\cdot}$, b) cyclized radical ends $C^{\equiv\cdot}$, c) tertiary radicals $T^{\equiv\cdot}$, d) branch-point radicals $B^{\equiv\cdot}$, and e) peroxidic radicals $O^{\equiv\cdot}$. Ellipses are used to highlight the end group of interested.

Figure B. 3 Experimental data ($\Delta$,○,×) and corresponding model predictions shown using curves with the same colors for experiments with oxygen at air level with a) $[BAPO] = 1$ wt%, $z_I$ targets 8 μm and $I_{l,0}$ targets 1000 W/m², b) $[BAPO] = 4$ wt%, $z_I$ targets 8 μm and $I_{l,0}$ targets 1000 W/m², and c) $[BAPO] = 4$ wt%, $z_I = 11.8$ μm, and light intensities $I_{l,0} = 200$ W/m² and $I_{l,0} = 500$ W/m².

Figure B. 4 Model predictions (curved lines) and the corresponding FTIR measured values for experiments in the absence (×) and in the presence of oxygen (Δ) using $[BAPO] = 4$ wt%, $z_I \sim 12$ μm, and $I_{l,0} \sim 750$ W/m².

Figure B. 5 Experimental data ($\Delta$,○,×) and corresponding model predictions shown using curves with the same colors for experiments with oxygen at air level with a) $[BAPO] = 1$ wt%, $z_I$ targets 16 μm, and
\( I_l, 0 \text{ targets 1000 W/m}^2, b) [BAPO] = 4 \text{ wt\%}, z_f \text{ targets 16 \mu m, and } I_{l,0} \text{ targets 1000 W/m}^2, \text{ and } c) [BAPO] = 4 \text{ wt\%, } z_f \text{ targets 16 \mu m, and } I_{l,0} \text{ targets 5000 W/m}^2. \)

**Figure E. 1** Boxplots for uncertainties of the parameters based on 200 bootstrap parameter estimates when both uncertain inputs \( F_{H_2}^0 \) and \( F_{c_6}^0 \) are uncertain: a) kinetic rate of the second reaction at the reference temperature, b) kinetic rate of the third reaction at the reference temperature, c) equilibrium constant at the reference temperature, d) activation energy for the first reaction, e) activation energy for the second reaction and f) activation energy for the third reaction. End of the whiskers correspond to an approximate 95\% confidence intervals.

**Figure E. 2** Prediction uncertainty boxplots for a) mean prediction of \( F_{3MP} \) and b) a single-future measurement of \( F_{3MP} \) constructed based on 200 MC simulations. Predictions correspond to the experimental condition at the 15\textsuperscript{th} run in Table E.1 (\( T = 513 \text{ K, } P = 1 \text{ MPa, } F_{H_2}^0 m = 1.3 \times 10^{-3} \text{ mol/s, and } F_{nC_6}^0 m = 2.56 \times 10^{-5} \text{ mol/s}. \)). The whisker ends correspond to approximate 95\% confidence intervals.

**Figure E. 3** Prediction uncertainty boxplots for a) mean prediction of \( F_{3MP} \) and b) a single-future measurement of \( F_{3MP} \) constructed based on 200 MC simulations corresponding to a new experimental condition (\( T = 580 \text{ K, } P = 1 \text{ MPa, } F_{H_2}^0 m = 1.3 \times 10^{-3} \text{ mol/s, and } F_{nC_6}^0 m = 2.56 \times 10^{-5} \text{ mol/s}. \)). The whisker ends correspond to approximate 95\% confidence intervals.

**Figure E. 4** Boxplots for yield predictions constructed based on 200 MC simulations: a) for a single-future yield predictions at the target experimental condition with \( T = 513 \text{ K, } P = 1 \text{ MPa, } F_{H_2}^0 m = 1.3 \times 10^{-3} \text{ mol/s, and } F_{nC_6}^0 m = 2.56 \times 10^{-5} \text{ mol/s}; \) b) for single-future yield predictions at a new experimental condition with higher temperature where \( T = 580 \text{ K, } P = 1.0 \text{ MPa, } F_{H_2}^0 m = 1.3 \times 10^{-3} \text{ mol/s, and } F_{nC_6}^0 m = 2.56 \times 10^{-5} \text{ mol/s. \) The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95\% confidence intervals).

**Figure E. 5** Boxplots for \( S2MP \) predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with \( T = 513 \text{ K, } P = 1 \text{ MPa, } F_{H_2}^0 m = 1.3 \times 10^{-3} \text{ mol/s, and } F_{nC_6}^0 m = 2.56 \times 10^{-5} \text{ mol/s}; \) c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where \( T = 580 \text{ K, } P = 1.0 \text{ MPa, } F_{H_2}^0 m = 1.3 \times 10^{-3} \text{ mol/s, and } F_{nC_6}^0 m = 2.56 \times 10^{-5} \text{ mol/s. \) The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95\% confidence intervals).

**Figure E. 6** Boxplots for \( S3MP \) predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with \( T = 513 \text{ K, } P = \)
1 \text{MPa}, F_{H_2}^0 = 1.3 \times 10^{-3} \text{mol s}^{-1}, \text{and } F_{nC_6}^0 = 2.56 \times 10^{-5} \text{mol s}^{-1}; \text{ c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where } T = 580 \text{K}, P = 1.0 \text{MPa}, F_{H_2}^0 = 1.3 \times 10^{-3} \text{mol s}^{-1}, \text{and } F_{nC_6}^0 = 2.56 \times 10^{-5} \text{mol s}^{-1}. \text{ The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).} \text{ Figure E. 7 Boxplots for uncertainties of the parameters based on 200 bootstrap parameter estimates when independent variable } F_{nC_6}^0 \text{ is uncertain: a) kinetic rate of the first reaction at the reference temperature, b) kinetic rate of the second reaction at the reference temperature and d) equilibrium constant at the reference temperature. End of the whiskers correspond to an approximate 95% confidence intervals.} \text{ Figure E. 8 Boxplots for uncertainties of the parameters based on 200 bootstrap parameter estimates when independent variable } F_{nC_6}^0 \text{ is uncertain: a) activation energy of the first reaction, b) activation energy of the second reaction, c) activation energy of the third reaction and d) enthalpy of physisorption. End of the whiskers correspond to approximate 95% confidence intervals.} \text{ Figure E. 9 Prediction uncertainty boxplots for a) mean prediction of F2MP, b) a single-future measurement of F2MP, c) mean prediction of F3MP, d) a single-future measurement of F3MP, e) mean prediction of FC3 and f) a single-future measurement of FC3 constructed based on 200 MC simulations. Predictions correspond to the experimental condition at the 15th run in Table E.1 (} T = 513 \text{K}, P = 1 \text{MPa}, \text{ and } F_{H_2}^0 = 1.3 \times 10^{-3} \text{mol s}^{-1}, \text{ and } F_{nC_6}^0 = 2.56 \times 10^{-5} \text{mol s}^{-1} \text{) when independent variable } F_{nC_6}^0 \text{ is uncertain. The whisker ends correspond to approximate 95% confidence intervals.} \text{ Figure E. 10 Prediction uncertainty boxplots for a) mean prediction of F2MP, b) a single-future measurement of F2MP, c) mean prediction of F3MP, d) a single-future measurement of F3MP, e) mean prediction of FC_3 and f) a single-future measurement of FC_3 constructed based on 200 MC simulations. Predictions correspond to the experimental at a new condition (} T = 580 \text{K}, P = 1 \text{MPa}, \text{ and } F_{H_2}^0 = 1.3 \times 10^{-3} \text{mol s}^{-1}, \text{ and } F_{nC_6}^0 = 2.56 \times 10^{-5} \text{mol s}^{-1} \text{) where independent variable FC60 is uncertain. The whisker ends correspond to approximate 95% confidence intervals.} \text{ Figure E. 11 Boxplots for yield predictions constructed based on 200 MC simulations: a) for mean yield predictions and b) for a single-new prediction at the target experimental condition with } T = 513 \text{K}, P = 1 \text{MPa}, \text{ and } F_{H_2}^0 = 1.3 \times 10^{-3} \text{mol s}^{-1}, \text{ and } F_{nC_6}^0 = 2.56 \times 10^{-5} \text{mol s}^{-1}; \text{ c) for mean yield predictions and d) a single-new prediction at a new experimental condition with higher temperature where } T = 580 \text{K}, P = 1.0 \text{MPa}, \text{ and } F_{H_2}^0 = 1.3 \times 10^{-3} \text{mol s}^{-1}, \text{ and } F_{nC_6}^0 = 2.56 \times 10^{-5} \text{mol s}^{-1}. \text{ The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).}
Figure E. 12 Boxplots for $S_{2MP}$ predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with $T = 513\ K, P = 1\ MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}, and F_{nc}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}; c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where $T = 580\ K, P = 1.0\ MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}, and F_{nc}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).

Figure E. 13 Boxplots for $S3MP$ predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with $T = 513\ K, P = 1\ MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}, and F_{nc}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}; c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where $T = 580\ K, P = 1.0\ MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}, and F_{nc}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).

Figure E. 14 Boxplots for uncertainties of the parameters based on 200 bootstrap parameter estimates when independent variable $F_{H_2}^0$ is uncertain: a) kinetic rate of the first reaction at the reference temperature, b) kinetic rate of the second reaction at the reference temperature, c) kinetic rate of the third reaction at the reference temperature and d) equilibrium constant at the reference temperature. End of the whiskers correspond to an approximate 95% confidence intervals.

Figure E. 15 Boxplots for uncertainties of the parameters based on 200 bootstrap parameter estimates when independent variable $FH20$ is uncertain: a) activation energy of the first reaction, b) activation energy of the second reaction, c) activation energy of the third reaction and d) enthalpy of physisorption. End of the whiskers correspond to an approximate 95% confidence intervals.

Figure E. 16 Prediction uncertainty boxplots for a) mean prediction of $F_{2MP}$, b) a single-future measurement of $F_{2MP}$, c) mean prediction of $F_{3MP}$, d) a single-future measurement of $F_{3MP}$, e) mean prediction of $F_{C_3}$ and f) a single-future measurement of $F_{C_3}$ constructed based on 200 MC simulations. Predictions correspond to the experimental condition at the 15th run in Table E.1 ($T = 513\ K, P = 1\ MPa$, and $F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}, and F_{nc}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$) when independent variable $FH20$ is uncertain. The whisker ends correspond to approximate 95% confidence intervals.

Figure E. 17 Prediction uncertainty boxplots for a) mean prediction of $F_{2MP}$, b) a single-future measurement of $F_{2MP}$, c) mean prediction of $F_{3MP}$, d) a single-future measurement of $F_{3MP}$, e) mean prediction of $F_{C_3}$ and f) a single-future measurement of $F_{C_3}$ constructed based on 200 MC simulations. Predictions correspond to the experimental at a new condition ($T = 580\ K, P = 1\ MPa$, and $F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}, and F_{nc}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$).
where independent variable $FH20$ is uncertain. The whisker ends correspond to approximate 95% confidence intervals.

**Figure E. 18** Boxplots for yield predictions constructed based on 200 MC simulations: a) for mean yield predictions and b) for a single-new prediction at the target experimental condition with $T = 513 K, P = 1 MPa, F_{H2}^0 m = 1.3 \times 10^{-3} mol_s, and F_{nC}_6^0 m = 2.56 \times 10^{-5} mol_s$; c) for mean yield predictions and d) a single-new prediction at a new experimental condition with higher temperature where $T = 580 K, P = 1.0 MPa, F_{H2}^0 m = 1.3 \times 10^{-3} mol_s, and F_{nC}_6^0 m = 2.56 \times 10^{-5} mol_s$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).

**Figure E. 19** Boxplots for $S_{2MP}$ predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with $T = 513 K, P = 1 MPa, F_{H2}^0 m = 1.3 \times 10^{-3} mol_s, and F_{nC}_6^0 m = 2.56 \times 10^{-5} mol_s$; c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where $T = 580 K, P = 1.0 MPa, F_{H2}^0 m = 1.3 \times 10^{-3} mol_s, and F_{nC}_6^0 m = 2.56 \times 10^{-5} mol_s$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).

**Figure E. 20** Boxplots for $S_{3MP}$ predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with $T = 513 K, P = 1 MPa, F_{H2}^0 m = 1.3 \times 10^{-3} mol_s, and F_{nC}_6^0 m = 2.56 \times 10^{-5} mol_s$; c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where $T = 580 K, P = 1.0 MPa, F_{H2}^0 m = 1.3 \times 10^{-3} mol_s, and F_{nC}_6^0 m = 2.56 \times 10^{-5} mol_s$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).

xviii
List of Tables

Table 2-1 A summary of modeling studies for photopolymerization of acrylate and methacrylate monomers ................................................................. 12
Table 2-2 Kinetic phenomena considered during four parameter-estimation scenarios .................. 13
Table 2-3 List of reactions involved in photopolymerization reaction scheme ............................. 22
Table 2-4 Modeling assumptions .................................................................................. 24
Table 2-5 Algebraic equations used in the model ............................................................. 26
Table 2-6 Dynamic material balance equations for chemical species and functional groups.......... 32
Table 2-7 List of parameter initial values and corresponding bounds used during parameter estimation 39
Table 2-8 Result of parameter estimation for different cases of parameter estimation ................. 41
Table 2-9 List of final parameter estimates ..................................................................... 42
Table 3-1 Summary of studies where EVM used for reactivity-ratio estimation ......................... 57
Table 3-2 Summary of studies involving other Chemical Engineering models where EVM used for parameter estimation ............................................. 59
Table 3-3 Error-in-variables model equations for different assumptions about replicate runs ........ 62
Table 3-4 Objective Functions for Parameter Estimation in Error-in-Variables Models with Different Assumptions about Replicate Runs ...................................................... 62
Table 3-5 Procedure for EVM parameter estimation using pseudo-replicate data ...................... 65
Table 3-6 Important reactions corresponding to a free-radical copolymerization terminal model .... 66
Table 3-7 The experimental data collected for copolymerization of n-Butyl Methacrylate and n-Butyl Acrylate from Ren et al ............................................................... 68
Table 3-8 EVM Mayo-Lewis for the AK, TR1 and PR situations .............................................. 70
Table 3-9 EVM-based objective functions for reactivity-ratio estimation corresponding to the AK, TR1 and PR situations ................................................................. 71
Table 3-10 Estimated (assumed) values of measurement variance $\sigma^2_{lnF_1}$ and estimated reactivity ratios corresponding to EVM situations with different underlying assumptions ................. 73
Table 4-1 Summary of studies involving prediction uncertainty quantification when only input uncertainties are important ................................................................. 88
Table 4-2 MC method for obtaining approximate 95% confidence intervals for mean model predictions at an experimental condition of interest ...................................................... 90
Table 4-3 Bootstrap steps to obtain parameter uncertainties when input uncertainties are important .... 93
**Table 4-4** Summary of studies involving parameter uncertainty quantification when input measurements are uncertain.................................................................................................................. 98

**Table 4-5** MC method for obtaining approximate 95% confidence intervals for mean model predictions at an experimental condition of interest .................................................................................................. 101

**Table 4-6** Summary of studies involving prediction uncertainty quantification from parameters and inputs uncertainties ............................................................................................................................................... 102

**Table 5-2** MC method for obtaining approximate 95% confidence intervals for mean model predictions at an experimental condition of interest .................................................................................................. 121

**Table 5-3** Bootstrap steps to obtain parameter uncertainties for multi-response error-in-variable models with several uncertain inputs using pseudo-replicate experiments .................................................................................................. 122

**Table 5-4** Reaction Scheme Proposed for nC₆ Hydroisomerization ............................................................................................................................................... 124

**Table 5-5** Algebraic equations based on the mole balance isomerization for different species ............ 124

**Table 5-6** Kinetic rate and equilibrium constant equations for solving balance equations and the corresponding parameters requiring estimation .................................................................................................. 125

**Table 5-7** Different uncertainty level for inputs with measurement uncertainties ................................ 126

**Table 5-8** The objective function and their corresponding decision variables for EVM and WLS objective functions ............................................................................................................................................... 129

**Table 5-9** Values of parameter estimates obtained by minimizing objective functions in Table 5-8 when different levels of uncertainties in model inputs .................................................................................................. 130

**Table A.1** The list of ranked parameters and their corresponding estimates for Case I parameter estimation ............................................................................................................................................... 151

**Table A.2** The list of ranked parameters and their corresponding estimates for Case II parameter estimation ............................................................................................................................................... 157

**Table A.3** The list of ranked parameters and their corresponding estimates for Case III parameter estimation ............................................................................................................................................... 164

**Table A.4** The list of ranked parameters and their corresponding estimates for Case IV parameter estimation ............................................................................................................................................... 170

**Table A.5** List of final parameter estimates and their approximate 95% confidence intervals .............. 180

**Table A.6** Correlation matrix elements greater than or equal to 0.8 or less than or equal to -0.8 ........ 182

**Table B.2** - List of reactions accounting for effect of oxygen on photopolymerization of HDDA ........ 196

**Table B.3** Material balances for chemical species and functional groups .............................................. 199

**Table B.4** List of parameters and their estimates obtained from Abdi’s model ........................................... 205

**Table B.5** List of parameters accounting for effects of oxygen inhibition .................................................. 207
Table C. 1 Proposed bootstrap steps to obtain uncertainties of reactivity-ratio parameters estimated from data in Table 7 using EVM when output measurement variances are assumed known.............................. 223
Table C. 2 Bootstrap steps to obtain uncertainties of parameter estimated from EVM in TR1 situation 224
Table C. 3 Bootstrap steps to obtain uncertainties in parameter estimates from EVM in PR situation... 225
Table E. 1 Experimental data from Toch et al. for hydroisomerization of n-hexane in a gas-phase, well-mixed reactor .......................................................... 235
# Nomenclature

## Nomenclature for Chapter 2

### Abbreviation

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAPO</td>
<td>bis-Acylphosphine oxide</td>
</tr>
<tr>
<td>HDDA</td>
<td>1,6-Hexanediol diacrylate</td>
</tr>
<tr>
<td>SSH</td>
<td>Stationary state hypothesis</td>
</tr>
</tbody>
</table>

### Symbols

- $A_C$: Free-volume parameter for the cyclization reaction
- $A_b$: Free-volume parameter for the branching reactions
- $A_{bb}$: Free-volume parameter for the backbiting reactions
- $A_f$: Free-volume parameter for the initiator efficiency
- $A_{in,j}$: Free-volume parameter for the initiation reactions ($j = IM, IV_p, IM_p, \text{ and } \bar{IV}_p$)
- $A_p$: Free-volume parameter for propagation reactions
- $A_{t_{in,j}}$: Free-volume parameters for termination reactions involving initiator radicals ($j = II, \bar{II}, \text{ and } I_p I_p$)
- $B^*$: Radicals at the branch points
- $c$: Speed of light
- $C^*$: Cyclized radical ends
- $f$: Efficiency for decomposition of the initiator molecules
- $f_0$: Efficiency for decomposition of the initiator molecules at the start of the polymerization
- $\bar{f}$: Efficiency for decomposition of carbon-phosphorous bonds $\bar{I}$ attached to polymer chains ($\bar{I}$)
- $h$: Planck’s constant
- $\bar{I}$: Carbon-phosphorous bonds attached to polymer molecules
- $I^*$: Carbon-centered initiator radicals
- $I^*$: Phosphorous-centered initiator radicals
- $I_p^*$: Phosphorous-centered initiator radicals generated from decomposition of $\bar{I}$
- $I_{t,0}$: Light intensity at the top surface of the film
- $I_{\bar{I}}$: Average light intensity within the film
- $J$: Objective function
- $k_{b,j}$: Rate coefficient for branching reactions involving $R^* (j = R), C^* (j = C), B^* (j = B)$, and $T^* (j = T)$
- $k_{b,j_0}$: Rate coefficient for branching reactions when they are kinetically-controlled
- $k_{bb,j}$: Rate coefficient for backbiting reaction involving $R^* (j = R)$ and $C^* (j = C)$
- $k_{bb_0}$: Rate coefficient for the backbiting reactions $k_{bb}$ when they are kinetically-controlled
- $k_C$: Rate coefficient for cyclization
- $k_{C_0}$: Rate coefficient for cyclization when it is kinetically-controlled
- $k_d$: Rate coefficient for decomposition of initiator
\( k_{\text{d}} \)  Rate coefficient for decomposition of \( \bar{I} \)
\( k_{\text{in},\text{IM}} \)  Initiation rate coefficient for initiation reaction involving a carbon-centred radical \( \bar{I}^* \) and a monomer
\( k_{\text{in},\text{JM}} \)  Initiation rate coefficient for initiation reaction involving a phosphorous-centered radical \( \bar{I}^* \) and a monomer
\( k_{\text{in},\text{IpM}} \)  Initiation rate coefficient for initiation reaction involving a phosphorous-centered radical \( \bar{I}_{pM} \) and a monomer
\( k_{\text{in},\text{IvP}} \)  Initiation rate coefficient for initiation reaction involving a carbon-centred radical \( \bar{I}^* \) and a pendant vinyl group
\( k_{\text{in},\text{IvP}} \)  Initiation rate coefficient for initiation reaction involving a phosphorous-centred radical \( \bar{I}_{p} \) and a pendant vinyl group
\( k_{\text{in},\text{IvPV}} \)  Initiation rate coefficient for initiation reaction involving a phosphorous-centered radical \( \bar{I}_{p} \) and a pendant vinyl
\( k_{\text{in},\text{j0}} \)  Initiation rate coefficient involving different species when the corresponding reaction is kinetically-controlled (\( j=\text{IM}, \bar{I}_{pM}, \text{IvP}, \bar{I}_{pV} \text{, and } \text{IvPV} \))
\( k_{p,j} \)  Rate coefficient for the propagation involving \( R_{\text{CH}} \) (\( j=\text{R}, \text{C}^* \), \( j=\text{R}^* \), \( j=\text{B}^* \), \( j=\text{B} \), and \( j=\text{T}^* \))
\( k_{p,j0} \)  Rate coefficient for the propagation reactions when they are kinetically controlled
\( k_{t,ij} \)  Self-termination rate coefficient for reactions involving large macromolecules \( j \) (\( j=\text{RR}, \text{CC}, \text{BB}, \text{and } \text{TT} \))
\( k_{t,ij0} \)  Rate coefficients for termination reactions \( k_{t,ij} \) when they are kinetically-controlled
\( k_{t,ij} \)  Cross-termination rate coefficients involving the radicals of type \( i \) (\( i=\text{I}, \bar{I}_{p}, \text{R}, \text{C}, \text{B}, \text{and } \text{T} \)) and \( j \) (\( j=\text{I}, \bar{I}_{p}, \text{R}, \text{C}, \text{B}, \text{and } \text{T} \))
\( k_{t,ij} \)  Self-termination rate coefficient for reactions involving initiator radicals (\( j=\text{I} \), \( \bar{I}_{p} \))
\( k_{t,ij0} \)  Self-termination rate coefficient for reactions involving initiator radicals when they are kinetically-controlled
\( M \)  Monomer
\( N_{X_v} \)  Number of measured values for overall-vinyl-group conversions
\( N_{X_M} \)  Number of measured values for monomer conversions
\( \tau_{\text{exp}} \)  Time duration of light exposure
\( \nu_{f,0} \)  Free volume fraction before polymerization starts (i.e., \( X_M = 0 \))
\( \nu_{f,1} \)  Free volume fraction when all monomer is consumed (i.e., \( X_M = 1.0 \))
\( \nu_f \)  Free volume fraction
\( \nu_{f,c,b} \)  Critical free volume parameter for branching reactions
\( \nu_{f,c,bb} \)  Critical free-volume fraction for backbiting reactions
\( \nu_{f,c,in,j} \)  Critical free-volume fraction for the \( j^{th} \) initiation reaction (\( j=\text{IM}, \text{IvP}, \bar{I}_{pM}, \text{IvPV} \text{, and } \text{IvPV} \))
\( \nu_{f,c,p} \)  Critical free-volume fraction for propagation reactions
\( \nu_{f,c,tip,ij} \)  Critical free volume parameters for termination reactions involving initiator radicals (\( j=\text{I}, \bar{I}_{p} \), and \( \text{I}_{p} \)).
\( X_M \)  Monomer conversion
\( X_{M,i} \)  Monomer conversion prediction corresponding to the \( i^{th} \) measurement
\( X_{M_{\text{mes}},i} \)  Measurement for the monomer conversion at condition corresponding to the \( i^{th} \) measurement
\( X_{V,i} \)  Overall vinyl-group conversion prediction corresponding to the \( i^{th} \) conversion data
\( X_{V_{\text{meas}},i} \)  Value of the \( i^{th} \) overall-vinyl group conversion measurement
$z$ \hspace{1em} Distance from the top surface of the film

$z_f$ \hspace{1em} Film thickness

**Greek symbols**

$\sigma_V^2$ \hspace{1em} Measurement variance for the vinyl group conversions

$\sigma_M^2$ \hspace{1em} Measurement variance for the monomer conversions

$\lambda$ \hspace{1em} Light wavelength

$\phi$ \hspace{1em} Quantum yield of initiator

$\epsilon$ \hspace{1em} Extinction coefficient
Nomenclature for Chapter 3

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK</td>
<td>Assumed known</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>EVM</td>
<td>Error-in-variables model</td>
</tr>
<tr>
<td>JC</td>
<td>Joint confidence</td>
</tr>
<tr>
<td>JCR</td>
<td>Joint confidence region</td>
</tr>
<tr>
<td>PR</td>
<td>Pseudo replicate</td>
</tr>
<tr>
<td>TR1</td>
<td>True replicate of type 1</td>
</tr>
<tr>
<td>TR2</td>
<td>True replicate of type 2</td>
</tr>
<tr>
<td>WLS</td>
<td>Weighted least squares</td>
</tr>
</tbody>
</table>

Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>Counter for the iterations of the bootstrapping algorithm</td>
</tr>
<tr>
<td>b_{max}</td>
<td>Maximum number of the bootstrapping algorithm</td>
</tr>
<tr>
<td>e</td>
<td>The relative change of parameter estimates in two subsequent iterations of the algorithm used in PR situations to estimate output measurement variances</td>
</tr>
<tr>
<td>f_1</td>
<td>Composition for comonomer 1</td>
</tr>
<tr>
<td>f_{1,i}</td>
<td>True value for the composition of the comonomer of type 1 at the (i^{th}) target condition</td>
</tr>
<tr>
<td>f_{1,ij}</td>
<td>The (j^{th}) true value for the composition of comonomer type 1 at the (i^{th}) target condition</td>
</tr>
<tr>
<td>f_{1m,i}</td>
<td>The measured value of for the composition of the comonomer type 1 obtained at the (i^{th}) target experimental condition</td>
</tr>
<tr>
<td>F_1</td>
<td>The copolymer composition for component 1</td>
</tr>
<tr>
<td>F_{1,ij}</td>
<td>The (j^{th}) repeated value of the copolymer composition of type 1 obtained at the (i^{th}) target condition</td>
</tr>
<tr>
<td>F_{1m,ij}</td>
<td>The (j^{th}) measurement for the copolymer composition of type 1 at the (i^{th}) target experimental condition</td>
</tr>
<tr>
<td>g</td>
<td>A vector of solutions for the model predictions of a multi-output model</td>
</tr>
<tr>
<td>g</td>
<td>A solution for the model prediction of a single-output model</td>
</tr>
<tr>
<td>g_{aug,s}</td>
<td>Scaled augmented model prediction vector containing scaled output and scaled input predictions</td>
</tr>
<tr>
<td>g_k</td>
<td>Solution for the prediction of the (k^{th}) output</td>
</tr>
<tr>
<td>I_N</td>
<td>An identity matrix of size (N)</td>
</tr>
<tr>
<td>J_{TR1}</td>
<td>EVM-based objective function corresponding to the experiments involving true replicates of type 1</td>
</tr>
</tbody>
</table>
$J_{TR2}$ EVM-based objective function corresponding to the experiments involving true replicates of type 2

$J_{PR}$ EVM-based objective function corresponding to the data set obtained from a pseudo-replicate situation

$\hat{J}_{\theta_p}$ The Jacobian matrix of the model predictions with respect to the primary parameters

$k_{11}$ Rate coefficient for the reaction between the growing polymer of type 1 and the comonomer of type 1

$k_{22}$ Rate coefficient for the reaction between the growing polymer of type 2 and the comonomer of type 2

$k_{12}$ Rate coefficient for the reaction between the growing polymer of type 1 and the comonomer of type 2

$k_{21}$ Rate coefficient for the reaction between the growing polymer of type 2 and the comonomer of type 1

$M_1$ Monomer of type 1

$M_2$ Monomer of type 2

$N$ Number of target experimental condition

$N_{U}$ Number of uncertain model input

$N_Y$ Number of output predictions at each experimental condition

$N_\theta$ Number of parameters

$n_i$ Number of replicate experiments at the $i^{th}$ run condition

$p$ Counter for the elements of parameter vector $\theta$

$\hat{Q}$ The weighting matrix for computing the covariance matrix of the primary parameters

$R_1^*$ Macroradical with terminal monomer of type 1

$R_2^*$ Macroradical with terminal monomer of type 2

$U_{ij}$ Input measurements random variable for a model with a single uncertain input for the $j^{th}$ replicate run at the $i^{th}$ experimental condition

$U_{ij}$ Vector of uncertain input measurements random variable for the $j^{th}$ replicate at the $i^{th}$ experimental condition

$r_1$ Reactivity ratio indicating the relative tendency of macroradical type 1 to react with monomers type 1 and type 2

$r_2$ Reactivity ratio indicating the relative tendency of macroradical type 2 to react with monomers type 1 and type 2

$s$ Iteration counter in calculating the output measurement variance using pseudo-replicate data

$u_{ij}$ True value of the input for $j^{th}$ replicate run of the $i^{th}$ target condition

$u_{ij}$ Vector of true values of uncertain inputs for the $j^{th}$ replicate run of $i^{th}$ target condition
Vector of measured values for uncertain inputs obtained from the $i^{th}$ experimental condition

Vector of perfectly-known model inputs for the $i^{th}$ target condition

Output measurement random variable corresponding to a single-output model for the $j^{th}$ replicate run at the $i^{th}$ experimental target condition

Output measurement corresponding for the $k^{th}$ output of a multi-output model obtained for the $j^{th}$ replicate run at the $i^{th}$ experimental target condition

Vector of measured values of outputs for the $j^{th}$ replicate run at the $i^{th}$ condition

Output measurement value for the $k^{th}$ model output obtained from the $j^{th}$ replicate run at the $i^{th}$ condition

The average of the $k^{th}$ model output measurements for replicate runs conducted at the $i^{th}$ target condition

Subscripts

$i$ Counter for the unique target experimental conditions

$j$ Counter for the repeated replicate experiments at each condition

$k$ Counter for the elements of a model-output vector in a multi-output model

$n$ Nuisance parameters

$p$ Primary parameters

Greek Symbols

$\delta_k$ Cut-off value for the estimates of variance for the $k^{th}$ output measurement

$\varepsilon$ Random measurement noise for a single-output model with perfectly known inputs

$\varepsilon_{lnf_1,i}$ Random measurement noise at the $i^{th}$ experimental condition for the natural logarithm of the measurements for the comonomer type 1 composition

$\varepsilon_{lnf_1,ij}$ Random measurement noise for the $j^{th}$ pseudo-replicated run at the $i^{th}$ experimental condition for the natural logarithm of the measurements for the comonomer type 1 composition

$\varepsilon_{lnF_1,i}$ Random measurement noise for the natural logarithm of the measurements of the copolymer composition at the $i^{th}$ experimental condition

$\varepsilon_{lnF_1,ij}$ Random measurement noise for the natural logarithm of the $j^{th}$ replicated measurement of the copolymer composition at the $i^{th}$ experimental condition

$\varepsilon_{U,ij}$ Vector of random measurement noises for the input measurements

$\varepsilon_{U,ij}$ Random noise for the input measurements of a model with a single uncertain input

$\varepsilon_{Y,ij}$ Vector of output random measurement noise for the $j^{th}$ replicated run at the $i^{th}$ experimental condition

$\varepsilon_{Y,ij}$ Random measurement noise for the $j^{th}$ output measurement of a single-output model at the $i^{th}$ experimental condition
Random noise for the $k^{th}$ model output measurement obtained for the $j^{th}$ replicated run at the $i^{th}$ target condition

Vector of true values of parameters

Output measurement random noise for a single-output model

Estimate of variance for the $k^{th}$ model output

Variance estimate for the $k^{th}$ model output obtained from measurements of replicate data corresponding to the $i^{th}$ experimental data

Variance for the natural logarithm of the comonomer type 1 composition measurements

Variance for the natural logarithm of the copolymer composition measurements

Variance of the $k^{th}$ output measurements at the $i^{th}$ target condition arising from output measurement noises and input uncertainties

Covariance matrix for the reactivity ratio value estimates

Covariance matrix for the uncertain input

Covariance matrix for the model outputs

Covariance matrix estimate for the primary parameters
Nomenclature for Chapter 4

### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI</td>
<td>Confidence intervals</td>
</tr>
<tr>
<td>EVM</td>
<td>Error-in-variables model</td>
</tr>
<tr>
<td>FIM</td>
<td>Fisher information matrix</td>
</tr>
<tr>
<td>JCR</td>
<td>Joint confidence regions</td>
</tr>
<tr>
<td>MC</td>
<td>Monte Carlo</td>
</tr>
</tbody>
</table>

### Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{C}$</td>
<td>Constant in likelihood function</td>
</tr>
<tr>
<td>$F(N_\theta, \nu_\theta, 0.95)$</td>
<td>The abscissa of a $F$ distribution with $N_\theta$ numerator degrees of freedom, $\nu_\theta$ denominator degrees of freedom, and 95% of the probability to the left</td>
</tr>
<tr>
<td>$g(.)$</td>
<td>Function of model predictions for the outputs</td>
</tr>
<tr>
<td>$g_{aug,s,q}(.)$</td>
<td>Function of model prediction corresponding to the $q^{th}$ element of augmented scaled vector containing all the input and output random variables (i.e., $YU_{aug,s}$)</td>
</tr>
<tr>
<td>$\hat{H}_{EVM}$</td>
<td>Hessian matrix of the $J_{EVM}$ with respect to augmented parameter vector elements, evaluated at $\hat{\theta}_{aug}$</td>
</tr>
<tr>
<td>$\hat{H}_{OLS}$</td>
<td>Hessian matrix of the $J_{OLS}$ with respect to parameter vector elements, evaluated at $\hat{\theta}$</td>
</tr>
<tr>
<td>$J_{EVM}$</td>
<td>EVM-based objective function</td>
</tr>
<tr>
<td>$J_g$</td>
<td>Jacobian matrix of model predictions with respect to the parameters in a model with perfectly-known inputs</td>
</tr>
<tr>
<td>$J_{gaug,s}$</td>
<td>Jacobian matrix of the augmented scaled prediction vector with respect to the augmented parameter vector, evaluated at $\hat{\theta}_{aug}$</td>
</tr>
<tr>
<td>$J_{OLS}$</td>
<td>Ordinary least squared based objective function</td>
</tr>
<tr>
<td>$L(.)$</td>
<td>Likelihood function</td>
</tr>
<tr>
<td>$N$</td>
<td>Number of experimental conditions</td>
</tr>
<tr>
<td>$N(.)$</td>
<td>Normally-distributed random variable</td>
</tr>
<tr>
<td>$N_u$</td>
<td>Number of elements in model input vectors $u_i$</td>
</tr>
<tr>
<td>$N_y$</td>
<td>Number of outputs of the multiresponse model $g(\theta, u_i, x_i)$</td>
</tr>
<tr>
<td>$N_\theta$</td>
<td>Number of parameters</td>
</tr>
<tr>
<td>$t_{\nu_\theta,0.025}$</td>
<td>The abscissa of a Student t-distribution with $\nu_\theta$ degrees of freedom and with 0.975 of the probability to the left</td>
</tr>
<tr>
<td>$U_i$</td>
<td>Vector of random variables corresponding to the measurements of uncertain inputs for the $i^{th}$ experimental run</td>
</tr>
<tr>
<td>$u_i$</td>
<td>Vector of true values of uncertain inputs for the $i^{th}$ experimental run</td>
</tr>
<tr>
<td>$u_i^{(c)}$</td>
<td>The value of sampled model input for the $c^{th}$ MC simulation run</td>
</tr>
<tr>
<td>$\bar{u}_k$</td>
<td>Vector of estimates for uncertain inputs for an arbitrary condition $k$</td>
</tr>
<tr>
<td>$u_{mi}$</td>
<td>Vector of measured values for the uncertain inputs at the $i^{th}$ experimental condition</td>
</tr>
<tr>
<td>$u_{mi}^{(b)}$</td>
<td>Vector of the $b^{th}$ sampled bootstrap data corresponding to the uncertain input measurements at the $i^{th}$ experimental run</td>
</tr>
<tr>
<td>$x_i$</td>
<td>Vector of perfectly known model inputs for the $i^{th}$ experimental run</td>
</tr>
<tr>
<td>$x_k$</td>
<td>Vector of perfectly known model inputs for an arbitrary condition $k$</td>
</tr>
<tr>
<td>$\hat{y}_{k,j}^{(c)}$</td>
<td>The value of $j^{th}$ model output for the $c^{th}$ mean response obtained from MC simulations for an arbitrary experimental condition $k$</td>
</tr>
<tr>
<td>$Y_i$</td>
<td>Vector of random variables corresponding to the output measurements for the $i^{th}$ experimental run</td>
</tr>
</tbody>
</table>
The value of $j^{th}$ model output for the $c^{th}$ future model response measurement obtained from MC simulations for an arbitrary experimental condition $k$

$YU_{aug,s}$ Vector of scaled augmented random variables of all uncertain output and input measurements

$yu_{m.aug,s}$ Vector of scaled augmented measured values of uncertain outputs and inputs

$YU_{aug,s,q}$ The $q^{th}$ element in vector $YU_{aug,s}$

$\hat{y}_k$ Vector of mean predictions for the model outputs corresponding to an arbitrary condition $k$

$y_{m,i}$ Vector of measured values for the model outputs at the $i^{th}$ experimental condition

$y^{(b)}$ Vector of the $b^{th}$ sampled bootstrap data corresponding to the output measurements at the $i^{th}$ experimental run

$\hat{Y}_{sk}$ Vector of model predictions for a single future measurement of model outputs at an arbitrary experimental condition $k$

$z_{0.025}$ The 97.5 percentile point of a standard normally-distributed random variable

Subscripts

- $i$: Counter for experimental conditions ($i = 1, \ldots, N$
- $j$: Counter for the elements of model output vector $Y_i$ ($j = 1, \ldots, N_Y$)
- $k$: Arbitrary condition $k$
- $l$: Counter for the elements of uncertain input vector $u_i$ ($l = 1, \ldots, N_U$)
- $p$: Counter for the elements in parameter vector $\theta$ ($p = 1, \ldots, N_\theta$)

Superscripts

- $b$: Counter for the bootstrap runs
- $b_{max}$: Total number of bootstrap runs
- $c$: Counter for MC simulations
- $c_{max}$: Number of MC simulation runs

Greek symbols

$\varepsilon_{Ui}$ Vector of random variables corresponding to the measurement random noises of model inputs for the $i^{th}$ experimental run

$\varepsilon_{Y_i}$ Vector of random variables corresponding to the measurement noises of model outputs for the $i^{th}$ experimental run

$\varepsilon^{(c)}_{Y_{k,j}}$ The value of $c^{th}$ sampled measurement noise for the $j^{th}$ model output

$\varepsilon_{Yk}$ Vector of random variables corresponding to the output measurement noises of an arbitrary condition $k$

$\varepsilon_{YU_{aug,s,q}}$ Random measurement noise for the $q^{th}$ element of the augmented, scaled measurement vector $YU_{aug,s,q}$

$\theta$: Vector of true values of model parameters

$\theta_{aug}$: Vector of augmented parameter vector

$\hat{\theta}^{(b)}$: Vector of parameter estimates corresponding to the $b^{th}$ bootstrap run

$\hat{\theta}^{(c)}$: Vector of parameter estimates corresponding to the $c^{th}$ MC simulations run

$\nu_\theta$: Degrees of freedom for parameter estimates covariance matrix

$\Sigma_u$: Diagonal covariance matrix for the uncertain input measurements $u_i$ ($i = 1, \ldots, N$

$\Sigma_Y$: Diagonal covariance matrix for the output measurements $Y_i$ ($i = 1, \ldots, N$)

$\Sigma_{\hat{\theta}}$: Covariance matrix for the estimates of parameters
\[ \Sigma_{\hat{\theta}_{\text{aug}}} \] Covariance matrix for the estimates of the augmented parameters vector

\[ \sigma^2_{\epsilon_{y_j}} \] Variance for the random measurement noise of the \( j^{th} \) model output

\[ \sigma_{\epsilon_{U_l}} \] Standard deviation for the random measurement noise of the \( l^{th} \) uncertain input

\[ \sigma_{\epsilon_{y_j}} \] Standard deviation for the random measurement noise of the \( j^{th} \) model output

\[ \tilde{\delta}^2_{y_{k,j}} \] Variance for the mean prediction response of the \( j^{th} \) model output \( (j = 1, \ldots, N_Y) \) for an arbitrary condition \( k \)

\[ \tilde{\delta}^2_{\hat{y}_{sk,j}} \] Variance for a future measurement of the \( j^{th} \) model output obtained at an arbitrary condition \( k \)

\[ \tilde{\delta}_{\hat{\theta}_p} \] Standard deviation of the \( p^{th} \) estimated parameters

\[ \chi^2(N_\theta, 0.95) \] the abscissa of a chi-squared distribution with \( N_\theta \) degrees of freedom and 95% of the probability to the left
### Nomenclature for Chapter 5

#### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2MP</td>
<td>2-methyl pentane</td>
</tr>
<tr>
<td>3MP</td>
<td>3-methyl pentane</td>
</tr>
<tr>
<td>C₃</td>
<td>Propane</td>
</tr>
<tr>
<td>EVM</td>
<td>Error-in-variables model</td>
</tr>
<tr>
<td>H₂</td>
<td>Hydrogen</td>
</tr>
<tr>
<td>MC</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>nC₆</td>
<td>n-Hexane</td>
</tr>
<tr>
<td>PR</td>
<td>Pseudo replicate</td>
</tr>
<tr>
<td>TR</td>
<td>True replicate</td>
</tr>
<tr>
<td>WLS</td>
<td>Weighted least squares</td>
</tr>
</tbody>
</table>

#### Symbols

- \( b \): Counter for the iterations of the bootstrapping algorithm
- \( b_{\text{max}} \): Maximum number of the bootstrap and Monte Carlo iterations
- \( c \): A condition of interest
- \( e \): The relative change of parameter estimates in two subsequent iterations of the algorithm used in PR situations to estimate output measurement variances
- \( E_{\text{a}r} \): Activation energy for the \( r^{th} \) reaction
- \( F_{2\text{MP}} \): Molar outflow rate of 2MP
- \( F_{3\text{MP}} \): Molar outflow rate of 3MP
- \( F_{C_3} \): Molar outflow rate of C₃
- \( F_{nC_6} \): Feed flow rate of nC₆
- \( F_{H_2} \): Feed flow rate of H₂
- \( F_{H_2}^{0} \): Measured value of feed flow rate of hydrogen
- \( F_{nC_6}^{0} \): Measured value of feed flow rate of n-hexane
- \( g(.) \): The function returning a vector of solutions for the model predictions of a multi-output model
- \( g_{k}(.) \): The function returning solution for the prediction of the \( k^{th} \) output
- \( J \): Objective function
- \( k_{r} \): Kinetic rate for the \( r^{th} \) reaction
- \( K_{\text{ref}} \): Equilibrium constant at the reference temperature \( T_{\text{ref}} \)
- \( k_{\text{ref}r} \): Kinetic rate of the \( r^{th} \) reaction at the reference temperature \( T_{\text{ref}} \)
- \( N \): Number of target experimental condition
- \( N_U \): Number of uncertain model input
- \( N_Y \): Number of output predictions at each experimental condition
- \( N_\theta \): Number of parameters
- \( n_i \): Number of replicate experiments at the \( i^{th} \) run condition
- \( P \): Reactor pressure
- \( P_i \): Reactor pressure at the \( i^{th} \) condition
- \( p \): Counter for the elements of parameter vector \( \theta \)
\( s \) Iteration counter in calculating the output measurement variance using pseudo-replicate data

\( S_q \) Selectivity of the \( q^{th} \) species (\( q = 2MP, 3MP \) and \( C_3 \))

\( T \) Reactor temperature

\( T_i \) Reactor temperature at the \( i^{th} \) experimental condition

\( T_{ref} \) Reference temperature

\( u_{ij} \) True value of the input for \( j^{th} \) replicate run of the \( i^{th} \) target condition

\( u_{tl} \) True value for the \( l^{th} \) uncertain input at the \( i^{th} \) experimental condition when there are no replicate data

\( u_{ij} \) Vector of true values of uncertain inputs for the \( j^{th} \) pseudo-replicate run at the \( i^{th} \) target condition

\( u_{ij}^{(b)} \) Bootstrap-generated value of the uncertain input for the \( j^{th} \) pseudo-replicate run at the \( i^{th} \) target condition

\( u_{mC_6} \) Measured value of \( F_{nc6}^0 \) at the \( i^{th} \) experimental condition

\( u_{mH_2} \) Measured value of \( F_{H_2}^0 \) at the \( i^{th} \) experimental condition

\( u_{ml} \) Vector of measured values for uncertain inputs obtained from the \( i^{th} \) experimental condition

\( u_{m ll} \) Measured value for the \( l^{th} \) uncertain input at the \( i^{th} \) experimental condition

\( W \) Catalyst weight

\( x_i \) Vector of perfectly-known model inputs at the \( i^{th} \) target condition

\( y_{2MP} \) Measured value of \( F_{2MP} \) for the \( j^{th} \) pseudo-replicate run at the \( i^{th} \) target condition

\( y_{3MP} \) Measured value of \( F_{3MP} \) for the \( j^{th} \) pseudo-replicate run at the \( i^{th} \) target condition

\( y_{C_3} \) Measured value of \( F_{C_3} \) for the \( j^{th} \) pseudo-replicate run at the \( i^{th} \) target condition

\( Y_{ij} \) Output measurement random variable corresponding to a single-output model for the \( j^{th} \) replicate run at the \( i^{th} \) experimental target condition

\( Y_{ijk} \) Output measurement corresponding for the \( k^{th} \) output of a multi-output model obtained for the \( j^{th} \) replicate run at the \( i^{th} \) experimental target condition

\( y_{m} \) Vector of measured values of outputs for the \( j^{th} \) pseudo-replicate run at the \( i^{th} \) condition

\( y_{m}^{(b)} \) Vector of bootstrap-generated output measurements corresponding to \( y_{mij} \)

\( y_{m k} \) Measured value for the \( k^{th} \) model response at the \( i^{th} \) experimental condition when there not replicates

\( y_{m ij} \) Output measurement value for the \( k^{th} \) model output obtained from the \( j^{th} \) pseudo-replicate run at the \( i^{th} \) condition

\( \bar{y}_{m k} \) The average of the \( k^{th} \) model output measurements for replicate runs conducted at the \( i^{th} \) target condition

\( \bar{y}_{2MP} \) The average of the measured values of \( F_{2MP} \) for replicate runs at the \( i^{th} \) target condition
Subscripts
\[ b \]
Counter for the iterations of the bootstrapping algorithm
\[ b_{max} \]
Maximum number of the bootstrap and Monte Carlo iterations
\[ c \]
A condition of interest
\[ i \]
Counter for the unique target experimental conditions
\[ j \]
Counter for the repeated replicate experiments at each condition
\[ k \]
Counter for the elements of a model-output vector in a multi-output model
\[ l \]
Counter for the elements of a model-input vector
\[ r \]
Counter for the reactions
\[ n \]
Nuisance parameters
\[ p \]
Primary parameters
\[ q \]
Letter for denoting different species \((q = 2\text{MP}, 3\text{MP}, \text{and} C_3)\)

Greek Symbols
\[ \alpha_q \]
The number of carbon atoms in the molecules of species \(q\) \((q = 2\text{MP}, 3\text{MP} \text{and} C_3)\).
\[ \delta_k \]
Cut-off value for the estimates of variance for the \(k^{th}\) model response
\[ \delta_{F_{2\text{MP}}} \]
Cut-off value for the estimates of variance for the output measurements of \(F_{2\text{MP}}\)
\[ \delta_{F_{3\text{MP}}} \]
Cut-off value for the estimates of variance for the output measurements of \(F_{3\text{MP}}\)
\[ \delta_{F_{C_3}} \]
Cut-off value for the estimates of variance for the output measurements of \(F_{C_3}\)
\[ \Delta H^0 \]
Enthalpy of the physisorption
\[ \varepsilon_{Uij} \]
Vector of random measurement noises for the input measurements
\[ \varepsilon_{Yij} \]
Vector of output random measurement noise for the \(j^{th}\) replicated run at the \(i^{th}\) experimental condition
\[ \theta \]
Vector of parameters
\[ \hat{\theta}^{(b)} \]
Parameter estimates
\[ \hat{\theta} \]
Vector of estimated values of parameters from the \(b^{th}\) bootstrap-generated data
\[ \hat{\sigma}^2_{F_{2\text{MP}}} \]
Output measurement variance for \(F_{2\text{MP}}\)
\[ \hat{\sigma}^2_{F_{3\text{MP}}} \]
Output measurement variance for \(F_{3\text{MP}}\)
\[ \hat{\sigma}^2_{F_{C_3}} \]
Output measurement variance for \(F_{C_3}\)
\[ \sigma^2_{n_{C_6}} \]
Measurement variance for uncertain input \(F^0_{n_{C_6}}\)
\[ \sigma^2_{o_{H_2}} \]
Measurement variance for uncertain input \(F^0_{o_{H_2}}\)
\[ \hat{\sigma}^2_{F_{2\text{MP}}} \]
Overall estimate of measurement variance for \(F_{2\text{MP}}\) accounting for input measurement noises \(\varepsilon_{Uij}\) and output measurement noises \(\varepsilon_{Yij}\)
\[ \sigma^2_{U_l} \]
Measurement variance for the \(l^{th}\) uncertain input
\[ \sigma^2_{Y_k} \]
Output measurement variance for the \(k^{th}\) model output
\[ \sigma^2_{Y_k} \]
Estimate of variance for the \(k^{th}\) model output measurements
\[ \sigma^2_{Y_{lk}} \]
Variance estimate for the \(k^{th}\) model output obtained from measurements of replicate data corresponding to the \(l^{th}\) experimental data
\[ \Sigma_U \]
Covariance matrix for the uncertain inputs
\[ \Sigma_Y \]
Covariance matrix for the model outputs

xxxiv
$\Sigma^{(b)}_y$ Estimated value of output measurements covariance matrix estimated from the $b^{th}$ bootstrap-generated data
Chapter 1

Introduction

Fundamental models are important for developing and improving chemical processes. These models are usually developed based on chemical engineering principles such as conservation of mass and energy, along with constitutive equations describing reaction kinetics and thermodynamic behaviour. Derivation of fundamental model equations is often followed by a parameter estimation step to determine appropriate values for the unknown parameters. This thesis is concerned with developing fundamental model equations and developing statistical methods for estimating fundamental model parameters. The research in the thesis focuses on two different situations, partly because it was sponsored by two different organizations. The research for the first part of the thesis which was sponsored by Technisch Industriele Procesbesturing (TIPb) and involves data from Canon Production Printing, involves the development of a new fundamental model for photopolymerization of 1,6-hexanediol diacrylate (HDDA) monomer used in high-speed printing applications. The resulting fundamental model contains too many parameters to be estimated reliably using the available conversion vs. time data provided by my industrial sponsor. In this situation, I focus on determining which types of reactions and associated parameters should be included in the model so that the predictions can fit the available data without overfitting the parameters. During this modeling and parameter estimation study, I rely on traditional least-squares assumptions: i) the independent variables are perfectly known, ii) the model equations have the correct structure, iii) the random errors in the dependent variables are independent, identically distributed and normally distributed.¹

The research in the second part of the thesis, which is sponsored in part by EUROKIN, involves new statistical methods for fundamental models where there is enough information to estimate all of the parameters. The complicating issue considered is that the independent variables
(experimental inputs) are not always perfectly known and may contain significant error when compared to the measurement errors in the dependent variables. Two different case studies involving experimental data from the literature are considered: i) copolymerization of n-butyl methacrylate and n-butyl acrylate; ii) hydroisomerization of n-hexane.

1.1 Detailed Thesis Objectives

1.1.1 Mathematical Modeling and Parameter Estimation for HDDA Photopolymerization with Bifunctional Initiator

Mathematical modeling of HDDA photopolymerization with bis-acylphosphine oxide (BAPO) is important because both the monomer and initiator are used in high-speed printing applications. A better understanding of the underlying reaction kinetics, developed through mathematical modeling, can help companies to systematically design their inks and printers. In this thesis, a comprehensive kinetic model is proposed to account for a wide variety of chemical phenomena. The importance of the proposed phenomena is evaluated using a parameter estimability analysis based on parameter ranking and parameter subset selection. The corresponding model for this photopolymerization is complicated because the monomer and initiator molecules are bifunctional. Further, the monomer is an acrylate, so macroradicals that are generated may undergo backbiting reactions. As such, a variety of different radical types and chemical phenomena might be important for predicting conversion vs. time behaviour for this photopolymerization process (e.g., crosslinking, backbiting, cyclization, and diffusion-dependent reactions). In Chapter 2 of the thesis my research goals are to:

i. Conduct a detailed literature review to identify important side reactions that may occur during HDDA polymerization;

ii. Develop a mathematical model to account for these phenomena;

iii. Use a systematic parameter subset selection and parameter estimation method to determine which parameters and phenomena should be included in the model;
iv. Compare model predictions with validation data not used for parameter estimation and make recommendations for future models that will account for oxygen contamination and spatial variations within the polymer film.

1.1.2 Output Measurement Variance Estimation, Parameter Estimation and Uncertainty Quantification in EVM situations

Error-in-variables models (EVMs) are used for parameter estimation in situations where some of the model inputs (independent variables) are not perfectly known.9,10 The corresponding EVM objective functions are more complicated than conventional parameter estimation objective functions such as weighted-least squares (WLS). For example, uncertain inputs are included along with the model parameters as extra decision variables in the parameter estimation optimization problem. Also, EVM objective functions have additional terms that account for input uncertainties. Information about input measurement variances and output measurement variances is needed to calculate appropriate weighting factors in the EVM objective function. Values of input measurement variances are usually assumed known based on engineering knowledge about uncertainties in analytical devices and in methods that are used to achieve target input values. Until now, most authors assume that output measurement variances are also known. Some research groups recommend estimation of output variances based on replicate experiments. However, until now, methods for estimating output variances from pseudo-replicate experiments that are influenced by input uncertainties have not been available. In Chapter 3 of the thesis my goals are to:

i. Conduct a detailed literature review on the types of replicate experiments that have been conducted by researchers. I wanted to determine whether true replicates were performed wherein the same unknown true value of an input was used in multiple experiments or whether pseudo replicates were conducted using different unknown true values. I also
wanted to investigate how users of EVM had estimated their output measurement variances.

\textit{ii.} Propose a linearization-based methodology for quantifying output measurement variances based on pseudo-replicate experiments.

\textit{iii.} Illustrate the methodology using a reactivity-ratio estimation problem involving literature data for n-butyl methacrylate/n-butyl acrylate copolymerization.

When error-in-variables models are used for making predictions, uncertainties in inputs and uncertainties in parameters both influence the model prediction uncertainties. In Chapter 4, I provide a detailed literature review about literature studies where uncertainties in model predictions were quantified. This literature review identified the following gaps: i) it was unclear when linearization-based and Monte Carlo methods should be used for uncertainty quantification; ii) researchers were not careful to distinguish between uncertainties in mean predictions at a particular target operating point and uncertainties associated with a new experimental result; iii) simplifying assumptions were made about output variances and parameter uncertainties that may not be valid in EVM estimation problems. As such, the main objectives for the research in Chapter 4 which is a literature review are to:

\textit{i.} Provide a summary of the chemical engineering literature where parameter uncertainties and prediction uncertainties were quantified in error-in-variables models.

\textit{ii.} Provide detailed recommendations on how modelers can use different techniques (i.e., linearization-based and Monte Carlo) to quantify parameters and prediction uncertainties.

\textit{iii.} Provide detailed information for modelers who want to quantify uncertainties in mean behaviour or uncertainties associated with predictions of a new experiment.

In Chapter 5, a multi-response case study is used to illustrate EVM parameter estimation and uncertainty quantification for model parameters and model predictions. This case study relies on a multi-response model for hexane hydroisomerization and associated data collected at
University of Gent. Because replicate experiments were performed on different days, I classified the repeated experiments as pseudo replicates, which might have different true values for uncertain hydrogen and hexane feed rates. The objectives for Chapter 5 are to:

1. Use the linearization-based method proposed in Chapter 3 to estimate output measurement variances from the pseudo-replicate data and then estimate parameters in the hexane hydroisomerization model.
2. Quantify parameter uncertainties when different levels of input uncertainties were assumed. Compare the resulting uncertainties with those obtained using WLS estimation, where input uncertainties were neglected.
3. Obtain two types of prediction uncertainties (i.e., uncertainties in mean response and in a single future experiment) for a variety of model outputs (i.e., reactor outflow rates, selectivities and yield).

Achieving objectives 1. and 2. required extension of linearization and bootstrapping methods that were developed in Chapter 3 and identified in the literature review in Chapter 4, so that they can be used for multivariate situations involving multiple uncertain inputs and multiple model responses.

1.2 Thesis Outline

The next four Chapters in this thesis are manuscripts that have been accepted in peer-reviewed journals or will be submitted soon. In Chapter 2, a comprehensive model is proposed to account for photopolymerization of HDDA with BAPO as initiator. Chapter 2 has been published in Chemical Engineering Science. In Chapter 2, first, a reaction scheme and corresponding model equations are proposed. Next, a systematic parameter subset selection and estimation study is used to determine important phenomena during photopolymerization, thereby enabling model simplification. Finally, the quality of model predictions is validated using experimental data that were not included in the parameter estimation. In Chapter 3, a detailed literature review is provided
about previous EVM studies in the Chemical Engineering literature with a focus on the types of replicate experiments that were attempted. Then, the attempted replicate experiments are classified as either true replicates or pseudo replicates. A linearization-based technique is proposed to estimate output measurement variances from pseudo-replicate data. Next, a single-response copolymerization model and literature data are used to illustrate EVM parameter estimation, measurement variance estimation, and parameter uncertainty quantification. The manuscript in Chapter 3 was published in *AIChe Journal*.

In Chapter 4, a literature review and detailed recommendations are provided for modelers who want to quantify parameter uncertainties and prediction uncertainties while accounting for input uncertainties using error-in-variable models. The manuscript in Chapter 4 have been submitted to *Industrial & Engineering Chemistry Research*.

In Chapter 5, parameter uncertainties and prediction uncertainties are obtained for a steady-state model describing n-hexane hydroisomerization in a well-mixed reactor. The available data set contains pseudo-replicate experiments. The results of the parameter estimation and uncertainty quantification are compared with corresponding results from conventional WLS parameter estimation where input uncertainties are neglected. Interesting trends in the prediction uncertainties are uncovered, due to the tradeoff between the assumed magnitudes of the input uncertainties and the resulting magnitudes of the output uncertainty estimates. The manuscript in Chapter 5 have been submitted to *AIChe* journal. In Chapter 6, conclusions and recommendations for this PhD thesis are provided.
1.3 References for Chapter 1


Chapter 2

Mathematical Modeling and Parameter Estimation for 1,6-Hexanediol Diacrylate Photopolymerization with Bifunctional Initiator*

2.1 Abstract

A dynamic model is proposed for photopolymerization of 1,6-hexanediol diacrylate (HDDA) using bifunctional initiator bis-acylphosphine oxide (BAPO). The proposed model accounts for branching, backbiting and cyclization reactions, and for diffusion-dependent reaction rates during photopolymerization. The proposed model contains 40 adjustable kinetic and free-volume parameters. Experimental data available for parameter estimation are vinyl group conversions obtained using a variety of light intensities and exposure times, and monomer conversions for three experiments. Systematic parameter ranking and estimation is used to evaluate the influence of phenomena included in the model on the quality of the fit. Estimation and ranking results indicate that branching, backbiting, and cyclization reactions have important influences on conversion. Reactions involving two large molecules and propagation reactions become diffusion-dependent. Incorporating diffusion-dependent initiator efficiency results in improved model predictions.

2.2 Introduction

Free-radical photopolymerization of multifunctional monomers is used for production of coatings, adhesives and printer inks.\textsuperscript{1,2} The resulting products are highly cross-linked, rigid, and glassy polymers with high modulus and solvent resistance.\textsuperscript{1} Photopolymerization processes have several advantages compared to conventional thermally-induced polymerization.\textsuperscript{1,3-5} For example, polymerization proceeds quickly, even at ambient temperature. Acrylate monomers are often used in photopolymerization applications such as printing because of their high reactivity.\textsuperscript{5} In this chapter, a comprehensive mathematical model is developed to describe photopolymerization inside a thin film containing the bifunctional monomer 1,6-hexanediol diacrylate (HDDA) and the bifunctional initiator bis-acylphosphine oxide (BAPO). BAPO, has two weak phosphorous-carbon bonds in its structure that break apart to produce free-radicals, which initiate polymerization.\textsuperscript{6} The proposed model tracks the concentration of unreacted initiator molecules and the concentration of initiator fragments that contain weak phosphorous-carbon bonds that may subsequently decompose to produce additional radicals. HDDA is a monomer of interest which is similar to monomers used in high-speed printing of posters.

The objective of the current research is to develop a model to predict the conversion-time behaviour of vinyl groups and monomer obtained using different light intensities and duration of the light pulse. The proposed model builds on previous HDDA photopolymerization models that considered the influence of light intensity and oxygen\textsuperscript{7} using a simplified reaction scheme involving a mono-functional initiator. In this simplified scheme, vinyl groups on the monomer and the polymer were assumed to have the same reactivities, and backbiting and cyclization reactions were ignored. In the current chapter, a comprehensive kinetic model is proposed and used to explore the relative importance of different types of reactions when predicting conversion vs. time behaviour.
Several research groups have proposed photopolymerization models involving acrylate and methacrylate monomers. A summary of these models is provided in Table 2-1. Twenty of the 23 studies in Table 2-1 involved monomers with two vinyl groups, which make crosslinking and cyclization reactions possible. However, only five of these modeling studies considered crosslinking reactions. Two additional studies account for the influence crosslinking reactions by considering that some free radicals may become “trapped” due to the formation of 3-D networks arising from crosslinking.\textsuperscript{1,8} Unfortunately, these models do not track pendant vinyl groups attached to the polymer, which are consumed during crosslinking reactions.\textsuperscript{1,8}

Cyclization reactions were only considered in 3 of 20 models involving diacrylates and dimethacrylates. In these studies, cyclization and cross-linking reactions are both assumed to be second-order reactions that may occur when a radical on a chain end is sufficiently close to a pendant double bond, which may be located on the same polymer molecule or on a nearby molecule. Backbiting was only considered in 1 study out of the 13 studies involving acrylate photopolymerization, even though backbiting has been demonstrated to have significant influence on the rate of acrylate polymerizations, even at low temperatures.\textsuperscript{9} In the proposed model, crosslinking, cyclization and backbiting reactions are included in the reaction scheme.

Fifteen of 23 studies in Table 2-1 account for the gel effect, which arises when termination reactions become diffusion-dependent. As polymerization proceeds, the viscosity of the reaction medium increases and the mobility of the reactants, especially large macromolecules, decreases. As a result, the rates of some reactions decrease with increasing conversion, especially the rates for termination reactions that involve two macromolecules. The 15 studies that accounted for gel-effect considered the influence of diffusion-dependent termination reactions and also the effect of diffusion-dependent propagation reactions. In 3 of the 15 studies, it was assumed that initiator efficiency was also influenced by the gel effect, with initiator efficiency decreasing as conversion increased due to reduced mobility of initiator radicals.\textsuperscript{1,8,10}
In the current chapter, we assess whether some of the above-mentioned phenomena (i.e., cyclization, backbiting, and diffusion-dependent reaction rates for small molecules) are important during HDDA photopolymerization. To do so, four different cases are considered, and parameter estimates are obtained for the corresponding models. A summary of the cases and the corresponding phenomena which are considered in each case is provided in Table 2-2. In Case I, only the crosslinking reactions are considered in addition to the main reactions that are conventionally considered in free-radical polymerization models (i.e., initiation, propagation and termination reactions). In Case I, the gel-effect is only considered for reactions involving two large molecules (i.e., crosslinking and termination reactions). In Case II, backbiting reactions are added to the simplified model in Case I. In Case III, cyclization is also considered. Finally, in Case IV, first-order reactions (i.e., cyclization and backbiting) and the reactions involving small molecules (i.e., initiation reactions, propagation reactions and termination reactions involving small molecules) are also considered diffusion-dependent, along with the initiator efficiency. In all four cases, the corresponding parameters for each case are estimated. Predictions from the different cases are compared to assess which phenomena are important for achieving a good fit to the data.

The remainder of this chapter is organized as follows. First, we describe the experimental data for HDDA polymerization in absence of oxygen (under nitrogen atmosphere) that are available to support parameter estimation and model validation. Next, the reaction mechanism for the proposed model is presented, along with the corresponding dynamic material balance equations for monomer, pendant vinyl groups, initiator, initiator fragments, and the different types of free radicals that are present in the system. The parameter-estimation process is described and simulation results are presented. Model predictions are compared with the data used for parameter estimation and with additional validation data sets so that the predictive ability of the proposed model can be assessed. Because a variety of phenomena (e.g., crosslinking, backbiting, cyclization, and gel effect) are included in the model, a large number of parameters (i.e., 40) appear in the model
equations. A systematic parameter ranking and estimation approach is used to determine which parameters are estimable from the available data and whether specific phenomena have noticeable effects on the fit to the conversion vs. time data.

Table 2-1 A summary of modeling studies for photopolymerization of acrylate and methacrylate monomers

<table>
<thead>
<tr>
<th>Authors</th>
<th>Monomers</th>
<th>Phenomena accounted for</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Crosslinking</td>
</tr>
<tr>
<td>Kurdikar and Peppas 11</td>
<td>Diethyleneglycol diacrylate</td>
<td>N</td>
</tr>
<tr>
<td>Lecamp et al. 12</td>
<td>Dimethacrylate oligomer</td>
<td>N</td>
</tr>
<tr>
<td>Elliott and Bowman 13</td>
<td>Diethyleneglycol dimethacrylate</td>
<td>Y</td>
</tr>
<tr>
<td>Wenand and McCormick 1</td>
<td>Diethyleneglycol dimethacrylate</td>
<td>N</td>
</tr>
<tr>
<td>Elliott et al. 14</td>
<td>Three different comonomer compositions involving methacrylates and dimethyl acrylates</td>
<td>Y</td>
</tr>
<tr>
<td>Elliott et al. 15</td>
<td>bis-GMA/triethyleneglycol dimethacrylate</td>
<td>Y</td>
</tr>
<tr>
<td>Boodhoo et al. 16</td>
<td>n-Butyl acrylate</td>
<td>N.A.</td>
</tr>
<tr>
<td>Achilias and Sideridou 8</td>
<td>Triethylene glycol dimethacrylate</td>
<td>N</td>
</tr>
<tr>
<td>Dickey et al. 17</td>
<td>Ethylene glycol diacrylate and t-butyl acrylate</td>
<td>N</td>
</tr>
<tr>
<td>Kizilel et al. 18,19</td>
<td>Ethylene glycol diacrylate with vinyl pyrroldinone comonomer</td>
<td>Y</td>
</tr>
<tr>
<td>Neo and Chan-Park 20</td>
<td>Polyethylene glycol diacrylate and trimethylpropane triacrylate</td>
<td>N</td>
</tr>
<tr>
<td>Johnson et al. 21,22</td>
<td>Hexane diol diacrylate with monoacrylate comonomers</td>
<td>N</td>
</tr>
<tr>
<td>Jachuck and Nekkanti 23</td>
<td>n-butyl acrylate</td>
<td>N.A.</td>
</tr>
<tr>
<td>Vinu and Madras 24</td>
<td>Alkyl methacrylates</td>
<td>N.A.</td>
</tr>
<tr>
<td>Boddapati et al. 25</td>
<td>Trimethylolpropane triacrylate</td>
<td>N</td>
</tr>
<tr>
<td>Study</td>
<td>Monomer</td>
<td>Crosslinking</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Taki et al. 26,27</td>
<td>Diurethane dimethacrylate</td>
<td>N</td>
</tr>
<tr>
<td>Iedema et al. 7,28</td>
<td>Hexanediol diacrylate</td>
<td>N</td>
</tr>
<tr>
<td>Anastasio 10</td>
<td>Bisphenol-A-ethoxylated dimethacrylate (SR540)</td>
<td>N</td>
</tr>
<tr>
<td>Christmann et al. 29</td>
<td>Ethoxylated bisphenol A diacrylate (SR349)</td>
<td>N</td>
</tr>
</tbody>
</table>

Table 2-2 Kinetic phenomena considered during four parameter-estimation scenarios
2.3 Experimental Data

Two types of experimental data (i.e., measured vinyl-group conversion and monomer conversion) were collected using FT-IR spectroscopy and 1H-NMR spectroscopy, respectively. Real-time vinyl-group conversion data were collected during polymerization of a 12-μm thick film. At the onset of polymerization this film contains nearly pure 1,6-hexanediol diacrylate (HDDA) monomer (Sigma Aldrich) with added bis(acylphosphine oxide) (BAPO) initiator (BASF, 4 wt%) and Tegorad 2250 surfactant (Evonik, 2 wt%). The photopolymerization was started by exposure of the film to light from one of two lamps (provided by a FireFly FF200 50x20 light source from Phoseon Technology; one with a peak wavelength at 405 nm and the other with a peak wavelength at 410 nm). The data were collected using two different types of experiments: i) where the film is exposed to 410 nm light throughout the entire duration of the polymerization experiment, and ii) where the film is exposed to 405 nm light for a short time at the start of the polymerization, after which the light exposure is stopped (i.e., light is added using a pulsed signal). The light intensity corresponding to the constant-light-exposure experiments was set at values between 105 and 1350 W m⁻². For the constant-exposure experiments, the UV light illuminated a 12 μm layer of HDDA solution placed on a Germanium crystal which was placed in a Horizontal QuickLock Transmittance Accessory (Bruker). For the short-time-exposure experiments, light with an intensity of 20 W/m² illuminated a 12 μm solution layer placed on a horizontal ZnSe-ATR crystal with angle 45° (see Boonen et al.³⁰ for details about this setup).

Vinyl-group conversion vs. time data were collected using a Bruker Vertex 80 FT-IR spectrometer with rapid scan option at a frequency of ten spectra per second and a spectral resolution of 4 cm⁻¹ (scan velocity = 240 kHz for constant-exposure experiments and 320 kHz for short-term exposure experiments).

Examples of the resulting vinyl conversion vs. time data are provided in Figure 2-1a, which shows data from 3 out of the 8 runs that were conducted using a constant light intensity during each
run. Figure 2-1b shows all of the available vinyl-group conversion data from the short-time-exposure experiments. Nine short-time-exposure experiments were conducted using three different pulse durations (i.e., three replicates of three runs). No replicate data are available for the constant-light-exposure runs. Details about the computation of vinyl-group conversions from the real-time FTIR measurements are provided elsewhere.\(^{30}\)

We determined the monomer conversion for the short-term-exposure experiments using 1H-NMR. The film was extracted in deuterated chloroform with an internal standard for two hours. The extract was analyzed on a Bruker 600 MHz NMR system equipped with a DCH-cryoprobe, using 1H-NMR with 64 scans and a delay time of 30 seconds. No monomer conversion data are available for constant-exposure experiments.

The extinction coefficient \(\epsilon\) for BAPO was determined by recording the light absorption spectrum of BAPO. The corresponding extinction coefficients at 405 nm and 410 nm are 135.7 m\(^3\)mol\(^{-1}\)m\(^{-1}\) and 104.8 m\(^3\)mol\(^{-1}\)m\(^{-1}\), respectively.

![Figure 2-1](image)

**Figure 2-1** Experimental data corresponding to a) constant-light-exposure experiments with the wavelength of 410 nm and different light intensities and b) short-time exposure photopolymerization with wavelength of 405 nm, intensity of 20 W/m\(^2\) and a variety of exposure times. In part b), data from three replicate runs are shown for each exposure time.
2.4 Model Development

2.4.1 Reaction Scheme

A list of all the reactions involved in the proposed model is provided in Table 2-3. As indicated in Table 2-3 and Figure 2-2a, the photopolymerization starts by decomposition of the initiator due to UV irradiation. Figure 2-2a shows the two phosphorous-carbon bonds in BAPO in red. When one of these weak bonds is broken, a carbon-centered radical \( I \) and a phosphorous-centered radical \( \tilde{I} \) are formed via reaction R1 in Table 2-3.\(^6\) This reaction is shown with an efficiency factor \( f \) to account for the cage effect.\(^{31,32}\) The phosphorous-centered initiator radical \( \tilde{I} \) has a remaining weak carbon-phosphorous bond that can decompose in a later reaction (see reaction R2 in Table 2-3, which has efficiency \( f \tilde{I} \), and the scheme in Figure 2-2b).\(^6\) It is assumed (assumption 2-4.5 in Table 2-4) that the radical on \( \tilde{I} \) is consumed via an initiation reaction (reaction R4 or R7) or a termination reaction (reactions R23 to R25) prior to decomposition of the remaining weak bond. After the radical on \( \tilde{I} \) is consumed, the weak bond available for decomposition is tracked using the symbol \( \tilde{I} \). For example, initiation reaction R4 is shown in Figure 2-3 with the portion of the product with the weak red bond labelled as \( \tilde{I} \). It is assumed that the phosphorous-centered radical \( \tilde{I}_p \) generated by decomposition of \( \tilde{I} \) in reaction R2 has the same reactivity as the phosphorous-centred radical \( \tilde{I}^* \) generated from decomposition of the initiator \( I \) in reaction R1 (assumption 2-4.6).
Initiation reaction involving a phosphorous-centered initiator radical $\tilde{I}$ and monomer, resulting in a growing polymer chain containing a $\tilde{I}$ end group (within the ellipse) that may later decompose.

Reactions R3 to R8 in Table 2-3 are initiation reactions which consume either a monomer $M$ or a pendant vinyl group $V_p$. The product of initiation reaction R3 is a regular radical end $R^{=\bullet}$. This type of end group (see Figure 2-4a) is produced whenever a monomer reacts with a free radical (i.e., via reactions R3 to R5 and R9 to R12). In these reactions, one of the two vinyl groups on the monomer is consumed and the other remains unreacted. For simplicity, the unreacted double bond
and the radical are tracked together as an end group denoted by \( R^{=\cdot} \) (assumption 2-4.7 in Table 2-4). \( R^{=\cdot} \) end groups may undergo first-order cyclization (reaction R17) where the radical in \( R^{=\cdot} \) reacts with its vinyl group to produce a cyclized end group \( C^{\cdot} \) (see Figure 2-4b). Consumption of the vinyl group in \( R^{=\cdot} \) by other reactions is ignored because the lifetime of each \( R^{=\cdot} \) end group is very short, with the free radical reacting much more quickly than the corresponding vinyl group. As indicated in assumption 2-4.8, the cyclization reaction involving \( R^{=\cdot} \) and pendant vinyl groups that are further from the chain end are ignored. The proposed reaction scheme converts the vinyl group associated with \( R^{=\cdot} \) into a regular pendant vinyl group \( V_p \) when \( R^{=\cdot} \) is consumed (i.e., via propagation reaction R9, branching reaction R13, backbiting reaction R18 and termination reactions, R29 and R33 to R35). Figure 2-5 shows propagation reaction R9, which consumes an \( R^{=\cdot} \) end group, producing a pendant vinyl group \( V_p \) and generating a new \( R^{=\cdot} \) end group.
Figure 2-4 Formation of different macroradicals including: a) regular radical ends $R^{=\cdot}$, b) cyclized radical ends $C^\cdot$, c) branch-point radicals $B^\cdot$, and d) tertiary radicals $T^\cdot$
Figure 2-5 The vinyl group that is part of a functional group $R^{=\cdot}$ (in the ellipse on the left-hand side) is treated as a pendant vinyl group $V_p$ (in the circle on the right-hand side) when $R^{=\cdot}$ is consumed by reaction.

Reactions R6 to R8 are initiation reactions in which one of the initiator radicals ($I^{\cdot}$, $I^{\cdot}$ or $I^{\cdot}_P$) reacts with a pendant vinyl $V_p$, resulting in a branch-point radical $B^{\cdot}$. As shown in Figure 2-4c, a branch-point radical $B^{\cdot}$ is produced whenever a radical of any kind reacts with $V_p$ (see reactions R6 to R8 and R13 to R16). Reactions R9 to R12 are propagation reactions involving different radical end groups. As seen in Table 2-3, reaction R12 can also be treated as a branching or crosslinking reaction, because the resulting product is a T-shaped branch-point macroradical. Branch-point macroradicals from the other crosslinking reactions (R13 to R16) lead to H-shaped branching. Tertiary radicals of type $T^{\cdot}$ may be generated by backbiting reactions R18 and R19 (see assumption 2-4.9). Reaction R18 is shown schematically in Figure 2-4d. In this reaction, a hydrogen atom is abstracted from a tertiary carbon after formation of a six-member intermediate ring structure.$^9$,$^{33,34}$
Reactions R20 to R38 are bimolecular termination reactions. In this reaction scheme, cross-termination reactions between initiator radicals (i.e., $I^*$, $\tilde{I}^*$, and $I'^{*}$) are ignored (assumption 2-4.10). Similarly, termination reactions between initiator radicals and $R^{{\equiv}^{*}}$ are ignored (assumption 2-4.11). Neglecting these relatively unimportant termination reactions permits the use of the stationary state hypothesis (SSH) which helps to reduce stiffness when solving the differential equations that arise from the reaction scheme in Table 2-3. Reaction R29 is a termination reaction involving two secondary radicals $R^{{\equiv}^{*}}$, resulting in two pendant vinyl groups $V_P$. These pendant vinyl groups were associated with the corresponding $R^{{\equiv}^{*}}$ groups before they were consumed. Note that reaction R29 is a termination-by-combination reaction, because secondary radicals in acrylate polymerizations tend to terminate by combination.\(^9\) In the proposed model, it is assumed that all termination reactions involving secondary macroradicals are combination reactions (see assumption 2-4.12). Reaction R32 is a termination-by-disproportionation reaction involving two tertiary radicals $T^*$, which generates a vinyl group at the end of one of the polymer chains. This terminal vinyl group is treated as if it were a regular pendant vinyl group $V_P$ in the proposed model (see assumption 2-4.14 in Table 2-4). It is assumed that all termination reactions involving a tertiary radical of type $T^*$ are disproportionation reactions (see assumption 2-4.13). In reactions R33 and R34, $R^{{\equiv}^{*}}$ undergoes termination by combination with other secondary radicals, producing a pendant vinyl (arising from $R^{{\equiv}^{*}}$) and a dead polymer molecule. Reaction R35 is interesting because two pendant vinyl groups are generated, one arising from $R^{{\equiv}^{*}}$ and one which is a terminal vinyl group due to termination by disproportionation.
<table>
<thead>
<tr>
<th>Reaction Description</th>
<th>Reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decomposition</td>
<td>$I \xrightarrow{2k_d} fI^* + fI^*$ (R1)</td>
</tr>
<tr>
<td></td>
<td>$\tilde{I} \xrightarrow{k_d} fI^* + f\tilde{I}^*$ (R2)</td>
</tr>
<tr>
<td>Initiation</td>
<td>$I^* + M \xrightarrow{2k_{in,IM}} R^*$ (R3)</td>
</tr>
<tr>
<td></td>
<td>$\tilde{I}^* + M \xrightarrow{2k_{in,IM}} R^* + \tilde{I}$ (R4)</td>
</tr>
<tr>
<td></td>
<td>$I_p^* + M \xrightarrow{2k_{in,IpM}} R^{**}$ (R5)</td>
</tr>
<tr>
<td></td>
<td>$I^* + V_p \xrightarrow{k_{in,IVP}} B^*$ (R6)</td>
</tr>
<tr>
<td></td>
<td>$\tilde{I}^* + V_p \xrightarrow{k_{in,IVP}} B^* + \tilde{I}$ (R7)</td>
</tr>
<tr>
<td></td>
<td>$I_p^* + V_p \xrightarrow{k_{in,IpVp}} B^*$ (R8)</td>
</tr>
<tr>
<td>Propagation</td>
<td>$R^{<strong>} + M \xrightarrow{2k_{p,R}} R^{</strong>} + V_p$ (R9)</td>
</tr>
<tr>
<td></td>
<td>$C^* + M \xrightarrow{2k_{p,C}} R^{**}$ (R10)</td>
</tr>
<tr>
<td></td>
<td>$B^* + M \xrightarrow{2k_{p,B}} R^{**}$ (R11)</td>
</tr>
<tr>
<td>Branching and</td>
<td>$T^* + M \xrightarrow{2k_{p,T}} R^{**}$ (R12)</td>
</tr>
<tr>
<td>Crosslinking</td>
<td>$R^{**} + V_p \xrightarrow{k_{b,R}} B^* + V_p$ (R13)</td>
</tr>
<tr>
<td></td>
<td>$C^* + V_p \xrightarrow{k_{b,C}} B^*$ (R14)</td>
</tr>
<tr>
<td></td>
<td>$B^* + V_p \xrightarrow{k_{b,B}} B^*$ (R15)</td>
</tr>
<tr>
<td></td>
<td>$T^* + V_p \xrightarrow{k_{b,T}} B^*$ (R16)</td>
</tr>
<tr>
<td>Cyclization</td>
<td>$R^{**} \xrightarrow{k_c} C^*$ (R17)</td>
</tr>
<tr>
<td>Backbiting</td>
<td>$R^{**} \xrightarrow{k_{bb,R}} T^* + V_p$ (R18)</td>
</tr>
<tr>
<td></td>
<td>$C^* \xrightarrow{k_{bb}} T^*$ (R19)</td>
</tr>
<tr>
<td>Termination</td>
<td>$I^* + C^* \xrightarrow{k_{LJ,C}} D$ (R20)</td>
</tr>
</tbody>
</table>
\( I^* + B^* \xrightarrow{k_{L,B}} D \) \hspace{1cm} (R21)
\( I^* + T^* \xrightarrow{k_{L,T}} D \) \hspace{1cm} (R22)
\( \tilde{I}^* + C^* \xrightarrow{k_{L,C}} D + \tilde{I} \) \hspace{1cm} (R23)
\( \bar{I}^* + B^* \xrightarrow{k_{L,B}} D + \bar{I} \) \hspace{1cm} (R24)
\( \bar{I}^* + T^* \xrightarrow{k_{L,T}} D + \bar{I} \) \hspace{1cm} (R25)
\( I_p^* + C^* \xrightarrow{k_{L,IpC}} D \) \hspace{1cm} (R26)
\( I_p^* + B^* \xrightarrow{k_{L,IpB}} D \) \hspace{1cm} (R27)
\( I_p^* + T^* \xrightarrow{k_{L,IpT}} D \) \hspace{1cm} (R28)
\( R^* + R^* \xrightarrow{k_{L,RR}} D + 2V_p \) \hspace{1cm} (R29)
\( C^* + C^* \xrightarrow{k_{L,CC}} D \) \hspace{1cm} (R30)
\( B^* + B^* \xrightarrow{k_{L,BB}} D \) \hspace{1cm} (R31)
\( T^* + T^* \xrightarrow{k_{L,TT}} 2D + V_p \) \hspace{1cm} (R32)
\( R^* + C^* \xrightarrow{k_{L,RC}} D + V_p \) \hspace{1cm} (R33)
\( R^* + B^* \xrightarrow{k_{L,RR}} D + V_p \) \hspace{1cm} (R34)
\( R^* + T^* \xrightarrow{k_{L,RT}} 2D + 2V_p \) \hspace{1cm} (R35)
\( C^* + B^* \xrightarrow{k_{L,CB}} D \) \hspace{1cm} (R36)
\( C^* + T^* \xrightarrow{k_{L,CT}} 2D + V_p \) \hspace{1cm} (R37)
\( B^* + T^* \xrightarrow{k_{L,BT}} 2D + V_p \) \hspace{1cm} (R38)
## Table 2-4 Modeling assumptions

<table>
<thead>
<tr>
<th>2-4.1</th>
<th>The thin film layer has constant volume and is spatially uniform.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-4.2</td>
<td>Temperature is constant at 25 °C.</td>
</tr>
<tr>
<td>2-4.3</td>
<td>The concentration of oxygen in the film is negligible.</td>
</tr>
<tr>
<td>2-4.4</td>
<td>The light intensity is uniform within the small film thickness.</td>
</tr>
<tr>
<td>2-4.5</td>
<td>Radicals of type $\tilde{I}^*$ are consumed by initiation reactions R4, R7 and termination reactions R23 to R25 before their weak carbon-phosphorous bond has an opportunity to break.</td>
</tr>
<tr>
<td>2-4.6</td>
<td>All phosphorous-centered radicals have the same kinetically-controlled reactivity, independent of whether they are generated by initiator decomposition or decomposition of initiator moieties $\tilde{I}$ attached to polymer molecules (i.e., $k_{inJM0} = k_{inJpM0}$ and $k_{inJvP0} = k_{inJpVp0}$).</td>
</tr>
<tr>
<td>2-4.7</td>
<td>End groups of type $R^{\equiv*}$ contain a vinyl group and a free radical. We assume that the short-lived free radical is consumed much more rapidly than the vinyl group. When the radical on $R^{\equiv*}$ is consumed by propagation, branching, backbiting and termination reactions, the vinyl group on $R^{\equiv*}$ becomes a pendant vinyl group that can participate in subsequent reactions.</td>
</tr>
<tr>
<td>2-4.8</td>
<td>Cyclization reactions may occur between a radical end $R^{\equiv*}$ and its vinyl group, but not with other vinyl groups on the polymer.</td>
</tr>
<tr>
<td>2-4.9</td>
<td>The only radical ends that participate in backbiting reactions are $R^{\equiv*}$ and $C^<em>$. Backbiting generates tertiary radicals $T^</em>$.</td>
</tr>
<tr>
<td>2-4.10</td>
<td>Self-termination reactions for initiator radicals (i.e., $I^<em>$, $\tilde{I}^</em>$, and $I^<em>_p$) are ignored so that the SSH can be readily applied to compute the concentrations of these short-lived radicals. Cross-termination reactions among $I^</em>$, $\tilde{I}^<em>$, and $I^</em>_p$ are also ignored.</td>
</tr>
</tbody>
</table>
2-4.11 Cross-termination reactions involving radical end $R^{=\cdot}$ and initiator radicals (i.e., $I^\star$, $I^\tilde{\star}$ or $I_P^\star$) are ignored so that the SSH assumption can be applied for $R^{=\cdot}$ (see equation 2-6.7 in Table 2-6). $I^\star$, $I^\tilde{\star}$ or $I_P^\star$ can terminate with other polymeric radicals. $R^{=\cdot}$ can terminate with polymeric radicals ($R^{=\cdot}$, $C^\star$, $B^\star$, and $T^\star$).

2-4.12 Termination reactions involving two secondary macroradicals are termination-by-combination reactions.

2-4.13 Termination reactions involving tertiary radical $T^\star$ and other radicals are termination-by-disproportionation reactions.

2-4.14 The reactivity of terminal vinyl groups resulting from termination by disproportionation is the same as the reactivity of pendant vinyl groups.

2-4.15 The initiator moieties remaining after initiator decomposition have the same absorbance properties as the initiator.

2-4.16 The free volume fraction of the film $v_f$ is assumed to change linearly with monomer conversion $X_M$.

2-4.17 Reaction diffusion, which influences termination rates, occurs predominantly by consumption of monomer vinyl groups rather than pendant vinyl groups.

2-4.18 Carbon-centred and phosphorous-centred initiator radicals have similar relative propensities when reacting with pendant vinyl groups compared with monomeric vinyl groups (i.e., $\frac{k_{in,JV_0}}{k_{in,JM_0}} = \frac{k_{in,IV_0}}{k_{in,IM_0}}$).

2-4.19 The free-volume parameters within the following classes of reactions are equal: i) reactions involving two small molecules (i.e., initiation reactions involving small initiator radicals $I^\star$ or $I^\tilde{\star}$ and monomer), ii) initiation reactions involving a small initiator radical and a large molecule with pendant vinyl groups, iii) reactions involving a small monomer molecule and a free-radical on a large molecule, iv)
branching reactions involving two large molecules, v) termination reactions involving two large molecules, vi) backbiting reactions, and vii) cyclization. The same free-volume parameters are used within each class of reaction to limit the number of parameters that appear in the model. Further, the free-volume parameters for initiation decompositions $f$ and $\tilde{f}$ are the same.

2.4.2 Rate Coefficients

Thirty-eight rate coefficients are required for the reactions listed in Table 2-3. Table 2-5 provides algebraic equations required to compute these rate coefficients, which appear in the dynamic material balances in Table 2-6. Equation 2-5.1 computes the average light intensity in the thin film. Because the film is thin, it is assumed that the light intensity throughout the film is uniform and is equal to $I_t$ (see assumption 2-4.4). In equation 2-5.1, $I_{t0}$ is the light intensity of the light source (i.e., light intensity at the top surface of the film), $\epsilon$ is the extinction coefficient, $[I]_0$ is the initial concentration of initiator, $z$ is the distance from the top surface of the film to any arbitrary depth, and $z_f$ is the film thickness. The initial concentration of initiator is used in equation 2-5.1, because it is assumed that, after decomposition of initiator, its fragments retain their absorbance properties (see assumption 2-4.15).

Table 2-5 Algebraic equations used in the model

| 2-5.1 | $\bar{I}_t = \frac{\int_0^{z_f} I_{t0} e^{-[I]_0 z} dz}{z_f}$ | Goodner and Bowman$^{35}$ |
| 2-5.2 | \[ k_d = \frac{\epsilon \phi \lambda \bar{I}_l}{N_A h c} \] | Goodner and Bowman \textsuperscript{35} |
| 2-5.3 | \[ v_f = v_{f,0} + (v_{f,1} - v_{f,0})X_M \] | Goodner et al. \textsuperscript{36} |
| 2-5.4 | \[ f = \frac{1}{\sqrt{1 - (1 - \frac{1}{f_0})e^{A_f \left(\frac{1}{v_f} - \frac{1}{v_{f,0}}\right)}}} \] | Wenand and McCormick \textsuperscript{1} |
| 2-5.5 | \[ k_{in,j} = \frac{k_{in,j_0}}{\sqrt{1 + e^{A_{in,j} \left(\frac{1}{v_f} - \frac{1}{v_{f,0}}\right)}}} \] | Anseth an Bowman\textsuperscript{37} |
| 2-5.6 | \[ k_{p,j} = \frac{k_{p,j_0}}{\sqrt{1 + e^{A_p \left(\frac{1}{v_f} - \frac{1}{v_{f,c'}}\right)}}} \] | Anseth an Bowman\textsuperscript{37} |
| 2-5.7 | \[ k_{b,j} = \frac{k_{b,j_0}}{\sqrt{1 + e^{A_b \left(\frac{1}{v_f} - \frac{1}{v_{f,b'}}\right)}}} \] | Anseth an Bowman\textsuperscript{37} |
| 2-5.8 | \[ k_C = \frac{k_C_0}{\sqrt{1 + e^{A_C \left(\frac{1}{v_f} - \frac{1}{v_{f,c}}\right)}}} \] | Anseth an Bowman\textsuperscript{37} |
| 2-5.9 | \[ k_{bb,j} = \frac{k_{bb,j_0}}{\sqrt{1 + e^{A_{bb} \left(\frac{1}{v_f} - \frac{1}{v_{f,bb}}\right)}}} \] | Anseth an Bowman\textsuperscript{37} |
| 2-5.10 | \[ k_{in,j} = \frac{k_{in,j_0}}{\sqrt{1 + e^{A_{in,j} \left(\frac{1}{v_f} - \frac{1}{v_{f,in,j}}\right)}}} \] | Anseth an Bowman\textsuperscript{37} |
| 2-5.11 | \[ k_{t,ij} = \sqrt{k_{t,ii} k_{t,jj}} \] | Hutchinson\textsuperscript{38} |
The computed value of $I_\text{t}$ from equation 2-5.1 is used to compute the rate coefficient $k_d$ for initiator decomposition in equation 2-5.2, where $\phi$ is the quantum yield of initiator (i.e., the number of initiator molecules that are dissociated per absorbed photon), $\lambda$ is the light wavelength, $N_A$ is Avogadro’s number, $h$ is Planck’s constant, and $c$ is the speed of light. In the proposed model, the initiation efficiencies $f$ and $\tilde{f}$ in decomposition reactions (i.e., R1 and R2) and the rate coefficients of reactions R3 to R38 may change as the polymerization proceeds due to changes in the free volume of the system.

In the proposed model, it is assumed that free volume fraction $v_f$ is a linear function of monomer conversion $X_M$ where $X_M = \frac{[M]_0 - [M]}{[M]_0}$ (see assumption 2-4.16). Equations 2-5.3 is used to compute the free volume fraction where $v_{f,0}$ is the free volume fraction before the polymerization starts (i.e., at $X_M = 0$), and $v_{f,1}$ is the free volume fraction when all of the monomer has been consumed (i.e., at $X_M = 1$). As the polymerization proceeds, free volume is consumed, so the initiator radicals may be more prone to recombination, because it is more difficult to escape from the cage. As a result, the initiator efficiencies $f$ and $\tilde{f}$ may decrease as polymerization proceeds. Equation 2-5.4 is used to account for the change in $f$ during polymerization, where $f_0$ is initiator efficiency at the start of the polymerization and $A_f$ is a dimensionless free-volume parameter. In the current chapter, it is assumed that free-volume parameters for $f$ and $\tilde{f}$ are the same, so the ratio of $f/\tilde{f}$ is constant during polymerization (assumption 2-4.19). A more important consequence of free-volume consumption during photopolymerization is that diffusion of a variety of species

| 2-5.12 | $k_{t,RR} = \left( k_{t,RRR} \right)^{-1}$ | Goodner et al. 

$$k_{t,RR} = k_{t,RRR} \left[ 1 + \frac{1}{e^{-A_{t,RR}} \left( \frac{1}{v_f} \frac{1}{v_{f,t,RRR}} \right)} + \frac{R_{rd} k_{p,R}[M]}{k_{t,RRR}} \right]^{-1}$$

---

The computed value of $I_t$ from equation 2-5.1 is used to compute the rate coefficient $k_d$ for initiator decomposition in equation 2-5.2, where $\phi$ is the quantum yield of initiator (i.e., the number of initiator molecules that are dissociated per absorbed photon), $\lambda$ is the light wavelength, $N_A$ is Avogadro’s number, $h$ is Planck’s constant, and $c$ is the speed of light. In the proposed model, the initiation efficiencies $f$ and $\tilde{f}$ in decomposition reactions (i.e., R1 and R2) and the rate coefficients of reactions R3 to R38 may change as the polymerization proceeds due to changes in the free volume of the system.

In the proposed model, it is assumed that free volume fraction $v_f$ is a linear function of monomer conversion $X_M$ where $X_M = \frac{[M]_0 - [M]}{[M]_0}$ (see assumption 2-4.16). Equations 2-5.3 is used to compute the free volume fraction where $v_{f,0}$ is the free volume fraction before the polymerization starts (i.e., at $X_M = 0$), and $v_{f,1}$ is the free volume fraction when all of the monomer has been consumed (i.e., at $X_M = 1$). As the polymerization proceeds, free volume is consumed, so the initiator radicals may be more prone to recombination, because it is more difficult to escape from the cage. As a result, the initiator efficiencies $f$ and $\tilde{f}$ may decrease as polymerization proceeds. Equation 2-5.4 is used to account for the change in $f$ during polymerization, where $f_0$ is initiator efficiency at the start of the polymerization and $A_f$ is a dimensionless free-volume parameter. In the current chapter, it is assumed that free-volume parameters for $f$ and $\tilde{f}$ are the same, so the ratio of $f/\tilde{f}$ is constant during polymerization (assumption 2-4.19). A more important consequence of free-volume consumption during photopolymerization is that diffusion of a variety of species
(especially large macro-radicals) within the film becomes slower, resulting in diffusion-dependent reaction rates. The influence of diffusional limitations on reaction rates can be modeled using:

\[
\frac{1}{k_i} = \frac{1}{k_{i0}} + \frac{1}{k_{mi}}
\] (2-1)

where \(k_i\) is the overall rate coefficient for the \(i^{th}\) reaction, \(k_{i0}\) is the value of this rate coefficient when the reaction is kinetically-controlled and \(k_{mi}\) accounts for diffusional resistance. In the proposed model, this type of expression can be used for initiation, propagation, cyclization, backbiting, branching, and termination rate coefficients. Equation 2-5.5 in Table 2-5 provides an expression for rate coefficients in initiation reactions R3 to R8. Subscript \(j\) \((j = IM, \bar{I}M, IpM, IVp, \bar{IV}p, and IpVp)\) in equation 2-5.5 denotes the type of initiation reaction. For example, when \(j = IM\), equation 2-5.5 is used to compute the rate coefficient \(k_{in,IM}\) for reaction R3 involving \(I^*\) and monomer. Notice that the reciprocal of equation 2-5.5 has a structure that is analogous to equation 2-1. As such, \(k_{in,j0}\) is the rate coefficient of the \(j^{th}\) initiation reaction when the reaction is kinetically-controlled. \(A_{in,j}\) is a dimensionless free-volume parameter for the \(j^{th}\) initiation reaction, and \(v_{fc,in,j}\) is the critical free-volume fraction for the \(j^{th}\) initiation reaction. At this critical free volume, the kinetic resistance is equal to the diffusional resistance. Equation 2-5.6 is used to compute the rate coefficients for propagation reactions R9 to R12. The subscript \(j\), where \((j = R, C, B, T)\), denotes the type of radical involved in the propagation reaction. For example, when \(j = R\), equation 2-5.6 computes propagation rate constant \(k_{p,R}\) for reaction R9 involving \(R^*\). Parameters \(A_p\) and \(v_{fc,p}\) are the corresponding free-volume parameter and the critical free-volume fraction for the propagation reactions. Similarly, equation 2-5.7 computes the branching rate coefficients for reactions R13 to R16.

Similar equations are also used to compute the rate coefficients for cyclization reaction R17, backbiting reactions R18 and R19, and self-termination reactions involving initiator radicals \(I^*\) and \(\bar{I}^*\) using equations 2-5.8, 2-5.9, and 2-5.10, respectively. Note that in equation 2-5.9, the
subscript $j$ ($j = R, C$) denotes the type of macroradical involved in backbiting reactions R18 and R19. In equation 2-5.10, the subscript $j$ ($j = II$ and $II$) denotes the type of radicals involved in self-termination reactions involving $I^*$ and $\hat{I}^*$. For example, when the $j = II$, equation 2-5.10 computes $k_{t_{in,II}}$, the self-termination rate coefficient involving two $I^*$ radicals. Note that the reactions involving the self-termination of initiator radicals (i.e., $I^*$, $\hat{I}^*$, and $I_P^*$) are not included in the list of reactions in Table 2-3, because these self-termination reactions are neglected in the reaction mechanism (assumption 2-4.10); however, the rate coefficients $k_{t_{in,II}}$, $k_{t_{in,\hat{I}}}$, and $k_{t_{in,IP}}$ are required to obtain the cross-termination reaction rates involving initiator radicals (i.e., reactions R20 to R28). The rate coefficients of cross-termination reactions are computed using a geometric mean as shown in equation 2-5.11. For example, the rate coefficient for the termination reaction involving a regular radical end $R^{=*}$ ($i = R$) and a cyclized radical end $C^*(j = C)$ is calculated from

$$k_{t,RC} = \sqrt{k_{t,R,R}k_{t,C,C}}.$$

The rate coefficient for self-termination reactions involving two large $R^{=*}$ macroradicals is computed using equation 2-5.12. The additional parameter $R_{rd}$ is used to account for reaction-diffusion, which is important near the end of the polymerization when the dominant mode of radical diffusion is by vinyl group consumption at the chain end, making it possible for large radicals to diffuse, meet and terminate. In the current chapter, it is assumed that only monomer vinyl groups are consumed through reaction-diffusion phenomena (assumption 2-4.17). The rate coefficient for self-termination of two large $I_P^*$ molecules is computed using equation 2-5.12 with the same diffusion-dependent parameters. Further, the kinetically-dependent rate coefficient used to calculate $k_{t_{in,IP}}$ is equal to $k_{t_{in,\hat{I}0}}$, because both $I_P^*$ and $\hat{I}^*$ are phosphorous-centered radicals (assumption 2-4.21).
2.4.3 Balance Equations

Table 2-6 contains dynamic mass balances for chemical species and end groups in the system. To obtain the balance equations in Table 2-6, it is assumed (assumption 2-4.1) that the film is sufficiently thin to be treated as spatially uniform. As shown in equation 2-6.1, the initiator concentration is computed using an analytical solution to the initiator material balance. Most of the other ordinary differential equations (ODEs) are solved numerically, except for the balances on $I^*, \tilde{I}^*, I_p^*$, and $R^{m*}$, which were converted to algebraic equations using the SSH. Initial values for the state variables in the ODEs in Table 2-6 are zero for all states except for the monomer which has $[M]_0 = 4.75 \text{ mol/L}$. The initial initiator concentration used in equation 2-6.1 is $[I]_0 = 0.1027 \text{ mol/L}$ for all of the experiments.
Table 2-6 Dynamic material balance equations for chemical species and functional groups

2-6.1 \[
\frac{d[I]}{dt} = -2k_d[I]
\]
which is solved to give: \([I] = [I]_0 \exp (-2k_d t)\)

2-6.2 \[
\frac{d[\tilde{I}]}{dt} = -k_d[\tilde{I}] + 2k_{in,JM}[\tilde{I}][M] + k_{in,IV,\rho}[\tilde{I}][V_p] + k_{t,JC}[\tilde{I}][C^*] + k_{t,IB}[\tilde{I}][B^*]
+ k_{t,IT}[\tilde{I}][T^*]
\]

2-6.3 \[
\frac{d[I^*]}{dt} = 2f_k_d[I] + \tilde{f}_k_d[I] - 2k_{in,JM} [I^*][M] - k_{in,IV,\rho}[I^*][V_p]
- k_{t,JC}[I^*][C^*] - k_{t,IB}[I^*][B^*] - k_{t,IT}[I^*][T^*]
\]
which becomes, after applying SSH:
\[
[I^*] = \frac{2f_k_d[I] + \tilde{f}_k_d[I]}{2k_{in,JM} [M] + k_{in,IV,\rho}[V_p] + k_{t,JC}[C^*] + k_{t,IB}[B^*] + k_{t,IT}[T^*]}
\]

2-6.4 \[
\frac{d[\tilde{I}^*]}{dt} = 2f_k_d[I] - 2k_{in,JM} [\tilde{I}^*][M] + k_{in,IV,\rho}[\tilde{I}^*][V_p] - k_{t,JC}[\tilde{I}^*][C^*]
- k_{t,IB}[\tilde{I}^*][B^*] - k_{t,IT}[\tilde{I}^*][T^*]
\]
which becomes, after applying SSH:
\[
[\tilde{I}^*] = \frac{2f_k_d[I]}{2k_{in,JM} [M] + k_{in,IV,\rho}[V_p] + k_{t,JC}[C^*] + k_{t,IB}[B^*] + k_{t,IT}[T^*]}
\]

2-6.5 \[
\frac{d[I_p^*]}{dt} = \tilde{f}_k_d [I] - 2k_{in,lp,M} [I_p^*][M] - k_{in,lp,V,\rho}[I_p^*][V_p] - k_{t,lp,C}[I_p^*][C^*]
- k_{t,lp,B}[I_p^*][B^*] - k_{t,lp,T}[I_p^*][T^*]
\]
which becomes, after applying SSH:
\[
[I_p^*] = \frac{\tilde{f}_k_d [I]}{2k_{in,lp,M} [M] + k_{in,lp,V,\rho}[V_p] + k_{t,lp,C}[C^*] + k_{t,lp,B}[B^*] + k_{t,lp,T}[T^*]}
\]
\[ \frac{d[M]}{dt} = -2k_{in,IM} [I^*][M] - 2k_{in,IM} [I^*][M] - 2k_{in,Ip,M}[I_p^*][M] - 2k_{p,R}[R^=][M] \\
- 2k_{p,C} [C^*][M] - 2k_{p,B} [B^*][M] - 2k_{p,T} [T^*][M] \]

\[ \frac{d[R^=]}{dt} = 2_{in,IM} [I^*][M] + 2k_{in,IM} [I^*][M] + 2k_{in,Ip,M}[I_p^*][M] + 2k_{p,C} [C^*][M] \\
+ 2k_{p,B} [B^*][M] + 2k_{p,T} [T^*][M] - k_{b,R} [R^=][V_p] - k_{C}[R^=] \\
- k_{bb}[R^=] - k_{t,RR}[R^=] - k_{t,RC}[R^=][C^*] - k_{t,RB}[R^=][B^*] \\
- k_{t,RT}[R^=][T^*] \]

which becomes, after applying SSH:

\[ [R^=] = \left( \frac{(k_{t,RC}[C^*] + k_{t,RB}[B^*] + k_{t,RT}[T^*] + k_{c} + k_{bb} + k_{b,C} + V_p)}{-2k_{t,RR}} \right)^2 \]

where:

\[ \Delta_R = (k_{t,RC}[C^*] + k_{t,RB}[B^*] + k_{t,RT}[T^*] + k_{c} + k_{bb} + k_{b,C} + V_p)^2 \]

\[ + 8k_{t,RR}(k_{in,IM} [I^*][M]) + 2k_{in,Ip,M}[I_p^*][M] \\
+ k_{p,C} [C^*][M] + k_{p,B} [B^*][M] + k_{p,T} [T^*][M] \]

\[ \frac{d[C^*]}{dt} = k_{C}[R^=] - 2k_{p,C} [C^*][M] - k_{b,C} [C^*][V_p] - k_{bb}[C^*] \\
- k_{t,CC}[C^*][C^*] - k_{t,RC}[R^=][C^*] - k_{t, CB}[C^*][B^*] - k_{t,CT}[C^*][T^*] \\
- k_{t,JC}[I^*][C^*] - k_{t,JC}[I^*][C^*] - k_{t,JP}[I_p^*][C^*] \]

\[ \frac{d[B^*]}{dt} = k_{in,IP} [I^*][V_p] + k_{in,jV_p} [I^*][V_p] + k_{in,Ip,V_p}[I_p^*][V_p] + k_{b,R}[R^=][V_p] \\
+ k_{b,C}[C^*][V_p] + k_{b,T} [T^*][V_p] - 2k_{p,B} [B^*][M] \\
- k_{t, BB}[B^*][B^*] - k_{t, RB}[R^=][B^*] - k_{t, CB}[C^*][B^*] - k_{t, BT}[B^*][T^*] \\
- k_{t, JB}[I^*][B^*] - k_{t, JB}[I^*][B^*] - k_{t, JP}[I_p^*][B^*] \]
2.5 Parameter Estimation and Selection of Important Phenomena

The weighted least-squares objective function that was used for parameter estimation is:

\[
J = \sum_{i=1}^{N_{X_V}} \frac{(X_{V,i} - X_{V,\text{meas},i})^2}{\sigma_V^2} + \sum_{i=1}^{N_{X_M}} \frac{(X_{M,i} - X_{M,\text{meas},i})^2}{\sigma_M^2}
\] (2-2)

where \(X_{V,i}\) is the model prediction of the \(i^{th}\) measured value of the vinyl conversion \(X_{V,\text{meas},i}\). Similarly, \(X_{M,i}\) is the model prediction for the \(i^{th}\) measured value of the monomer conversion \(X_{M,\text{meas},i}\). The total number of vinyl conversion measurements used for parameter estimation is \(N_{X_V} = 2950\), arising from 15 experimental runs used for parameter estimation. The weighting factor in the denominator of the first term of the objective function (i.e., \(\sigma_V^2 = 3.65 \times 10^{-4}\)) is a pooled variance estimate computed using replicate responses from 9 short-time-exposure experiments. The number of monomer-conversion measurements in the second term of the objective function is only \(N_{X_M} = 3\) because the final monomer conversion was measured at the end of only three of the short-time-exposure polymerization experiments. The value of \(\sigma_M^2 = 6.289 \times 10^{-5}\) was set somewhat arbitrarily so that the standard deviation \(\sigma_M\) corresponds to 1.0% of the average measured monomer conversion for the three runs. Note that conversion data near the
start of each run (when reaction rates are changing quickly) are of more scientific and industrial importance than data points near the end of each run. As such, all of the data (sampled every 0.05 s) from early in the runs were used in the objective function. However, only a portion of the data for vinyl conversions at long times were used (i.e., at times beyond 95% of the corresponding final vinyl conversion). The sampling interval for the long-term data was selected so that the number of data values corresponding to times before 95% of the final vinyl conversion matches the number of the data values used in the objective function for the long-term data. Experimental data from two runs (with light intensities 240 and 320 $W/m^2$) were not used for parameter estimation but were saved for model validation.

A list of the 40 adjustable parameters required to solve the equations in Table 2-6 is provided in Table 2-7 along with initial guesses and bounds that were enforced during parameter estimation. Several simplifying assumptions were made to prevent the number of parameters requiring estimation from becoming even larger. For example, it is assumed (assumption 2-4.18) that $\frac{k_{in,VP0}}{k_{in,JM0}} = \frac{k_{in,VP0}}{k_{in,JM0}}$, indicating that carbon-centred and phosphorous-centred radicals have similar preferences when reacting with either pendant vinyl groups or vinyl groups on the monomer. It is also assumed (assumption 2-4.19) that the same free-volume parameters are applicable when calculating $k_{in,JM}$ and $k_{in,JM}$. Similarly, free-volume parameters for initiation reactions involving pendant vinyl groups are assumed to be the same when computing $k_{in,VP}$ and $k_{in,VP}$. Note that it is assumed (assumption 2-4.6) that the kinetically-controlled rate coefficient for initiation reaction involving $I^*_p$ is equal to the corresponding kinetically-controlled rate coefficient of initiation reaction involving $I^*$ (i.e., $k_{in,IM0} = k_{in,Ip0}$ and $k_{in,IP0} = k_{in,IPp0}$). The free-volume parameters for these initiation reactions are not the same, because the size of phosphorous-centered radicals $I^*$ and $I^*_p$ are different. $I^*$ is a small radical which can diffuse through the medium more quickly compared to $I^*_p$, which is a macroradical. There are other classes of reactions which
are assumed to have the same free-volume parameters within the class: i) all propagation reactions with monomer (R9 to R12) and initiation reaction R5 involving $I_P$ and monomer (i.e., all reactions involving a large macroradical and a monomer), ii) all branching reactions (R13 to R16) and initiation reaction R8 involving $I_P$ and pendant vinyl (i.e., all reactions involving a large macroradical and a pendant vinyl), iii) all termination reactions involving two large macroradicals (R26 to R38), and iv) backbiting reactions (R18 and R19).

Further, it is assumed (assumption 2-4.20) that $\frac{k_{b,R0}}{k_{p,R0}} = \frac{k_{b,C0}}{k_{p,C0}} = \frac{k_{b,B0}}{k_{p,B0}} = \frac{k_{b,T0}}{k_{p,T0}}$ so that $C^*, B^*, T^*$ and $R^*$ radicals have similar relative preferences for reacting with pendant vinyl groups compared with vinyl groups on the monomer. The initial guesses and lower and upper bounds for some of the parameters in Table 2-7 were obtained based on previous experimental studies (see references in the final column). Initial guesses and bounds for other parameters were chosen based on our engineering judgment and preliminary simulation results. As shown in Table 2-7, most of the parameters in the current study are estimated using ratios, making it easier to specify reasonable parameter bounds and initial guesses for parameters that have not been estimated previously.

Because there are so many adjustable parameters (i.e., 40), and there is limited information in the available conversion data, it is important to use a systematic method to determine which parameters can and should be estimated while preventing over-fitting. In the current study, an estimability-based parameter subset-selection technique is used to determine which parameters are estimable, based on the sensitivity of the model predictions to the parameters. This analysis also accounts for the uncertainty in the initial parameter guesses (obtained using the bounds in Table 2-7) and for the correlated influences of different parameters. The parameters with the largest influences on predictions of available data tend to be those that are most estimable. Parameters that are not estimable are fixed at their initial guesses or are removed from the model equations via further simplifying assumptions.
Studies of HDDA photopolymerization are limited, so it is difficult to know whether some of the phenomena described in section 2-4.1 (e.g., cyclization and backbiting reactions) are important or not. As a result, we consider four different cases in this parameter-estimation study, with each case becoming more complex as additional reactions and parameters are included. In Case I, which has the least complexity, backbiting (R18 and R19) and cyclization reactions (reactions R17) are not included in the model. Also, the initiator efficiency and the rate coefficients for reactions involving small molecules (i.e., reactions R3 to R7, R9 to R12 and R17 to R25), are assumed to be kinetically rather than diffusionally-controlled. As such, diffusion control is only considered for reactions involving two large molecules in Case I. The corresponding model for Case I has 21 adjustable parameters (i.e., parameters 1 to 4, 6 to 9, 12, 15, 19, 21 to 23, 30, 33 to 37 and 39 in Table 2-7). In Case II, backbiting reactions (reaction R18 and R19) are added to the phenomena considered in the Case I, so that parameters 20, 27 and 40 (see Table 2-7) are included in the list of adjustable parameters. In Case III, the cyclization reaction (R17) is included so that parameters 18, 24, and 38 influence the model predictions. Finally, in Case IV, the initiator efficiency and reactions involving small molecules (i.e., reactions R3 to R7, R9 to R12 and R17 to R25) are assumed to be diffusionally-controlled so that all 40 parameters in Table 2-7 are adjustable.

A summary of the parameter estimability analysis and estimation results for each case is provided in Table 2-8 with additional detail provided in Appendix A. As indicated in Table 2-8, 16 out of 21 parameters were estimable in Case I. The resulting parameter estimates (see Appendix A) were used as initial guesses for the parameter estimation in Case II, because we judged that they are more reliable than the initial guesses shown in Table 2-7. In Case II, 6 parameters of 24 adjustable parameters were estimable, including the ratio between rate coefficients for propagation reaction involving \( T^* \) and propagation reaction involving \( R^* \) (i.e., \( \frac{k_{p,T^*}}{k_{p,R^*}} \)). As shown in Table 2-8, the objective function value decreased by 0.5% between Case I and Case II. As a result, we decided
to include backbiting reactions R18 and R19 and the corresponding reactions arising from backbiting (i.e., reactions R12, R16, R22, R25, R28, R32, R35, and R37 to R38) within the proposed reaction scheme. The parameters estimated from Case II were then used as initial guesses for parameter ranking and estimation in Case III. Note that, at the start of Case III, we increased the upper bound for quantum yield $\phi$ from 0.8 to 0.9, because the estimated value for $\phi$ was at its corresponding upper bound in Cases I and II. In Case III, 27 of the 27 model parameters were estimable, including the rate coefficient for the extra reactions which were added due to cyclization. As shown in Table 2-8, the optimal objective function value in Case III decreased by 63% compared to the value in Case II, revealing that cyclization reactions are important for achieving a good fit to the data. In Case IV, 39 of 40 parameters were estimable. Of the 13 new parameters, which were introduced to account for diffusional effects in reactions involving small molecules and in first-order reactions R17 to R19, only parameters accounting for a decrease in initiator efficiency and propagation reactions were important (i.e., the free volume parameter $A_f$ that accounts for changes in initiator efficiency, and the ratio of critical free volume of propagation reactions to the critical free volume of branching reactions $\frac{v_{fc,p}}{v_{fc,b}}$). Therefore, we decided to run another parameter estimation problem where the diffusion limitations were only accounted for in initiator efficiency, propagation reactions, branching reactions, and termination reactions involving large radicals.

We used the parameter estimates from Case IV as the initial guesses for the parameter estimates that appear in the final model. The results of the parameter estimation for the final model are summarized in the final row of Table 2-8. As shown, only one parameter is estimable in this step, because the initial guesses used were reliable. The estimated value for this estimated parameter is nearly the same as its corresponding initial guess. The final parameter estimates are shown in Table 2-9. The final model accounts for all of the following phenomena: multiple types of vinyl groups, backbiting, cyclization, crosslinking, bifunctional initiator, diffusion-dependence for propagation and for reactions involving two large molecules, and diffusion-dependent initiator
efficacy. Information about parameter significance and correlations is provided in the supporting information.

**Table 2-7** List of parameter initial values and corresponding bounds used during parameter estimation

<table>
<thead>
<tr>
<th>#</th>
<th>Parameter</th>
<th>Initial Guess</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Units</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$v_{f,0}$</td>
<td>0.065</td>
<td>0.055</td>
<td>0.0750</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$v_{f,1}/v_{f,0}$</td>
<td>0.30</td>
<td>0.20</td>
<td>0.50</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$\phi$</td>
<td>0.6</td>
<td>0.5</td>
<td>0.8</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$f_0$</td>
<td>0.7</td>
<td>0.3</td>
<td>1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>$A_f$</td>
<td>0.7</td>
<td>0.5</td>
<td>1</td>
<td>-</td>
<td>Wenand and McCormick(^1)</td>
</tr>
<tr>
<td>6</td>
<td>$k_d/k_d$</td>
<td>0.8</td>
<td>0.1</td>
<td>1.0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>$\tilde{f}/f$</td>
<td>1</td>
<td>0.5</td>
<td>1.2</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>$k_{in,IM0}/k_{in,IM0}$</td>
<td>0.01</td>
<td>0.005</td>
<td>0.1</td>
<td>-</td>
<td>Meng et al.(^40) and Kajiwara et al.(^41)</td>
</tr>
<tr>
<td>9</td>
<td>$k_{in,MO}$</td>
<td>$5 \times 10^7$</td>
<td>$1 \times 10^6$</td>
<td>$5 \times 10^8$</td>
<td>$\frac{L}{mol.s}$</td>
<td>Sumiyoshi and Schnabel(^42)</td>
</tr>
<tr>
<td>10</td>
<td>$A_{in,IM}$</td>
<td>0.7</td>
<td>0.5</td>
<td>1.0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>$v_{fc,in,IM}/v_{fc,IRR}$</td>
<td>0.15</td>
<td>0.1</td>
<td>0.3</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>$k_{in,IV,po}/k_{in,IM0}$</td>
<td>0.8</td>
<td>0.2</td>
<td>1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>$A_{in,IV}$</td>
<td>0.7</td>
<td>0.5</td>
<td>1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>$v_{fc,in,IV}/v_{fc,IRR}$</td>
<td>0.25</td>
<td>0.15</td>
<td>0.4</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>$k_{p,R0}$</td>
<td>$5 \times 10^4$</td>
<td>$1 \times 10^4$</td>
<td>$1 \times 10^5$</td>
<td>$\frac{L}{mol.s}$</td>
<td>Ballard and Asua(^9) Iedemat et al.(^7) and Wen and Junkers(^43)</td>
</tr>
<tr>
<td>16</td>
<td>$A_p/A_b$</td>
<td>0.7</td>
<td>0.5</td>
<td>1.0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>$v_{fc,p}/v_{fc,b}$</td>
<td>0.7</td>
<td>0</td>
<td>1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>$k_{p,CO}/k_{p,R0}$</td>
<td>1</td>
<td>0.5</td>
<td>1.2</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
</tbody>
</table>
|  19 | $k_{p,B0}/k_{p,R0}$  
$= k_{b,CO}/k_{b,R0}$ |  0.1 |  0.01 |  0.5 | - |
|  20 | $k_{p,TE}/k_{p,R0}$  
$= k_{b,Te}/k_{b,R0}$ |  0.001 |  0.0005 |  0.05 | -  
Kattner and Buback\textsuperscript{44} |
|  21 | $k_{b,R0}/k_{p,R0}$ |  0.8 |  0.1 |  1.0 | - |
|  22 | $A_b$ |  0.6 |  0.2 |  1.0 | - |
|  23 | $v_{f,c,b}/v_{f,c,TRR}$ |  0.6 |  0.33 |  1.2 | - |
|  24 | $k_{CO}/k_{p,R0}$ |  0.5 |  0.01 |  5 | - |
|  25 | $A_c$ |  0.7 |  0.5 |  1 | - |
|  26 | $v_{f,c,e}/v_{f,c,TRR}$ |  0.2 |  0.1 |  0.3 | - |
|  27 | $k_{bb,0}/k_{p,R0}$ |  0.008 |  0.001 |  0.01 | -  
Wenn and Junkers\textsuperscript{43} |
|  28 | $A_{bb}$ |  0.7 |  0.5 |  1.0 | - |
|  29 | $v_{f,c,bb}/v_{f,c,TRR}$ |  0.2 |  0.1 |  0.3 | - |
|  30 | $k_{t_{in,II}}$ | $5 \times 10^7$ | $1 \times 10^6$ | $1 \times 10^9$ | $L/mol/s$ |
|  31 | $A_{t_{in,II}}$ |  0.7 |  0.5 |  1 | - |
|  32 | $v_{f,c,t_{in,II}}/v_{f,c,TRR}$ |  0.2 |  0.1 |  0.3 | - |
|  33 | $k_{t_{in,II}}/k_{t_{in,II}}$ |  1 |  0.1 |  1.5 | - |
|  34 | $k_{t,RR,0}$ | $5 \times 10^8$ | $1 \times 10^7$ | $1 \times 10^9$ | $L/mol/s$  
Barth et al. and\textsuperscript{45}  
Iedema et al.\textsuperscript{7} |
|  35 | $A_{t,RR}$ |  0.69 |  0.5 |  1 | -  
Iedema et al.\textsuperscript{7} |
|  36 | $v_{f,c,TRR}$ |  0.093 |  0.07 |  0.1 | -  
Iedema et al.\textsuperscript{7} |
<p>|  37 | $R_{rd}$ |  4 |  0.1 |  8 | $L/mol$ |
|  38 | $k_{f,CCO}/k_{f,RR0}$ |  1 |  0.5 |  1.5 | - |</p>
<table>
<thead>
<tr>
<th>Parameter estimation cases</th>
<th>Optimized objective function value</th>
<th>Total Parameters</th>
<th>Estimable parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case I</td>
<td>$8.60 \times 10^3$</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>Case II</td>
<td>$8.56 \times 10^3$</td>
<td>24</td>
<td>6</td>
</tr>
<tr>
<td>Case III</td>
<td>$3.19 \times 10^3$</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Case IV</td>
<td>$2.26 \times 10^3$</td>
<td>40</td>
<td>39</td>
</tr>
<tr>
<td>Final Model</td>
<td>$2.26 \times 10^3$</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>#</td>
<td>Parameters</td>
<td>Estimate</td>
<td>Unit</td>
</tr>
<tr>
<td>----</td>
<td>-----------------------------------------------</td>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>1</td>
<td>$v_{f,0}$</td>
<td>0.064</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>$v_{f,1}/v_{f,0}$</td>
<td>0.50</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>$\phi$</td>
<td>0.90</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>$f_0$</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>$A_f$</td>
<td>0.50</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>$k_{d}/k_d$</td>
<td>0.57</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>$f/f$</td>
<td>1.20</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>$k_{in,IM0}/k_{in,IM0}$</td>
<td>0.005</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>$k_{in,IM0}$</td>
<td>$1.0 \times 10^6$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>10</td>
<td>$k_{in,IVp0}/k_{in,IM0} = k_{in,IVp0}/k_{in,IM0}$</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>$k_{p,R0}$</td>
<td>$1.00 \times 10^4$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>12</td>
<td>$A_p/A_b$</td>
<td>0.95</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>$v_{f,c,p}/v_{f,c,b}$</td>
<td>0.87</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>$k_{p,C0}/k_{p,R0} = k_{B,C0}/k_{B,R0}$</td>
<td>0.50</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>$k_{p,B0}/k_{p,R0} = k_{B,B0}/k_{B,R0}$</td>
<td>0.0483</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>$k_{p,T0}/k_{p,R0} = k_{B,T0}/k_{B,R0}$</td>
<td>$5.00 \times 10^{-4}$</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>$k_{BR,0}/k_{PR,0}$</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>$A_b$</td>
<td>0.59</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>$v_{f,c,b}/v_{f,c,RR}$</td>
<td>0.53</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>$k_{c,0}/k_{p,R0}$</td>
<td>3.35</td>
<td>mol L$^{-1}$</td>
</tr>
<tr>
<td>21</td>
<td>$k_{B,D0}/k_{p,R0}$</td>
<td>$1.00 \times 10^{-3}$</td>
<td>mol L$^{-1}$</td>
</tr>
<tr>
<td>22</td>
<td>$k_{t,II}/k_{t,II}$</td>
<td>$6.98 \times 10^8$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>23</td>
<td>$k_{t,III}/k_{t,III}$</td>
<td>1.15</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>$k_{t,RR}$</td>
<td>$1.69 \times 10^8$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>25</td>
<td>$A_{t,RR}$</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>26</td>
<td>$v_{f,c,RR}$</td>
<td>0.086</td>
<td>-</td>
</tr>
<tr>
<td>27</td>
<td>$R_{rd}$</td>
<td>7.31</td>
<td>L mol$^{-1}$</td>
</tr>
<tr>
<td>28</td>
<td>$k_{t,CC0}/k_{t,RR0}$</td>
<td>1.50</td>
<td>-</td>
</tr>
<tr>
<td>29</td>
<td>$k_{t,B0}/k_{t,RR0}$</td>
<td>0.01</td>
<td>-</td>
</tr>
<tr>
<td>30</td>
<td>$k_{t,TT0}/k_{t,RR0}$</td>
<td>0.088</td>
<td>-</td>
</tr>
</tbody>
</table>
2.6 Final Simulation Results and Model Validation

Figure 2-6 shows the model fit to the vinyl group conversion data, which was obtained using the final parameter estimates in Table 2-9. The model predictions are very good for all runs including the short-time-exposure experiments and constant-exposure experiments. The proposed model slightly underpredicts the measurements obtained at light intensity $I_{l,0} = 115 \, W/m^2$. This is not surprising, however, considering the variability in overall vinyl group conversions, which is noticeable in the measurements from short-term-exposure experiments. Figure 2-7 shows the fit to the final-monomer-conversion data for the three short-time-exposure experiments. The fit is quite good, with the model slightly underpredicting the final monomer conversion for all three runs where monomer conversion data are available. Figure 2-8 shows model predictions for two validation runs that were not used for parameter estimation (i.e., the collected data shown were not used in the objective function in equation 2-2). Model predictions for validation runs are very good, indicating that the model and the parameter estimates shown in Table 2-9 did not overfit the data. The Supporting Information in Appendix A provides additional plots of initiator, monomer, pendant-vinyl-group and radical concentrations as a function of time for several runs.
Figure 2-6 Experimental data (•, ••, •••, and ×) and corresponding model predictions shown using curves with the same colors obtained using final parameter estimates for a) short-time-exposure experiments with light intensity $I_{L0} = 20 \text{ W/m}^2$ and time exposures $t_{exp} = 0.5, 1, \text{ and } 2 \text{ s}$, and for constant-light-exposure experiments with light intensities b) $I_{L0} = 105 \text{ and } 310 \text{ W/m}^2$, c) $I_{L0} = 115 \text{ and } 760 \text{ W/m}^2$, and d) 195 and 1350 W/m$^2$. 
Figure 2-7 Comparison of model predictions and final monomer conversion data for three short-time-exposure experiments conducted using $I_{l,0} = 20 \text{ W/m}^2$. Straight lines between predicted conversions are shown to guide the eye.

Figure 2-8 Model predictions (curved lines) obtained from final estimated parameters and the corresponding FTIR measured values (● and ○) for validation data obtained using light intensities: a) $I_{l,0} = 240 \text{ W/m}^2$ and b) $I_{l,0} = 320 \text{ W/m}^2$
2.7 Conclusions

A fundamental dynamic model is proposed for the photopolymerization of 1,6-hexanediol diacrylate (HDDA) with a bifunctional initiator, which is important for understanding the underlying phenomena in high-speed printing applications. The proposed model accounts for branching, backbiting and cyclization reactions, as well as the influence of diffusional limitations on polymerization reactions. The model also accounts for different types of radicals that form when weak carbon-phosphorous bonds on the BAPO initiator decompose. Model parameters are estimated using vinyl-group conversion data and a limited amount of monomer conversion data obtained using FTIR and 1H-NMR, respectively. A systematic parameter ranking and estimation methodology was utilized to determine which reactions and diffusional effects are important for achieving accurate predictions of the data. This analysis revealed that including branching, backbiting and cyclization reactions in the reaction scheme is important for achieving a good fit to the data and reliable predictions of validation data. Furthermore, the parameter-estimation results indicate that diffusional limitations have an important influence on termination, branching, and propagation reactions, but not for backbiting and cyclization. Accounting for diffusional effects on initiator efficiency is also important for achieving reliable conversion predictions.

Model predictions are very good for experiments whose corresponding data were used for parameter estimation (i.e., within ± 1.61% for overall vinyl group conversion) and for experiments whose corresponding data were used for validation (i.e., within ± 1.34% for overall vinyl group conversion).

An important contribution of the current modeling study is that it helps to justify simplifying assumptions about diffusional effects that significantly reduce model complexity and the number of adjustable parameters (i.e., from 40 to 30). This is an important step prior to new modeling research that will explore the influence of oxygen and the variation in species concentrations over the depth of the film. We recommend that additional data should be collected
to support future parameter-estimation and model-validation studies. For example, monomer-conversion data should be collected for additional runs. Also, it will be particularly helpful if concentrations of cyclized groups or branch points can be determined experimentally. Additional replicate experiments should be performed because variability in experimental inputs, such as light intensity, pulse duration, film thickness, and oxygen contamination may lead to run-to-run variability, which should be considered during parameter estimation.
2.8 References for Chapter 2


Chapter 3

Estimation of Output Measurement Variances for EVM Parameter Estimation†

3.1 Abstract

Error-in-variables model (EVM) methods require information about variances of input and output measured variables when estimating the parameters in mathematical models for chemical processes. In EVM, using replicate experiments for estimating output measurement variances is complicated because true values of inputs may be different when multiple attempts are made to repeat an experiment. To address this issue, we categorize attempted replicate experiments as: i) true replicates (TRs) when uncertain inputs are the same in replicated runs and ii) pseudo replicates (PRs) when measured inputs are the same, but unknown true values of inputs are different. We propose methodologies to obtain output measurement variance estimates and associated parameter estimates for both situations. We also propose bootstrap methods for obtaining joint-confidence information for the resulting parameter estimates. A copolymerization case study is used to illustrate the proposed techniques. We show that different assumptions noticeably affect the uncertainties in the resulting reactivity-ratio estimates.

† This chapter is published as: Abdi, K. and McAuley, K.B., 2022. AIChE Journal, p.e17735.
3.2 Introduction

Fundamental mathematical models are widely used for chemical process development and improvement. These models usually contain unknown parameters that require estimation.\(^1\),\(^2\) In conventional parameter-estimation methodologies such as weighted least-squares (WLS) estimation, the model inputs are assumed to be perfectly known (i.e., measured without error), while random measurement errors are considered for the model outputs.\(^3\),\(^4\) Nevertheless, there are important situations where model inputs are not perfectly known, but contain significant uncertainties. The Error-in-Variables Model (EVM) technique was developed to account for measurement uncertainties in both inputs and outputs during parameter estimation.\(^5\),\(^6\) In EVM, the true values of the measured inputs are estimated along with the model parameters. EVM parameter estimation requires more information about the measurement uncertainties than WLS, and more decision variables are involved when minimizing EVM objective functions. As a result, in situations where the input uncertainties are small compared to uncertainties in measured outputs, we recommend using WLS. However, in situations where the uncertainties in inputs are relatively large, using WLS parameter estimation might result in unrealistic parameter estimates and underprediction of parameter uncertainties. EVM\(^7\)\(^-\)\(^30\) has been used in a variety of chemical engineering modeling studies (see Table 3-1 and Table 3-2).

The following single-response model can be used to illustrate assumptions that are made during an EVM parameter estimation wherein \(n_i\) replicated experiments have been attempted at \(N\) different target operating conditions (for \(i = 1, \ldots, N\)):

\[
Y_{ij} = g(u_{ij}, x_i, \theta) + \varepsilon_{Y_{ij}} \quad \text{(3-1)}
\]

\[
U_{ij} = u_{ij} + \varepsilon_{U_{ij}} \quad \text{(3-2)}
\]

Note that we use target conditions to describe unique experimental points at which we aim to collect output measurements. The number of data points used for parameter estimation will be larger than \(N\) if replicates have been attempted at some of the target conditions. In equation (3-1), \(Y_{ij}\) is the measured value of the response for the \(j^{th}\) \((j = 1, \ldots, n_i)\) run conducted using the \(i^{th}\)
target experimental condition, \( g \) is a solution of nonlinear model equation(s), \( u_{ij} \) is the true value of the uncertain input for the \( j^{th} \) run at the \( i^{th} \) target condition, \( x_i \in \mathbb{R}^{N_x} \) is a vector containing any perfectly-known model inputs for the \( i^{th} \) experimental condition, \( \theta \in \mathbb{R}^{N_{\theta}} \) is the vector of model parameters, and \( \varepsilon_{Y,ij} \) is the random measurement noise for the response \( g( u_{ij}, x_i, \theta ) \). The variables \( x_i \) and \( u_{ij} \) correspond to the independent variables in the experiments, and \( Y_{ij} \) corresponds to the dependent variables in the experiments. In equation (3-2), \( U_{ij} \) is the measured value of the uncertain input \( u_{ij} \) and \( \varepsilon_{U,ij} \) is the corresponding random measurement noise. We assume that \( \varepsilon_{Y,ij} \) and \( \varepsilon_{U,ij} \) in equations (3-1) and (3-2) are independent random variables with \( \varepsilon_{Y,ij} \sim \mathcal{N}(0, \sigma_Y^2) \) and \( \varepsilon_{U,ij} \sim \mathcal{N}(0, \sigma_U^2) \).

Based on the model in equations (3-1) and (3-2), a maximum-likelihood approach can be used to obtain the following objective function for use in EVM parameter estimation:

\[
J_{EVM} = \sum_{i=1}^{N} \sum_{j=1}^{r_i} \left( \frac{(y_{m,ij} - g(x_i, u_{ij}, \theta))^2}{\sigma_Y^2} + \frac{(u_{m,ij} - u_{ij})^2}{\sigma_U^2} \right) \quad (3-3)
\]

where \( y_{m,ij} \) is the measured value of output from the \( j^{th} \) run conducted using the \( i^{th} \) target run condition, and \( u_{m,ij} \) is the measured value of the corresponding uncertain input. The objective function in equation (3-3) is minimized to simultaneously determine the parameter estimates \( \hat{\theta} \) and estimates \( \hat{u}_{ij} \) for the true values of the uncertain inputs. As shown in equation (3-3), variances of the measured model inputs \( \sigma_{\theta}^2 \) and outputs \( \sigma_Y^2 \) are weighting factors in the EVM objective function and therefore influence the parameter estimates. Also, confidence intervals for the parameters depend on these variances.\(^{31}\) The correlation between measurements of the uncertain inputs and outputs is neglected in objective function (3-3), which is a common assumption in EVM parameter estimation studies. When the errors in uncertain inputs \( u_{ij} \) are negligible, the second term in objective function (3-3) can be removed, resulting in the WLS objective function, which is mostly used to estimate model parameters. In conventional WLS parameter estimation with perfectly-known inputs, information is only required about output measurement variances \( \sigma_Y^2 \), making it more
straight-forward than EVM parameter estimation, which requires additional information about input measurement variances $\sigma_U^2$.

Furthermore, in WLS parameter-estimation studies, where inputs are perfectly known, measurement variances are often estimated from replicate experiments.\textsuperscript{32-34} The influence of all uncertainties and disturbances associated with setup of the equipment, measurement and mixing of reactants, conduct of the experiments, collection of any samples for analysis, and measurement of the desired responses is lumped together and treated as part of the output uncertainty.\textsuperscript{34} However, when some of the inputs are acknowledged to contain important uncertainties, so that EVM parameter estimation is required, the issue of replicate experiments and output variance estimation is more complex. In some situations, the experimentalist may believe that true replicate experiments have been conducted (i.e., all experiments conducted at the $i^{th}$ target run condition have the same unknown value of $u_{ij}$). For example, the modeler will have true replicates when repeated experiments are conducted using the same batch of input material that has uncertain properties. In other situations, it may not be appropriate to assume that $u_{ij}$ is the same for all runs conducted at the $i^{th}$ target condition (e.g., when attempted replicate experiments use different batches of input material with uncertain compositions). In the current chapter, we refer to these runs with the same target settings, but different values of $u_{ij}$ as pseudo-replicate runs. When a modeler is confronted with pseudo-replicate experiments, the issue of estimating $\sigma_Y^2$ from data becomes complicated because the measured values $y_{m,ij}$ are influenced by random errors in the inputs as well as the corresponding output.

Table 3-1 and Table 3-2 summarize previous EVM studies in the chemical engineering literature where data from experiments with uncertain inputs were used for parameter estimation.\textsuperscript{4,7} Studies in Table 3-1 involve reactivity-ratio estimation in copolymerization and terpolymerization studies. The authors used EVM for these studies because important model inputs (i.e., initial feed compositions or comonomer concentrations in the reactor)
have significant uncertainties compared to the measurement uncertainty in the resulting copolymer or terpolymer composition. Other studies involving different types of chemical engineering models are summarized in Table 3-2. These studies are related to vapour-liquid-equilibrium, gas-phase catalytic hydrogenation, gas-solid adsorption, oxygen consumption in wastewater, water gas-shift reaction, ion-exchange equilibrium, gas purification, mass-transfer in liquid mixtures, a low-density polyethylene reactor, and a natural gas network. The uncertain inputs in these systems include measured inlet temperatures, inlet flowrates, inlet mole fractions, partial pressures, and liquid-phase concentrations.

In most EVM parameter-estimation studies, variances of the uncertain model inputs are estimated from data or are assumed to be known based on previous studies. Users of EVM can usually obtain reliable values of $\sigma_U^2$ for use in their objective functions, but it is much more difficult to obtain a reliable estimate of $\sigma_Y^2$. Repeating measurements of the same output sample only captures the portion of the variability that is attributable to noise in the analytical device. Other types of variability that should be included in $\sigma_Y^2$ (e.g., due to equipment set up and experimental procedures) can only be detected when full experiments are repeated. This difficult issue has not received adequate consideration in the EVM literature.

As shown in Table 3-1 and Table 3-2, values of $\sigma_Y^2$ were reported in 29 of the 39 studies, while the remaining authors were silent about the weighting factors used for parameter estimation. Authors from 23 of the 39 studies assumed that $\sigma_Y^2$ was known a priori (indicated using AK in the 4th column), while 14 studies were silent about how $\sigma_Y^2$ was obtained (indicated using ? in the 4th column). In three of the 39 studies, the authors estimated $\sigma_Y^2$ from replicate data (indicated using R in the 4th column). Replicate experimental results were reported in 16 out of the 39 EVM studies. In one of these 16 studies, it is clear that true replicate experiments were conducted so that $\sigma_Y^2$ was relatively easy to estimate. In this study, Keeler and Reilly purchased several standard ethylene/propylene copolymer samples and used them to fit a calibration curve between IR (the
model input) and NMR (the model output) measurements. Repeated measurements of the NMR response were obtained for each standard sample, which has a constant composition and therefore a fixed true value of the unknown (noise-free) IR response. In another of the 16 studies, it is clear that Scott and Penlidis obtained pseudo-replicate data. In this terpolymerization study, they made multiple solutions of monomers with the same target concentrations and used them to repeat some experiments. As such, their reported input measurements might have different true values, even though the corresponding reported input values are the same. In the remaining 14 studies with replicates, the authors do not provide sufficient details about how the repeated runs were conducted, so it is difficult to ascertain whether true replicates or pseudo replicates were performed (indicated using R? in the 5th columns). Based on the articles summarized in Table 3-1 and Table 3-2, it is apparent that users of EVM are not attempting to obtain reliable estimates of \( \sigma^2 \) from data. Note that EVM parameter estimation has been used by engineers and scientists in a variety of fields beyond chemical engineering. To our knowledge, reliable methods for estimating \( \sigma^2 \) from data have not been reported in these other EVM studies.

Researchers who perform EVM parameter estimation are interested in quantifying the uncertainties in their parameter estimates. As shown in Table 3-1, all research groups who estimated reactivity ratios obtained joint confidence regions (JCRs) and/or confidence intervals (CIs) for their parameters. Only three of the 28 studies in Table 3-1 provide details about how their parameter uncertainty information was obtained. All three studies relied on the assumption that output measurement variances were perfectly known.

The objectives of the current study are: i) to propose approaches for estimating \( \sigma^2 \) from true-replicate and pseudo-replicate data, ii) to use a copolymerization case study to illustrate how \( \sigma^2 \) can be reliably estimated in situations involving true replicates and pseudo replicates, and iii) to estimate joint confidence (JC) information for the estimated parameters in the copolymerization model. The remainder of the chapter is organized as follows. First, we describe three different
situations involving replicate experiments in EVM and provide information about suitable EVM objective functions. Next, methods for calculating $\hat{\sigma}_Y^2$ in EVM situations involving different types of replicate data are explained. Finally, copolymerization of n-butyl methacrylate and n-butyl acrylate is used as a case study to show how $\sigma_Y^2$ and $\theta$ should be estimated in EVM situations involving true replicates and pseudo replicates. We also propose a bootstrapping method to obtain JC information for reactivity-ratio estimates, based on the value for $\hat{\sigma}_Y^2$.

Table 3-1 Summary of studies where EVM used for reactivity-ratio estimation

<table>
<thead>
<tr>
<th>Authors</th>
<th>Monomers</th>
<th>Variance estimates reported</th>
<th>How variances obtained</th>
<th>Type of replicate experiments</th>
<th>Parameter uncertainty estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dube et al. 14,38</td>
<td>Styrene/butyl acrylate</td>
<td>Yes</td>
<td>?</td>
<td>R?</td>
<td>JCR</td>
</tr>
<tr>
<td>Brar et al. 4</td>
<td>Styrene/butyl acrylate</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
<td>JCR and CIs</td>
</tr>
<tr>
<td>Brar and Charan 56</td>
<td>Vinyl acetate/methyl acrylate</td>
<td>No</td>
<td>?</td>
<td>None</td>
<td>CIs</td>
</tr>
<tr>
<td>Brar and Charan 57</td>
<td>Vinyl acetate/ethyl methacrylate</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
<td>CIs</td>
</tr>
<tr>
<td>Dube and Penidis 38; Scott and Penidis 39</td>
<td>Butyl acrylate/methyl methacrylate, butyl acrylate / vinyl acetate, and methyl methacrylate/vinyl acetate</td>
<td>Yes</td>
<td>AK</td>
<td>R?</td>
<td>JCR and CIs</td>
</tr>
<tr>
<td>Schoonbrood et al. 25</td>
<td>Styrene/2-Hydroxyethyl methacrylate</td>
<td>No</td>
<td>?</td>
<td>None</td>
<td>JCR and CIs</td>
</tr>
<tr>
<td>Brar et al. 40</td>
<td>Acrylonitrile/methacrylic acid</td>
<td>Yes</td>
<td>?</td>
<td>None</td>
<td>JCR and CIs</td>
</tr>
<tr>
<td>Brar and Malhorta 41</td>
<td>Vinylidene chloride/methyl acrylate</td>
<td>Yes</td>
<td>?</td>
<td>None</td>
<td>CIs</td>
</tr>
<tr>
<td>McManus and Penidis 42; Scott and Penidis 39</td>
<td>Styrene/ethyl acrylate</td>
<td>Yes</td>
<td>AK</td>
<td>R?</td>
<td>JCR</td>
</tr>
<tr>
<td>Suddaby et al. 16</td>
<td>Methyl methacrylate/n-butyl methacrylate</td>
<td>Yes</td>
<td>AK</td>
<td>R?</td>
<td>JCR</td>
</tr>
<tr>
<td>Brar et al. 27</td>
<td>Acrylonitrile/glycidyl methacrylate</td>
<td>No</td>
<td>?</td>
<td>None</td>
<td>JCR and CIs</td>
</tr>
<tr>
<td>Authors</td>
<td>Formula</td>
<td>Polymers</td>
<td>Data Availability</td>
<td>Neoprene</td>
<td>JCR Status</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>----------</td>
<td>------------------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>Hakim et al.</td>
<td>Butyl acrylate/methyl methacrylate</td>
<td>No</td>
<td>?</td>
<td>R?</td>
<td>JCR</td>
</tr>
<tr>
<td>Baradie et al.</td>
<td>Tetrafluoroethylene</td>
<td>Yes</td>
<td>R</td>
<td>R?</td>
<td>JCR and CIs</td>
</tr>
<tr>
<td>Kazemi et al.</td>
<td>Tetrafluoroethylene</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
<td>JCR</td>
</tr>
<tr>
<td>Zhang and Dube</td>
<td>N-Butyl Methacrylate/d-limonene</td>
<td>No</td>
<td>?</td>
<td>R?</td>
<td>JCR</td>
</tr>
<tr>
<td>Kazemi et al.</td>
<td>N-Butyl Methacrylate/d-limonene</td>
<td>Yes</td>
<td>AK</td>
<td>R?</td>
<td>JCR</td>
</tr>
<tr>
<td>Mathew and Duever</td>
<td>Di-n-butyl itaconate (DBI)/methyl methacrylate</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
<td>JCR</td>
</tr>
<tr>
<td>Ren et al.</td>
<td>N-Butyl acrylate/n-Butyl methacrylate</td>
<td>Yes</td>
<td>AK</td>
<td>R?</td>
<td>JCR</td>
</tr>
<tr>
<td>Scott and Penlidis</td>
<td>2-methylene-1,3-dioxepane/vinyl acetate</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
<td>JCR</td>
</tr>
<tr>
<td>Scott and Penlidis</td>
<td>2-acrylamido-2-methylpropene sulfonic acid/acrylamide / acrylic acid</td>
<td>No</td>
<td>?</td>
<td>Pseudo replicate</td>
<td>JCR</td>
</tr>
<tr>
<td>Gabriel and Dube</td>
<td>N-Butyl Acrylate/2-Ethylhexyl Acrylate/Methyl Methacrylate</td>
<td>No</td>
<td>?</td>
<td>R?</td>
<td>JCR</td>
</tr>
<tr>
<td>Yousefi et al.</td>
<td>Triisopropylsilyl Acrylate/ Methyl Methacrylate/ Butyl Acrylate</td>
<td>No</td>
<td>?</td>
<td>R?</td>
<td>JCR</td>
</tr>
<tr>
<td>Krieger et al.</td>
<td>Itaconic Acid/N-Vinyl-2-Pyrrolidone</td>
<td>No</td>
<td>?</td>
<td>None</td>
<td>JCR</td>
</tr>
<tr>
<td>Santos et al.</td>
<td>Styrene/VeoVa-10 copolymerization</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
<td>JCR and CIs</td>
</tr>
</tbody>
</table>
Table 3-2 Summary of studies involving other Chemical Engineering models where EVM used for parameter estimation

<table>
<thead>
<tr>
<th>Authors</th>
<th>Model of study</th>
<th>Variance estimates are reported</th>
<th>How variances were determined</th>
<th>Type of replicate experiments reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keeler and Reilly</td>
<td>Copolymer composition calibration between IR and NMR measurements</td>
<td>No</td>
<td>R</td>
<td>True replicate</td>
</tr>
<tr>
<td>Sutton and MacGregor</td>
<td>Vapour-liquid equilibrium for ethanol-isoctane and benzene- heptane systems</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
</tr>
<tr>
<td>Duever et al.</td>
<td>Vapor-liquid equilibrium</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
</tr>
<tr>
<td>Rod and Hancil</td>
<td>Gas-phase hydrogenation of phenol on a palladium catalyst</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
</tr>
<tr>
<td>High and Danner</td>
<td>Gas-solid adsorption</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
</tr>
<tr>
<td>Valko and Vajda</td>
<td>Biological oxygen consumption and chemical oxygen consumption relationship in a wastewater</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
</tr>
<tr>
<td>Kim et al.</td>
<td>Vapour-liquid equilibrium for methanol and 1,2-dichloroethane</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
</tr>
<tr>
<td>Bardow and Marquardt</td>
<td>Liquid-liquid diffusion</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
</tr>
<tr>
<td>Kim et al.</td>
<td>Water-gas shift reaction</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
</tr>
<tr>
<td>Vamos and Haas</td>
<td>Binary ion-exchange equilibrium</td>
<td>Yes</td>
<td>R</td>
<td>R?</td>
</tr>
<tr>
<td>Faber and Wozny</td>
<td>Coke-oven-gas purification process</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
</tr>
<tr>
<td>Zavala and Biegler</td>
<td>Low-density polyethylene tubular reactor</td>
<td>No</td>
<td>?</td>
<td>None</td>
</tr>
<tr>
<td>Leung et al.</td>
<td>Gas lateral network of TransCanada Pipeline system</td>
<td>Yes</td>
<td>AK</td>
<td>None</td>
</tr>
</tbody>
</table>
3.3 Proposed Classification of Replicate Experiments in EVM

Table 3-3 and Table 3-4 provide information about model equations and objective functions corresponding to three different sets of assumptions about replicate (or nearly replicated) experiments. The first two classifications, TR1 and TR2, correspond to “true replicate” experiments. By true replicates, we mean that, when several experiments are conducted using the same target settings, the true values of the uncertain inputs can be assumed the same for the replicated runs. The third classification corresponds to “pseudo-replicate” experiments, denoted by PR. By pseudo-replicate runs, we mean that the true values for the uncertain inputs are different when several experiments are conducted using the same target conditions. Two different situations are considered for true replicates because sometimes modelers will have only one measured value for an uncertain input used in the repeated runs (situation TR1) and sometimes modelers will have new measured values for uncertain inputs in each replicated run, even though the modeler assumes that the underlying true values have not changed.

Equation (3-3.1) in Table 3-3 is the model equation used to describe the measured outputs for situations TR1 and TR2. Here $Y_{ij} \in R^{N_y}$ is the vector of measured responses for the $j^{th}$ replicate experiment conducted at the $i^{th}$ target run condition, where $N$ is the number of distinct run conditions in the data set (i.e., $i = 1, \ldots, N$) and $n_i$ is the number of repeated runs performed at the $i^{th}$ target run condition (i.e., $j = 1, \ldots, n_i$). In equation (3-3.1), $g \in R^{N_y}$ is a vector of solutions of nonlinear equations, $u_i \in R^{Nu}$ is the vector of unknown true values for the uncertain model inputs for the $i^{th}$ target condition and $\epsilon_{Y_{ij}} \in R^{N_y}$ is a vector of random noise for the corresponding vector of output measurements $Y_{ij}$. Equation (3-3.2) in Table 3-3 is the model equation used to describe the measured values of the uncertain inputs in situation TR1. $U_i$ is the vector of uncertain input measurements for the $i^{th}$ target experimental condition and $\epsilon_{U_{i}} \in R^{Nu}$ is the corresponding vector of random measurement noise. In the TR1 situation, only one measurement is obtained for each uncertain input. As a result, the vector of input measurement noise $\epsilon_{U_{i}}$ in Equation (3-3.1)
has a single index $i$. However, the symbol for the vector of output measurement random noise $\varepsilon_{Y,ij}$ has two indices, because $n_i$ measurements are obtained for the model outputs at each target condition.

Similarly, equation (3-3.3) is used to describe measured values of the uncertain inputs for situation TR2. Here $U_{ij}$ is the $j^{th}$ vector of uncertain input measurements for the $i^{th}$ target experimental condition, and $\varepsilon_{U,ij}$ is the corresponding vector of measurement noise. The extra subscript $j$ compared with equation (3-3.2) appears because, in situation TR2, each repeated run has its own set of measurements for the unknown inputs. Equations (3-3.4) and (3-3.5) are the model equations for the measured outputs and measured uncertain inputs for the situation where pseudo-replicate runs have been conducted. Notice that equation (3-3.4) is the same as equation (3-3.1) except that the vector of true values for the uncertain inputs $u_{ij}$ has an extra subscript, $j$. Here, the modeler assumes that the true inputs are different for each attempted replicate experiment. As a result, equation (3-5) also contains the vector $u_{ij}$.

Objective functions for parameter estimation using the three different assumptions about replicates are provided in Table 3-4. The objective function in equation (3-4.1) is minimized to simultaneously to obtain the vector of parameter estimates $\hat{\theta}$ and the estimates of uncertain inputs in vector $\hat{u}_i$ for each target run condition. In objective function (3-4.1), $y_{m,ij}$ is the vector of output measurements from the $j^{th}$ run conducted using the $i^{th}$ target condition, and $u_{m,i}$ is the vector of uncertain input measurements for the $i^{th}$ target run. Note that the lower-case letters $y$ and $u$ in $y_{m,ij}$ and $u_{m,i}$, respectively, are used to show that these quantities are samples obtained from random variables $Y_{ij}$ and $U_i$. The total number of data values used for parameter estimation in situation TR1 is $N_Y \sum_{i=1}^{N} n_i + NN_{Uj}$ as shown in the 3$^{rd}$ column of Table 3-3. Similarly, equation (3-4.2) is the objective function for situation TR2, where $u_{m,ij}$ is the vector of uncertain input measurements for the $j^{th}$ run using the $i^{th}$ experimental target. The total number of data values in situation TR2...
is larger than in TR1 because more uncertain input measurements are available. Equation (3-4.3) is used when true inputs are assumed to be different whenever replicate experiments are attempted.

As a result, more unknown input values (in vectors $u_{ij}$) require estimation than in situations TR1 and TR2. As shown in Table 3-3, the PR situation requires estimation of $N_U \sum_{i=1}^N n_i$ unknown input values, while situations TR1 and TR2 require estimation of $NN_U$ unknown inputs.

**Table 3-3** Error-in-variables model equations for different assumptions about replicate runs

<table>
<thead>
<tr>
<th>Case</th>
<th>EVM Equations</th>
<th>Number of Data Values</th>
<th>Number of parameters and inputs requiring estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR1</td>
<td>$Y_{ij} = g(x_i, u_i, \theta) + \varepsilon_{Y,ij}$</td>
<td>$N_Y \sum_{i=1}^N n_i + NN_U$</td>
<td>$N_{\theta} + NN_U$</td>
</tr>
<tr>
<td></td>
<td>$U_i = u_i + \varepsilon_{U,i}$</td>
<td>(3.3.2)</td>
<td></td>
</tr>
<tr>
<td>TR2</td>
<td>$Y_{ij} = g(x_i, u_i, \theta) + \varepsilon_{Y,ij}$</td>
<td>$\sum_{i=1}^N n_i (N_Y + N_U)$</td>
<td>$N_{\theta} + NN_U$</td>
</tr>
<tr>
<td></td>
<td>$U_{ij} = u_{ij} + \varepsilon_{U,ij}$</td>
<td>(3.3.3)</td>
<td></td>
</tr>
<tr>
<td>PR</td>
<td>$Y_{ij} = g(x_i, u_{ij}, \theta) + \varepsilon_{Y,ij}$</td>
<td>$\sum_{i=1}^N n_i (N_Y + N_U)$</td>
<td>$N_{\theta} + N_U \sum_{i=1}^N n_i$</td>
</tr>
<tr>
<td></td>
<td>$U_i = u_{ij} + \varepsilon_{U,ij}$</td>
<td>(3.4.1)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3-4 Objective Functions for Parameter Estimation in Error-in-Variables Models with Different Assumptions about Replicate Runs

<table>
<thead>
<tr>
<th>Case</th>
<th>EVM objective function and decision variables requiring estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR1</td>
<td>$J_{TR1} = \sum_{i=1}^{N} \sum_{j=1}^{n_i} \left( y_{m,ij} - g(x_i, u_i, \theta) \right)^T \Sigma_Y^{-1} \left( y_{m,ij} - g(x_i, u_i, \theta) \right) + \sum_{i=1}^{N} (u_{mi} - u_i)^T \Sigma_U^{-1} (u_{mi} - u_i)$</td>
</tr>
<tr>
<td></td>
<td>Decision variables: $u_i$ ($i = 1, ..., N$) and $\theta$</td>
</tr>
<tr>
<td>TR2</td>
<td>$J_{TR2} = \sum_{i=1}^{N} \sum_{j=1}^{n_i} \left( y_{m,ij} - g(x_i, u_i, \theta) \right)^T \Sigma_Y^{-1} \left( y_{m,ij} - g(x_i, u_i, \theta) \right) + \left( u_{mij} - u_i \right)^T \Sigma_U^{-1} (u_{mij} - u_i)$</td>
</tr>
<tr>
<td></td>
<td>Decision variables: $u_i$ ($i = 1, ..., N$) and $\theta$</td>
</tr>
<tr>
<td>PR</td>
<td>$J_{PR} = \sum_{i=1}^{N} \sum_{j=1}^{n_i} \left( y_{m,ij} - g(x_i, u_{ij}, \theta) \right)^T \Sigma_Y^{-1} \left( y_{m,ij} - g(x_i, u_{ij}, \theta) \right) + \left( u_{mi} - u_{ij} \right)^T \Sigma_U^{-1} (u_{mi} - u_{ij})$</td>
</tr>
<tr>
<td></td>
<td>Decision variables: $u_{ij}$ ($j = 1, ..., n_i$ and $i = 1, ..., N$) and $\theta$</td>
</tr>
</tbody>
</table>
3.4 Proposed Output Measurement-Variance Estimation in Situations Involving Replicate Experiments

In this section, methods for estimation of output measurement variances are proposed for situations involving data values from replicate or pseudo-replicate runs. In situations TR1 and TR2, where true replicate experiments have been performed, the methods for calculation of variance estimates are similar. The estimated variance in the measurements for the $k^{th}$ model output obtained using the $i^{th}$ run condition is:

$$\hat{\sigma}^2_{Y_{ik}} = \frac{\sum_{j=1}^{n_i} (y_{m,ijk} - \bar{y}_{m,ik})^2}{n_i - 1} \quad (i = 1, ..., N \text{ and } k = 1, ..., N_Y) \quad (3-4)$$

where $y_{m,ijk}$ is a measured value obtained from the $j^{th}$ replicate run and $\bar{y}_{m,ik}$ is the corresponding average for the measurements obtained from the $n_i$ repeated runs. For situations TR1 and TR2, measurement variances for each model output can be estimated by pooling variances obtained from different run conditions:

$$\hat{\sigma}^2_{Y_k} = \frac{\sum_{i=1}^{N} (n_i - 1) \hat{\sigma}^2_{Y_{ik}}}{\sum_{i=1}^{N} (n_i - 1)} \quad (k = 1, ..., N_Y) \quad (3-5)$$

When pseudo replicates are conducted, variance estimates obtained from equations (3-4) and (3-5) could be too large, because uncertainties associated with the different true inputs may inflate the variability of the measured outputs at each target condition.

We propose a linearization-based approach for the PR situation so that more-accurate estimates of $\hat{\sigma}^2_{Y_k}$ can be computed. Using a Taylor-series expansion and Equation (3-3.4) in Table 3-3, the $k^{th}$ element of measurement vector $Y_{ij}$ can be estimated:

$$Y_{ijk} \approx g_k(x_i, u_{m,i}, \theta) + \frac{\partial g_k}{\partial u} \bigg|_{u_{m,i},x_i,\theta} (u_{ij} - u_{m,i}) + \varepsilon_{Y_{ijk}} \quad (3-6)$$

by linearizing around measured values $u_{m,i}$ from a PR run. In equation (3-6), the Jacobian $\frac{\partial g_k}{\partial u} \bigg|_{u_{m,i},x_i,\theta}$ is a row vector with $N_{ij}$ columns, which contains derivatives of model predictions for the $k^{th}$ response variable with respect to each of the uncertain inputs and $\varepsilon_{Y_{ijk}}$ is the random noise.
for the measurements of the $k^{th}$ output obtained at the $j^{th}$ replicate of the $i^{th}$ target condition.

Taking variances of both sides of equation (3-6) gives:

$$
\sigma_{PR,ik}^2 \approx \left( \frac{\partial g_k}{\partial u} \bigg|_{u_{m,i},x_i,\theta} \right) \Sigma_U \left( \frac{\partial g_k}{\partial u} \bigg|_{u_{m,i},x_i,\theta} \right)^T + \sigma_{Y,ik}^2
$$

(3-7)

Assuming that uncertainties in measured inputs and outputs are independent. In equation (3-7), $\sigma_{PR,ik}^2$ is the variance of the $k^{th}$ model output due to variability in both inputs and outputs, whereas $\sigma_{Y,ik}^2$ is the measurement variance for the $k^{th}$ response variable.

In PR situations, $\hat{\sigma}_{Y,ik}^2$ in equation (3-4) provides an estimate of the overall output variance $\sigma_{PR,ik}^2$. Substituting this expression for $\sigma_{PR,ik}^2$ in equation (3-7), replacing $\theta$ with $\hat{\theta}$, and rearranging gives:

$$
\hat{\sigma}_{Y,ik}^2 = \frac{\sum_{j=1}^{n_i} (y_{m,ijk} - \tilde{y}_{m,ik})^2}{n_i - 1} - \left( \frac{\partial g_k}{\partial u} \bigg|_{u_{m,i},x_i,\hat{\theta}} \right) \Sigma_U \left( \frac{\partial g_k}{\partial u} \bigg|_{u_{m,i},x_i,\hat{\theta}} \right)^T \hat{\delta}_k \quad (3-8)
$$

In our experience, occasionally $\hat{\sigma}_{Y,ik}^2$ computed using equation (3-8) can be a negative number, especially if it is computed using a small number of pseudo-replicate runs, and uncertainties in the model inputs are relatively large. As a result, we recommend a cut-off value $\delta_k$ be used, based on the modeler’s knowledge about a reasonable size for the output measurement variance:

$$
\sigma_{Y,ik}^2 = \max \left( \frac{\sum_{j=1}^{n_i} (y_{m,ijk} - \tilde{y}_{m,ik})^2}{n_i - 1} - \left( \frac{\partial g_k}{\partial u} \bigg|_{u_{m,i},x_i,\hat{\theta}} \right) \Sigma_U \left( \frac{\partial g_k}{\partial u} \bigg|_{u_{m,i},x_i,\hat{\theta}} \right)^T, \delta_k \right) \quad (k = 1, ..., N_Y) \quad (3-9)
$$

Equation (3-5) can then be used to pool these output-measurement variance estimates, resulting in an appropriate value of $\hat{\sigma}_{Y,ik}^2$ for use in EVM parameter estimation.

Because parameter estimates $\hat{\theta}$ are required to calculate the Jacobian matrix $\frac{\partial g_k}{\partial u}$ in equation (3-9), a procedure that updates $\frac{\partial g_k}{\partial u} \bigg|_{u_{m,i},x_i,\hat{\theta}}$ based on the most recent estimate $\hat{\theta}$ is recommended.

Table 3-5 provides suggested steps for this iterative procedure. In Steps 2 and 3, initial estimates for the output variances and parameters are obtained by neglecting the influence of the uncertain
inputs. Steps 4 to 6 involve iterations that improve these estimates by taking the input uncertainties into account.

Table 3-5 Procedure for EVM parameter estimation using pseudo-replicate data

1- Assign appropriate cut-off values $\delta_k$ ($k = 1, ..., N_Y$) for each of the measured outputs variances. Set the step counter to $s=0$.
2- Obtain an initial guess for the output measurement variances using equations (3-4) and (3-5).
3- Obtain initial parameter estimates $\hat{\theta}^{(0)}$ using weighted-least squares parameter estimation.
4- For each PR target condition and measured output, calculate $\hat{\sigma}_{Y_{ik}}^2$ from equation (3-9) using the most recent parameter estimates $\hat{\theta}^{(s)}$.
5- Use equation (3-5) to pool the variance estimates obtained in step 4.
6- Use $J_{PR}$ in Table 3-4 to perform an EVM parameter estimation, using the $\hat{\sigma}_{Y_{ik}}^2$ as diagonal elements of $\hat{\Sigma}_Y$, resulting in updated parameter estimates $\hat{\theta}^{(s+1)}$.
7- Calculate the relative change in the parameter values $e = \sqrt{\sum_{p=1}^{N_\theta} \left( \theta_p^{(s+1)} - \theta_p^{(s)} \right)^2} / \theta_p^{(s)}$ where subscript $p$ denotes the $p^{th}$ element in $\hat{\theta}$. If $e$ is smaller than a tolerance set by the user, stop and report the parameter values. Otherwise, increase the value of $s$ by one and return to step 4.

3.5 Case Study: EVM Reactivity-ratio Estimation Using Mayo-Lewis Equation

The Mayo-Lewis equation describes the relationship between instantaneous comonomer composition in the reaction medium $f_1$ and the corresponding copolymer composition $F_1$:

$$ F_1 = \frac{r_1 f_1^2 + f_1 (1-f_1)}{r_1 f_1^2 + 2 f_1 (1-f_1) + r_2 (1-f_1)^2} \quad (3-10) $$

Many researchers have used the Mayo-Lewis equation to estimate reactivity ratios from low-conversion batch-reactor data (i.e., at conversions below 10%), where the initial monomer composition and cumulative copolymer composition are assumed to be similar to the instantaneous compositions $f_1$ and $F_1$ in equation (3-10), respectively.\textsuperscript{35,63} In equation (3-10), $F_1$ is the dependent
variable (i.e., \( y \) in equations (3-3.1) and (3-3.4)) and \( f_1 \) is the uncertain independent variable (i.e., \( u \) in equations (3-3.2), (3-3.3), and (3-3.5)). Reactivity ratios \( r_1 \) and \( r_2 \) are model parameters used to account for the relative rates of the four propagation reactions shown in Table 3-6, where \( R_1^* \) and \( R_2^* \) are growing polymer chains with terminal monomers \( M_1 \) and \( M_2 \), respectively:

\[
\begin{align*}
    r_1 &= \frac{k_{11}}{k_{12}} \\ \\
    r_2 &= \frac{k_{22}}{k_{21}}
\end{align*}
\]

\[ (3-11) \quad (3-12) \]

**Table 3-6** Important reactions corresponding to a free-radical copolymerization terminal model

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Rate Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_1^* + M_1 \xrightarrow{k_{11}} R_1^* )</td>
<td>(3-6.1)</td>
</tr>
<tr>
<td>( R_1^* + M_2 \xrightarrow{k_{12}} R_2^* )</td>
<td>(3-6.2)</td>
</tr>
<tr>
<td>( R_2^* + M_1 \xrightarrow{k_{21}} R_1^* )</td>
<td>(3-6.3)</td>
</tr>
<tr>
<td>( R_2^* + M_2 \xrightarrow{k_{22}} R_2^* )</td>
<td>(3-6.4)</td>
</tr>
</tbody>
</table>

In the current study, we use a Mayo-Lewis equation case study, because it has been widely used (see Table 3-1) in EVM-based parameter estimation problems to account for the uncertainties in input \( (f_1) \) measurements and uncertainties in the model outputs \( (F_1) \). We use copolymerization data provided by Ren et al., with n-butyl methacrylate as component 1 and n-butyl acrylate as component 2 to estimate reactivity ratios \( r_1 \) and \( r_2 \), which are the model parameters (i.e., elements of \( \theta \) in equations (3-3.1) and (3-3.4)). Data values obtained by Ren et al. are provided in Table 3-7. Their data set contains replicate data at two initial monomer feed compositions (i.e., \( f_1 = 0.487 \) and \( f_1 = 0.196 \), respectively). The information provided by Ren et al. about how they conducted their experiments is not sufficiently detailed so that we can ascertain whether their replicated experiments are true replicates or pseudo replicates. Nevertheless, because the reported measured inputs at each replicate condition are the same, each a TR1 or PR situation is plausible.
A TR2 situation would usually lead to different reported values of $f_1$ when replicate experiments are attempted.
Table 3-7 The experimental data collected for copolymerization of n-Butyl Methacrylate and n-Butyl Acrylate from Ren et al.²

<table>
<thead>
<tr>
<th>$f_1$</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.100</td>
<td>0.187</td>
</tr>
<tr>
<td>0.200</td>
<td>0.335</td>
</tr>
<tr>
<td>0.300</td>
<td>0.459</td>
</tr>
<tr>
<td>0.410</td>
<td>0.620</td>
</tr>
<tr>
<td>0.501</td>
<td>0.668</td>
</tr>
<tr>
<td>0.601</td>
<td>0.762</td>
</tr>
<tr>
<td>0.700</td>
<td>0.820</td>
</tr>
<tr>
<td>0.801</td>
<td>0.882</td>
</tr>
<tr>
<td>0.897</td>
<td>0.968</td>
</tr>
<tr>
<td>0.487</td>
<td>0.656</td>
</tr>
<tr>
<td>0.487</td>
<td>0.654</td>
</tr>
<tr>
<td>0.487</td>
<td>0.651</td>
</tr>
<tr>
<td>0.487</td>
<td>0.655</td>
</tr>
<tr>
<td>0.196</td>
<td>0.334</td>
</tr>
<tr>
<td>0.196</td>
<td>0.348</td>
</tr>
<tr>
<td>0.196</td>
<td>0.344</td>
</tr>
<tr>
<td>0.196</td>
<td>0.353</td>
</tr>
</tbody>
</table>

Based on the results in Table 3-7, Ren et al. might have prepared a mixture of comonomers and used it to run four copolymerization experiments at the corresponding replicate condition where $f_1 = 0.487$, resulting in situation TR1. Alternatively, they might have run four copolymerization experiments at that target replicate condition using four different monomer mixtures, after making some adjustments to ensure that the corresponding measured values of $f_1$ are the same, resulting in a PR situation. When estimating reactivity ratios from their data, Ren et al. assumed that the input and output measurement variances were both already known (AK). They did not explicitly estimate $\sigma_Y^2$ using their replicate experiments.

Below we demonstrate how to obtain output measurement variance estimates from the replicates provided by Ren et al. in two different ways: i) assuming that replicate data in Table 3-7. are true replicates of type 1, and ii) assuming that pseudo replicates were performed. In our analysis, we make the common assumption (used by Ren et al.) that random errors in the
measured comonomer feed composition and copolymer composition are multiplicative rather than
additive (i.e., the anticipated percentage error is constant, so that larger measured values tend to
have larger errors). Writing the model equation in terms of natural logarithms of the measured
inputs and responses results in additive random-error terms with constant variances. Model
equations (3-8.1) and (3-8.2) in Table 3-8 are used for situations where the output measurement
variance is assumed known (indicated by AK in the first column of Table 3-8). In this situation, we
do not need to pay attention to whether replicates have been conducted or not, and each pair of
\((f_1, F_1)\) values can be treated as if it were obtained at a distinct run condition. In Equations (3-8.1)
and (3-8.2) the index \(i\), where \(i = 1 \ldots 17\), is the index for the 17 experimental runs listed in Table
3-7. On the left-hand side of Equation (3-8.1), we use a capital \(L\) in the symbol \(L_{nF_1,i}\) to indicate
that the natural logarithm of the \(i^{th}\) measured output is a random variable due to the random
measurement error \(\varepsilon_{L_{nF_1,i}}\). On the right-hand side of Equation (3-8.1), lower-case symbols \(r_1, r_2\)
and \(f_{1,i}\) are used to indicate unknown true values of the reactivity ratios and the input comonomer
compositions that require estimation. Similarly, in Equation (3-8.2), \(Ln f_{1,i}\) is a random variable,
\(ln f_{1,i}\) on the right-hand side is the corresponding true value, and \(\varepsilon_{L_{nF_1,i}}\) is the random measurement
error.

The third row of Table 3-8 shows the EVM equations for the situation where TR1 replicates
are conducted. In this situation, there are 11 unique experimental conditions (i.e., \(i = 1, \ldots, 11\)).
Data from replicated runs are shown for the last two target conditions (i.e., \(n_{10} = n_{11} = 4\)). In
equation (3-8.3), \(L_{nF_1,ij}\) is the output for the \(j^{th}\) replicate at the \(i^{th}\) target condition.
Equation (3-8.2) shows the relationship between input measurements and their corresponding true
values in the TR1 situation. The difference between Equation (3-8.2) when it is used for the TR1
and AK situations is that, in the TR1 situation, the index \(i\) has the values from 1 to 11, whereas it
has the values from 1 to 17 in the AK situation. Equations (3-8.4) and (3-8.5) are used for the PR
situation.
Table 3-8 EVM Mayo-Lewis for the AK, TR1 and PR situations

<table>
<thead>
<tr>
<th>Case</th>
<th>EVM equation</th>
<th>Counter values</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK</td>
<td>( \ln F_{1,i} = \ln \left( \frac{r_1 f_1^2 + f_{1,i}(1-f_{1,i})}{r_1 f_1^2 + 2f_{1,i}(1-f_{1,i}) + r_2(1-f_{1,i})} \right) + e_{\ln F_{1,i}} ) (3-8.1)</td>
<td>( i = 1, \ldots, 17 )</td>
</tr>
<tr>
<td>TR1</td>
<td>( \ln F_{1,ij} = \ln \left( \frac{r_1 f_{1,ij}^2 + f_{1,ij}(1-f_{1,ij})}{r_1 f_{1,ij}^2 + 2f_{1,ij}(1-f_{1,ij}) + r_2(1-f_{1,ij})} \right) + e_{\ln F_{1,ij}} ) (3-8.3)</td>
<td>( i = 1, \ldots, 11, j = 1, \ldots, n_l )</td>
</tr>
<tr>
<td>PR</td>
<td>( \ln F_{1,ij} = \ln \left( \frac{r_1 f_{1,ij}^2 + f_{1,ij}(1-f_{1,ij})}{r_1 f_{1,ij}^2 + 2f_{1,ij}(1-f_{1,ij}) + r_2(1-f_{1,ij})} \right) + e_{\ln F_{1,ij}} ) (3-8.4)</td>
<td>( n_l = \begin{cases} 1 &amp; \text{if } 1 \leq i \leq 9 \ 4 &amp; \text{if } 10 \leq i \leq 11 \end{cases} )</td>
</tr>
</tbody>
</table>

The corresponding objective functions for the models in Table 3-8 are provided in Table 3-9. Equation (3-9.1) is the objective function when the output variance is assumed known, where \( F_{1m,i} \) and \( f_{1m,i} \) are measurements for copolymer composition and comonomer composition, respectively. A lower-case letter \( l \) is used in \( \ln F_{1m,i} \) and \( \ln f_{1m,i} \), because they are numerical values obtained from an experiment rather than random variables. Equations (3-9.2) and (3-9.3) are the corresponding objective functions for the TR1 and PR situations. In all of the objective functions, we assume a known variance for the inputs \( \sigma_{\ln F_{1}}^2 = 1 \times 10^{-4} \), which was reported by Ren et al.\(^2\)

In the AK situation, we also use their assumed output measurement variance \( \sigma_{\ln F_{1}}^2 = 2.5 \times 10^{-3} \).

For the TR1 and PR situations, we estimate the output measurement variances using the proposed methodology, obtaining the results shown in Table 3-10. The estimate of \( \hat{\sigma}_{\ln F_{1}}^2 \) for the PR situation is 47% smaller than the estimate for the TR1 situation. This result makes sense because, in the PR situation, some of the variation in the measured outputs is attributed to uncertainties in \( \ln f_{1,m} \). However, in the TR1 situation, all of the variation in the measured outputs for the true replicate runs is assigned to output measurement noise (and to common-cause variation not associated with \( f_1 \)).
Table 3-9 EVM-based objective functions for reactivity-ratio estimation corresponding to the AK, TR1 and PR situations

<table>
<thead>
<tr>
<th>Case</th>
<th>Objective function and decision variables requiring estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK</td>
<td>[ J_{AK} = \sum_{i=1}^{17} \left( \frac{(\ln F_{1,i} - \ln F_{1,1})^2}{\sigma^2_{\ln F_1}} + \frac{(\ln f_{1,i} - \ln f_{1,1})^2}{\sigma^2_{\ln f_1}} \right) ] (3-9.1)</td>
</tr>
<tr>
<td></td>
<td>Decision variables: ( \theta = [r_1, r_2]^T ) and ( u_i = \ln f_{1,i} ) (( i = 1, \ldots, 17 ))</td>
</tr>
<tr>
<td>TR1</td>
<td>[ J_{TR1} = \sum_{i=1}^{11} \sum_{j=1}^{n_i} \left( \frac{(\ln F_{1,ij} - \ln F_{1,1})^2}{\sigma^2_{\ln F_1}} + \sum_{i=1}^{11} \left( \frac{(\ln f_{1,ij} - \ln f_{1,1})^2}{\sigma^2_{\ln f_1}} \right) ] (3-9.2)</td>
</tr>
<tr>
<td></td>
<td>where ( n_i = \begin{cases} 1 &amp; \text{if } 1 \leq i \leq 9 \ 4 &amp; \text{if } 10 \leq i \leq 11 \end{cases} )</td>
</tr>
<tr>
<td></td>
<td>Decision variables: ( \theta = [r_1, r_2]^T ) and ( u_i = \ln f_{1,i} ) (( i = 1, \ldots, 11 ))</td>
</tr>
<tr>
<td>PR</td>
<td>[ J_{PR} = \sum_{i=1}^{11} \sum_{j=1}^{n_i} \left( \frac{(\ln F_{1,ij} - \ln F_{1,1})^2}{\sigma^2_{\ln F_1}} + \sum_{i=1}^{11} \sum_{j=1}^{n_i} \left( \frac{(\ln f_{1,ij} - \ln f_{1,1})^2}{\sigma^2_{\ln f_1}} \right) ] (3-9.3)</td>
</tr>
<tr>
<td></td>
<td>where ( n_i = \begin{cases} 1 &amp; \text{if } 1 \leq i \leq 9 \ 4 &amp; \text{if } 10 \leq i \leq 11 \end{cases} )</td>
</tr>
<tr>
<td></td>
<td>Decision variables: ( \theta = [r_1, r_2]^T ) and ( u_{ij} = \ln f_{1,ij} ) (( j = 1, \ldots, n_i ) and ( i = 1, \ldots, 11 ))</td>
</tr>
</tbody>
</table>

Estimated values of the parameters obtained using three different assumptions about uncertainties are shown in columns 3 and 4 of Table 3-10. The estimated values of the reactivity ratios for the different situations are similar. However, uncertainties in these parameter estimates are noticeably different. Figure 3-1a shows a linearization-based 95% JCR (the red ellipse) corresponding to the AK situation. Detailed information about how this JCR was computed is provided in Appendix C. We obtained the red ellipse by extending a methodology developed by Hamilton who obtained JCRs for important subsets of model parameters in nonlinear least-squares regression. While constructing these JCRs, Hamilton accounted for uncertainties in nuisance parameters that were not part of the important subsets. In our proposed method, described in Appendix C, the uncertain inputs \( f_1 \) are treated as nuisance parameters and \( r_1 \) and \( r_2 \) are important parameters. The JCR shown in Figure 3-1a is constructed using a chi-squared
distribution with 2 degrees of freedom (corresponding to the 2 important parameters being estimated).

When output measurement variances are not assumed known, it is appropriate to use an F distribution for computing the corresponding JCR. However, it is not obvious how a modeler should determine the denominator degrees of freedom for this F distribution (corresponding to the degrees of freedom for the estimated parameter covariance matrix). As a result, we opted to use an empirical bootstrapping technique to obtain information about parameter uncertainties. Information about the proposed bootstrapping algorithm is provided in Appendix C. In this method, we generate bootstrap composition measurements and use them to determine plausible reactivity-ratio estimates, which are shown by green dots in Figures 3-1a, 3-1b, and 3-1c. We validated the proposed bootstrapping method by comparing the results obtained for the AK situation with the red ellipse. Bootstrapping results are in good agreement with this analytical result, because only 4 out of 100 bootstrap estimates falls outside of the linearization-based 95% JCR. Using 10000 random seeds for bootstrapping resulted in 533 plausible parameter estimates falling outside of the red ellipse. We attribute the small deviation between 533/10000 and the theoretical value of 5.0% to nonlinearity of the model in equation (3-8.1).

Results obtained using the proposed bootstrap methods for the TR1 and PR situations are shown by the green dots in Figures 3-1b and Figure 3-1c, respectively, based on the corresponding estimates of $\hat{\sigma}_{\text{Lin}F_i}^2$ shown in Table 3-10. The resulting reactivity-ratio uncertainties are noticeably smaller than for the AK situation (the red ellipse for the AK situation is repeated in Figures 3-1b and 3-1c for comparison). It makes sense that cloud of green dots obtained using the PR assumption is smaller than the corresponding cloud for the TR1 assumption, because the variance estimate obtained using the PR assumption is smaller. These results reveal that experimentalists should be careful when they make assumptions about uncertainties and repeated experiments in EVM.
parameter estimation. These assumptions can have an important influence on the resulting parameter uncertainties.

Table 3-10 Estimated (assumed) values of measurement variance $\sigma_{LNF1}^2$ and estimated reactivity ratios corresponding to EVM situations with different underlying assumptions

<table>
<thead>
<tr>
<th>Case</th>
<th>Estimated (Assumed) output variance</th>
<th>$r_1$</th>
<th>$r_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK</td>
<td>$2.5 \times 10^{-3}$</td>
<td>2.0121</td>
<td>0.4615</td>
</tr>
<tr>
<td>TR1</td>
<td>$8.1 \times 10^{-5}$</td>
<td>2.0556</td>
<td>0.4749</td>
</tr>
<tr>
<td>PR</td>
<td>$4.3 \times 10^{-5}$</td>
<td>2.0081</td>
<td>0.4612</td>
</tr>
</tbody>
</table>
Figure 3-1 JC information for reactivity ratios obtained using data from Table 3-7 and different uncertainty assumptions. Different reactivity-ratio estimate, indicated by • are obtained for situations where a) the output measurement variance $\sigma^2_{\ln F_1} = 2.5 \times 10^{-3}$ is assumed known, b) the output measurement variance $\hat{\sigma}^2_{\ln F_1}$ is estimated from true replicate runs and c) the output measurement variance $\hat{\sigma}^2_{\ln F_1}$ is estimated assuming pseudo replicate runs. The red ellipse shown in all three plots is the linearization-based JCR obtained by assuming that $\sigma^2_{\ln F_1}$ is perfectly known. This ellipse is repeated in b) and c) as a helpful reference for comparison purposes. Green dots in a), b) and c) are plausible alternative reactivity-ratio estimates obtained using bootstrapping.
3.6 Conclusions

Conducting replicate experiments is complicated in error-in-variables (EVM) situations. Depending on how the experiments are performed, the true values of the uncertain inputs may or may not be the same for multiple runs conducted at the same target conditions. We call replicate experiments “true replicates”, when true values of uncertain inputs are assumed to be the same for all runs conducted at a target condition. When measured values of uncertain inputs are the same for repeated runs, but the underlying true inputs may be different, we call these experiments pseudo replicates.

Obtaining accurate estimates of output measurement variances is important in EVM parameter estimation because these variance estimates appear in the EVM objective function, thereby influencing the parameter estimates. It is relatively straightforward to obtain output measurement variance estimates from true replicates, but situations involving pseudo replicates are more complex. In the current study, a linearization-based approach is proposed to estimate output measurement variances in pseudo-replicate (PR) situations.

We use a Mayo-Lewis case study involving literature data for n-butyl methacrylate and n-butyl acrylate copolymerization to illustrate the proposed methodology. Reactivity-ratio parameters in the Mayo-Lewis equation are estimated using a data set containing replicate runs. We estimate the reactivity ratios making three different assumptions: i) output measurement variance is assumed known, ii) output measurement variance is unknown and replicate data are true replicates, and iii) output measurement variance is unknown and replicate data are pseudo replicates. We show similar point estimates for the reactivity ratios are obtained in all three situations.

To obtain information about parameter uncertainties, we extend a linearization-based method to obtain JCRs in situations where the output measurement variance is assumed known and we show that the parameter results are consistent with those from a proposed bootstrapping method. We develop variations for this bootstrapping technique to provide parameter uncertainty.
information for true-replicate and pseudo-replicate situations. The results of these bootstrapping calculations reveal that parameter uncertainty estimates are influenced to a great extent by the assumed type of replicates. In the current case study, uncertainties in the parameters are larger when true replicates are assumed than when pseudo replicates are assumed, because of the noticeable difference (by a factor of ~2) in the corresponding estimates for the output measurement variance.

In the future, the proposed methodologies in this study will be applied to more complex case studies. For example, a multi-input/multi-output model for hexane hydro-isomerization will be considered using real data and synthetic data.
3.7 References for Chapter 3


77


80
Chapter 4

Propagating Input Uncertainties into Parameter Uncertainties and Model Prediction Uncertainties - A Review‡

4.1 Abstract

A review of uncertainty quantification techniques is provided for a variety of situations involving uncertainties in model inputs (independent variables). The situations of interest are divided into three categories: i) when model prediction uncertainties are quantified based on uncertainties in uncertain inputs, ii) when parameter estimate uncertainties are calculated by propagation of uncertainties from measured inputs and outputs, and iii) when model prediction uncertainties are quantified based on corresponding uncertainties in measured inputs and uncertain parameter estimates. For all three situations, linearization-based and Monte Carlo based techniques are reviewed and details for their corresponding algorithms are presented. Recommendations are provided on which uncertainty quantification techniques are best for different types of chemical engineering models based on the amount of input uncertainty and nonlinearity over the range of plausible input and parameter values.

‡ This chapter has been submitted to Industrial and Engineering Chemistry Research with B. Celse and K. B. McAuley as coauthors
4.2 Introduction

Fundamental models are used for simulating the behavior of chemical processes and for designing new processes. The reliability of predictions from mathematical models needs to be assessed to ascertain uncertainties associated with model-based process operating and design decisions. Uncertainties in model predictions may arise from uncertainties in parameter estimates as well as from model inputs that are not perfectly known. The objective of the current chapter is to provide a literature review for three situations involving uncertain model inputs. First, we consider the situation where uncertainties in inputs are important, but uncertainties in model parameters can be neglected. Next, we consider the influence of input uncertainties on uncertainties on the estimates of model parameters. Finally, we review the literature on the joint influence of input uncertainties and parameter uncertainties on model predictions. For all three situations, we describe relatively simple literature methods based on model linearization and more-accurate methods that use Monte Carlo (MC) techniques to account for nonlinearity.

Uncertainties in inputs can arise from different sources. For example, the properties of process inflow streams (e.g., composition, flow rate, temperature) may be poorly known or may fluctuate. Other process operating conditions that are used as model inputs may also be uncertain. For example in many models of copolymerization systems, uncertain comonomer concentrations in the reactor are used as input variables for predicting the copolymer composition. Model developers and model users require reliable methods to quantify the influence of these uncertain inputs on model predictions and model parameters.

Specialized methods have been developed to account for uncertain inputs when estimating model parameters. These methods have been shown to give improved parameter estimates compared with traditional least-squares estimation methods in situations where input uncertainties are important compared with output uncertainties. A variety of techniques have been developed for propagating uncertainties in inputs and model parameters into prediction
uncertainties. However, to the best of our knowledge, there are no review articles in the chemical engineering literature that focus on uncertainty propagation and quantification. In the current review chapter, we describe available methods for handling input uncertainties and we provide advice on which linearization-based and MC methods are best in different situations. In addition, we make recommendations about how existing uncertainty propagation techniques can be extended to reliably account for simultaneous input and parameter uncertainties.

### 4.3 Propagating Input Uncertainties into Prediction Uncertainties

Consider the following multi-response model:

\[
Y_i = g(\theta, u_i, x_i) + \epsilon_{Y_i} \quad (4-1)
\]
\[
U_i = u_i + \epsilon_{U_i} \quad (4-2)
\]

where \( Y_i \in R^{N_Y} \) is a vector of \( N_Y \) random variables for responses obtained using experimental condition \( i \), \( u_i \in R^{N_U} \) is a vector of \( N_U \) unknown true values for uncertain model inputs used to conduct the \( i^{th} \) experimental run, \( x_i \) is a vector of perfectly-known inputs for the \( i^{th} \) run, and \( \epsilon_{Y_i} \) is a vector of random measurement errors, which are normally distributed with a covariance matrix \( \Sigma_Y \). The vector \( g(\theta, u_i, x_i) \) is the solution to the model equations that would be obtained from unknown true inputs \( u_i \) and model parameters \( \theta \in R^{N_\theta} \). In equation (4-2), \( U_i \in R^{N_U} \) is a vector of uncertain input measurements for the \( i^{th} \) run, and \( \epsilon_{U_i} \in R^{N_U} \) is the corresponding random noise vector, which belongs to a multivariate normal distribution with covariance matrix \( \Sigma_U \). We assume that the elements of \( \epsilon_{Y_i} \) and \( \epsilon_{U_i} \) are independent so that \( \Sigma_Y \) and \( \Sigma_U \) are diagonal.

In this section, we focus on prediction uncertainties for situations where uncertainties in parameters can be neglected, because they have a small influence compared to uncertainties in \( u \). Furthermore, we assume that estimates \( \hat{u} \) for the uncertain inputs are available from experimental data or engineering knowledge, along with estimates for the elements of \( \Sigma_U \). A model user may calculate model predictions:

\[
\hat{y}_k = g(\theta, \hat{u}_k, x_k) \quad (4-3)
\]
for an operating condition \( k \) of interest and want to quantify the uncertainties associated with the elements of \( \hat{y}_k \). Alternatively, a model user may want to obtain prediction uncertainties for future measurements to be obtained from a single new experiment conducted at condition \( k \):
\[
\hat{Y}_{sk} = g(\theta, \hat{u}_k, x_k) + \varepsilon_{Yk} \quad (4-4)
\]

The uncertainty associated with \( \hat{Y}_{sk} \) is larger than the uncertainty associated with \( \hat{y}_k \) in equation (4-3) because equation (4-4) accounts for additional measurement error \( \varepsilon_{Yk} \) that will influence the outcome of a new experiment. The uncertainty associated with \( \hat{y}_k \) in equation (4-3) is the uncertainty in the mean response that would be obtained from a very large number of runs with measured values of uncertain inputs given by \( \hat{u}_k \), whereas the uncertainty associated with \( \hat{Y}_{sk} \) in equation (4-4) accounts for both input uncertainty and output measurement uncertainty. To the best of our knowledge, distinguishing between these two types of prediction uncertainties has been overlooked in the chemical engineering literature that focuses on the effects of input uncertainties (even though popular undergraduate textbooks do talk about two types of prediction uncertainties when model parameters are uncertain but inputs are perfectly known). (e.g., 34, 35)

Table 4-1 provides a summary of chemical engineering studies where input uncertainties are important and parameter uncertainties are neglected when making model predictions. In all 22 of these studies, the authors seemed to be concerned with uncertainties in the mean prediction \( \hat{y}_k \) rather than \( \hat{Y}_{sk} \). As shown in Table 4-1, uncertainties in model predictions were obtained for models of a wide variety of chemical processes including crude oil distillation, crystallization of pharmaceuticals, wastewater treatment, a propylene glycol plant, and several combustion processes. In all studies in Table 4-1, MC-based techniques were used to obtain the prediction uncertainties, presumably because MC techniques are straightforward to use and relatively easy to explain. (36, 37) Basic steps to implement MC uncertainty analysis to obtain 95% confidence intervals for the elements of \( \hat{y}_k \) are summarized in Table 4-2. The algorithm in Table 4-2 assumes that the uncertainties in the unknown inputs are specified using normal distributions. However, it is
relatively straightforward to assume a different type of distribution for some of the unknown inputs.\textsuperscript{5-9} For example, if a uniform distribution is used for one of the unknown inputs, the modeler must specify a corresponding upper and lower bound rather than a variance. Most studies summarized in Table 4-1 consider normally-distributed uncertain inputs, but several consider inputs with uniform or beta distributions.\textsuperscript{8,9,38,39} The algorithm in Table 4-2 results in 95\% confidence intervals for the mean model predictions, but it is straightforward to make a small adjustment to obtain 90\% or 99\% confidence intervals or to obtain an empirical probability histogram (or boxplot) for the predicted outputs.

In situations where confidence intervals for the elements in $\hat{Y}_{sk}$ are desired, an additional step can be added to the algorithm in Table 4-2. After $\hat{y}_{k,j}^{(c)}$ is computed using equation (4-2.1), randomly-selected measurement errors $\varepsilon_{Y_{k,j}}^{(c)}$ (where $c = 1, \ldots, c_{\text{max}}$ for all values of $j=1,\ldots,N_Y$) should be sampled from the probability distribution for the corresponding measurement noise and added to the output predictions to obtain plausible measured responses:

$$\hat{y}_{sk,j}^{(c)} = \hat{y}_{k,j}^{(c)} + \varepsilon_{Y_{k,j}}^{(c)} \quad (4-5)$$

Confidence intervals, histograms or boxplots can then be constructed from the MC simulation results.

A benefit of MC methods is that they readily handle model nonlinearities as well as skewed distributions of uncertain inputs.\textsuperscript{40} The only disadvantage of the basic MC method is that a large number of random samples for the uncertain inputs can be required to obtain accurate and smooth probability distribution information for the model predictions and parameters.\textsuperscript{40-43} As a result, MC methods can be computationally intensive, especially for complex model equations that require significant time and effort to solve.\textsuperscript{43} A popular method that has been used to reduce the computational intensity is Latin hypercube sampling (LHS), which is often used instead of regular random sampling.\textsuperscript{5,37,44-48} Using LHS, a smaller number of MC simulations is typically needed to provide reliable confidence intervals and probability distributions.\textsuperscript{39,47} To understand how LHS is
implemented, consider a simple model with two uncertain inputs. Imagine that the measured value of the first input is $\hat{u}_{k,1} = 2.0$, and this uncertain input is assumed to be normally distributed with a standard deviation of 1.0 and that the second input has a measured value of $\hat{u}_{k,2} = 2.5$ and it belongs to a uniform distribution between 2.0 and 3.0. To generate five sets of candidate input values using LHS, the corresponding domains of the plausible inputs $u_{k,1}$ and $u_{k,2}$ are divided into five equally-probable intervals, as shown in Figure 4-1. Dividing the domains for the input values in this fashion results in 25 equal-probability boxes. LHS sampling is performed by first randomly selecting one of the five intervals corresponding to the first uncertain input on the horizontal axis, and one of the five intervals corresponding to the second uncertain input on the vertical axis, resulting in a randomly-selected box. Next, second intervals are randomly selected (without replacement) from among the remaining four intervals corresponding to each uncertain input $u_{k,1}$ and $u_{k,2}$, resulting in a second randomly-selected box. Selection continues until five boxes have been selected, as shown in Figure 4-1. Note that each row in Figure 4-1 contains one selected box and each column in Figure 4-1 contains one selected box. Next, input values are randomly chosen from each selected box and used to obtain model predictions. In this way, random samples are guaranteed from all parts of the corresponding distributions without taking a very large number of samples.
Figure 4-1 Latin hypercube sampling for two uncertain inputs where the first uncertain input is normally distributed $u_1 \sim N(2,1)$ and the second uncertain input belongs to a uniform distribution $u_2 \sim Uniform(2,3)$

Surrogate models have also been used to cope with the computationally intensive MC simulations involving fundamental models,\textsuperscript{7} especially models that are described by partial differential equations (PDEs).\textsuperscript{38,48,50-56} In this approach, an empirical surrogate model is fitted to a set of measurement-error-free fundamental-model responses obtained from a limited number of plausible uncertain inputs (e.g. 10 to 500).\textsuperscript{57,58} The fitted surrogated model is then used for model evaluation in step 2 of the algorithm in Table 4-2, instead of the complex original model. Polynomial chaos expansions are the most popular surrogate models used for prediction uncertainty quantification.\textsuperscript{48,52-55}
<table>
<thead>
<tr>
<th>Authors</th>
<th>System</th>
<th>Types of Uncertain Inputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sherer et al.</td>
<td>Sterilization process</td>
<td>-Timing of cell division and death</td>
</tr>
<tr>
<td>Zhang and Sahinidis</td>
<td>CO₂ Sequestration</td>
<td>-Porosity and permeability of the geological layers</td>
</tr>
<tr>
<td>Gel et al.</td>
<td>Fluidized bed</td>
<td>-Solid circulation rate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Superficial gas velocity</td>
</tr>
<tr>
<td>Gel et al.</td>
<td>Coal devolatilization in a gasifier</td>
<td>-Heating rate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Particle temperature</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Gasifier pressure</td>
</tr>
<tr>
<td>Sarangi et al.</td>
<td>Manifold microchannel heat sink.</td>
<td>-Width, depth, manifold depth, manifold inlet length, and manifold outlet length of the channels</td>
</tr>
<tr>
<td>Weber et al.</td>
<td>Rapid compression machine in combustion of methylcyclohexane and propene</td>
<td>-Initial and final pressure of the compression cycle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Initial temperature in the compression cycle</td>
</tr>
<tr>
<td>Duong et al. and Minh et al.</td>
<td>Propylene glycol production and lean dry gas processing plant</td>
<td>-Flow rates of the feeds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Temperature at a mixer outlet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Temperature at a cooler outlet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Pressure of inlet streams</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Pressure at a mixer outlet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Reactor temperature</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Reflux ratio in a distillation tower</td>
</tr>
<tr>
<td>Duong et al.</td>
<td>Lean dry gas sales production from crude natural gas</td>
<td>-Crude natural gas flow rate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Chiller heat removal duty</td>
</tr>
<tr>
<td>Sahraei et al.</td>
<td>Pilot-scale gasifier</td>
<td>-Volatile percentage of the fuel</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Solid particle diameters</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Angle of multiphase flow jet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Recirculation ratio</td>
</tr>
<tr>
<td>Zhang et al.</td>
<td>Moving-bed chromatography</td>
<td>-Flow rate</td>
</tr>
<tr>
<td>Minh et al.</td>
<td>Crude oil distillation and the corresponding revenue per unit of processes crude oil</td>
<td>-Crude oil feed flow rate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Heated crude oil temperature</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Flow rate of distillation products (i.e., naphtha, kerosene, diesel, and atmospheric gas oil)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Flow rate and temperature for the pump around streams</td>
</tr>
<tr>
<td>Authors</td>
<td>Field</td>
<td>Parameters</td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Shahaneh et al.</td>
<td>3D natural convection</td>
<td>Rayleigh and Prandtl numbers, Temperature at one of the boundaries</td>
</tr>
<tr>
<td>Mohammadi and Raisee et al.</td>
<td>Heat transfer in a gas turbine</td>
<td>Position of holes on the turbine blades and their corresponding diameters, Stream mean velocity, Film cooling hole length to diameter ratio, Blowing ratio, Streamwise and spanwise angle of hole film, Density ratio, Stream angle of attack</td>
</tr>
<tr>
<td>Öner et al.</td>
<td>Pharmaceutical crystallization</td>
<td>Mean of seed particle diameter distribution, Variance of seed particle diameter distribution, Seed particles mass, Cooling time, Cooling profile</td>
</tr>
<tr>
<td>Rajabi et al.</td>
<td>Heat and mass transfer in porous media</td>
<td>Rayleigh number, Diffusional buoyancy to thermal buoyancy ratio, Lewis number</td>
</tr>
<tr>
<td>Al et al.</td>
<td>Wastewater treatment</td>
<td>Surface flow rate of clarifier tank, Height of clarifier tank, Hydraulic retention time in aerobic tanks, Solid retention time in activated sludge, Ratio of anoxic tank volume to aerobic tank volume</td>
</tr>
<tr>
<td>Sterr et al.</td>
<td>Heat transfer in a microchannel heat sink</td>
<td>Surface roughness of a heat sink</td>
</tr>
<tr>
<td>Kim et al.</td>
<td>NOx emissions from waste materials incineration</td>
<td>Waste particle size, Waste mass ratio</td>
</tr>
</tbody>
</table>
Table 4-2 MC method for obtaining approximate 95% confidence intervals for mean model predictions at an experimental condition of interest

1- Specify experimental settings for inputs of interest (i.e., values of $x_k$ and $\hat{u}_k$) and the elements of $\Sigma_U$. Sample from the distribution of plausible true values $u_k$ for the uncertain inputs to obtain C candidate values for the uncertain input vector (i.e., $u_k^{(1)}$, $u_k^{(2)}$ ... $u_k^{(c_{\text{max}})}$), where $c_{\text{max}}$ is a large number (i.e., typically 1000 or 10000).

2- For $c=1$ to $c_{\text{max}}$, calculate the corresponding predicted values for the model outputs:

$$\hat{y}_{k,j}^{(c)} = g_j(\theta, u_k^{(c)}, x_k) \quad (j = 1, ..., N_Y) \quad (4-2.1)$$

3- Obtain the 2.5 and 97.5 percentile values of $\hat{y}_{k,j}^{(c)}$. These values are approximate bounds for the 95% confidence intervals for the predicted response $\hat{y}_{k,j}$.

An alternative to MC simulations would be to use a linearization-based method that reduces the computational requirements of the uncertainty quantification problem. In this technique, the model is linearized around the unknown true values of uncertain inputs using a Taylor series expansion:

$$\hat{y}_{k,j}(\theta, \hat{u}_k, x_k) \approx g_j(\theta, u_k, x_k) + \left. \frac{\partial g_j}{\partial u} \right|_{\hat{u}_k, x_k} (\hat{u}_k - u_k) \quad (4-6)$$

where $\left. \frac{\partial g_j}{\partial u} \right|_{\hat{u}_k, x_k}$ is a row vector containing derivatives of the model prediction for the $j^{th}$ response variable with respect to each of the uncertain inputs. Taking variances of both sides of equation (4-6), assuming that input uncertainties are independent, gives:

$$\hat{\sigma}_{\hat{y}_{k,j}}^2 \left( \left. \frac{\partial g_j}{\partial u} \right|_{\hat{u}_k, x_k} \right) \Sigma_U \left( \left. \frac{\partial g_j}{\partial u} \right|_{\hat{u}_k, x_k} \right)^T \quad (j = 1, ..., N_Y) \quad (4-7)$$

Use of equation (4-7) to estimate model prediction uncertainties is sometimes called the “delta method” when it is applied to situations involving uncertain model parameters rather than uncertain inputs. Symmetric 95% confidence intervals for the model predictions

$$\left[ \hat{y}_{k,j}(\theta, \hat{u}_k, x_k) \pm z_{0.025} \hat{\sigma}_{\hat{y}_{k,j}} \right]$$

are then obtained from normally-distributed input uncertainties.
where $z_{0.025}$ is the 97.5 percentile point of a standard normally-distributed random variable.\textsuperscript{34,40,65-67} Variance estimates associated with future experimental data are larger than $\hat{\sigma}_{y_{k,j}}^2$ due to the corresponding measurement uncertainty.\textsuperscript{68-70}

$$\hat{\sigma}_{y_{sk,j}}^2 = \hat{\sigma}_{y_{k,j}}^2 + \sigma_{\epsilon_{y_{j}}}^2 \quad (4-8)$$

where $\sigma_{\epsilon_{y_{j}}}^2$ is the variance of the random measurement noise for the $j^{th}$ response variable.

Use of linearization-based approaches has not been reported in uncertainty quantification studies involving only input uncertainties (see Table 4-1). This is probably because MC methods are easy to use, and modelers may have been concerned about significant nonlinearities in their model predictions. Note, however, that linearization-based uncertainty quantification is popular for assessing parameter uncertainties, as will be discussed in the next section.

In summary, MC-based or linearization-based techniques can be used to obtain model prediction uncertainties when model inputs are uncertain. MC-based simulation methods are computationally intensive, especially when the corresponding model equations are complex (e.g., involving PDEs or many ODEs). In this situation, Latin hypercube sampling can be used to efficiently reduce the required number of MC runs. Surrogate models (e.g., polynomial chaos expansions) can also be used instead of the original model for MC simulations to reduce computational effort. One benefit of MC-based techniques is that they can readily handle situations where input uncertainties belong to distributions that are not normal. Linearization-based methods for assessing prediction uncertainties require relatively low computational effort to obtain prediction uncertainties. In situations where uncertainties in model inputs are normally distributed and the model is only mildly nonlinear over the range of typical input values, we recommend the use of linearization-based uncertainty quantification for model predictions.

### 4.4 Propagating Input Uncertainties into Parameter Uncertainties

Parameter uncertainties in mathematical models arise from measurement noise associated with uncertain inputs and outputs during experimentation. Consider the case where variances of the
output measurements are perfectly known (i.e., $\Sigma_Y$ is diagonal with known elements). In this situation, maximum likelihood arguments can be used to obtain the corresponding EVM objective function for parameter estimation:

$$J_{EVM} = \sum_{i=1}^{N}(y_{m,i} - g(x_i, u_i, \theta))^T \Sigma^{-1}_Y \sum_{i=1}^{N}(y_{m,i} - g(x_i, u_i, \theta)) + (u_{m,i} - u_i)^T \Sigma^{-1}_U (u_{m,i} - u_i)$$  \hspace{1cm} (4-9)

where the subscript EVM indicates that an error-in-variables model like the one shown in equations (4-1) and (4-2) is being used.\textsuperscript{30} In equation (4-9), $y_{m,i}$ and $u_{m,i}$ are vectors of output and input measurements, respectively, for the $i^{th}$ run. Minimizing the objective function in equation (4-9) results in estimates for parameters $\hat{\theta}$ and uncertain inputs $\hat{u}_i$.

Table 4-3 shows the steps required to obtain uncertainties for the resulting parameter estimates using a parametric bootstrap approach (i.e., a MC approach for generating synthetic data).\textsuperscript{71} A large number ($b_{max}$) of plausible parameter sets is generated using steps 2 to 4, so that 95% confidence intervals for the parameters can be generated in step 5. Alternatively, the cloud of plausible parameter values can be used to construct empirical histograms, boxplots, or even joint confidence regions for the parameters.

This bootstrap uncertainty quantification technique relies on the assumption that $\hat{\theta}$ and $\hat{u}_i$ are close enough to the corresponding true values so that the method can accurately account for nonlinearity of the model equations. The algorithm in Table 4-3 assumes that the elements of covariance matrix $\Sigma_Y$ are perfectly known. If the output measurement variances are instead estimated from replicate or pseudo-replicate data, slightly different bootstrap algorithm is recommended.\textsuperscript{71}
Bootstrap steps to obtain parameter uncertainties when input uncertainties are important

1- Specify the covariance matrix of the output responses $\Sigma_Y$ and uncertain input measurements $\Sigma_U$. Use the real data to estimate parameters $\hat{\theta}$ and uncertain inputs $\hat{u}_i$ ($i = 1, \ldots, N$) based on the objective function in equation (4-9). Specify the number $b_{max}$ of bootstrap runs that will be used for assessing parameter uncertainty and assign the bootstrap counter $b = 1$.

2- Generate the $b^{th}$ set of synthetic data by treating the $\hat{\theta}$ and $\hat{u}_i$ ($i = 1, \ldots, N$) from step 1 as if they were true values:

$$y_{m,i}^{(b)} = g(x_i, \hat{u}_i, \hat{\theta}) + \varepsilon_{Y_i} \quad (i = 1, \ldots, N) \quad (4-3.1)$$

$$u_{m,i}^{(b)} = \hat{u}_i + \varepsilon_{U_i} \quad (4-3.2)$$

where

$$\varepsilon_{Y_i} \sim N(0, \Sigma_Y) \quad (4-3.3)$$

$$\varepsilon_{U_i} \sim N(0, \Sigma_U) \quad (4-3.4)$$

3- Substitute the synthetic data from equations (4-3.1) and (4-3.2) into the $J_{EVM}$ objective function:

$$J_{EVM}^{(b)} = \sum_{i=1}^{N} \left( y_{m,i}^{(b)} - g(x_i, u_i, \theta) \right)^T \Sigma_Y^{-1} \left( y_{m,i}^{(b)} - g(x_i, u_i, \theta) \right) + \left( u_{m,i}^{(b)} - u_i \right)^T \Sigma_U^{-1} \left( u_{m,i}^{(b)} - u_i \right) \quad (4-3.5)$$

and estimate plausible parameter estimates $\hat{\theta}^{(b)}$ corresponding to the bootstrap data.

4- If $b = b_{max}$, proceed to step 5. Otherwise, increase $b$ by one and go back to step (2).

5- Obtain 95% confidence intervals for the parameter estimates using 2.5% and 97.5% percentiles from the cloud of $b_{max}$ sets of bootstrap parameter estimates.

Table 4-4 shows a summary of previous studies where EVM parameters were estimated. Only 4 out of 37 studies shown in the table used bootstrap-based techniques to quantify uncertainties in model parameters. For example, Sutton and MacGregor estimated the parameters in a Wilson equation used for vapor-liquid equilibrium calculations by minimizing a simplified EVM objective function. In their study, they used a bootstrap technique to generate synthetic data.
experimental data and to evaluate the performance of their parameter-estimation method compared to several competing techniques. A related parameter-uncertainty characterization method that accounts for nonlinearity involves sampling from a posterior distribution of parameters during Bayesian EVM parameter estimation.32

We believe that all the studies in Table 4-4 relied on linearization-based methods to obtain information about parameter uncertainties. Many of the studies (i.e., 22 out of 37, which are denoted using * in the final column) did not provide details about how the reported parameter uncertainties were obtained. Based on the elliptical shapes of joint confidence regions and the symmetric confidence intervals reported, we believe that linearization-based methods were used. Three different linearization-based approaches were used to obtain uncertainties in the parameters. The first linearization-based method uses the Hessian of objective function $J_{EVM}$ to obtain the covariance matrix for the parameters and unknown inputs. However, details about how they obtained their Hessian matrices were not provided. We believe that the following method was used. Consider the augmented parameter vector:

$$ \theta_{aug} = [\theta^T \ u_1^T \ u_2^T \ ... \ u_N^T]^T $$  \hspace{1cm} (4-10)

which contains the parameters in the model $g$ and the uncertain inputs requiring estimation. The Hessian matrix of $J_{EVM}$ with respect to the elements of $\theta_{aug}$ is denoted by $H_{J_{EVM}}$ and can be obtained numerically, based on the estimated parameters $\hat{\theta}$ and estimated uncertain inputs $\hat{u}_i$. Each element in $H_{J_{EVM}}$ is a second derivative of $J_{EVM}$ with respect to the elements of $\theta_{aug}$ (see Appendix D for details). The inverse of the Hessian matrix can be used to obtain the covariance matrix for the estimated values of the elements of $\theta_{aug}$:

$$ \Sigma_{\theta_{aug}} = \left( \frac{1}{2} \hat{H}_{J_{EVM}} \right)^{-1} $$  \hspace{1cm} (4-11)

The corresponding covariance matrix $\Sigma_{\tilde{\theta}}$ for the parameter estimates appears as a block in the upper left corner of $\Sigma_{\theta_{aug}}$. 

94
The other option for determining the matrix $\Sigma_{\widehat{\theta}_{\text{aug}}}$ uses the Jacobian matrix of the scaled augmented model prediction vector $g_{\text{aug},s}$ with respect to augmented parameter vector $J_{g_{\text{aug},s}}$ rather than the Hessian:\textsuperscript{13, 21, 75}

$$\Sigma_{\widehat{\theta}_{\text{aug}}} = \left(J_{g_{\text{aug},s}}^T J_{g_{\text{aug},s}}\right)^{-1} \quad (4-12)$$

where the elements in $g_{\text{aug},s}$ are:

$$g_{\text{aug},s} = \begin{bmatrix} \frac{\partial g_1(u_1, x_1, \theta)}{\sigma_{EY_1}} & \frac{\partial g_2(u_1, x_1, \theta)}{\sigma_{EY_2}} & \ldots & \frac{\partial g_{NY}(u_1, x_1, \theta)}{\sigma_{EY_NY}} & \ldots & \frac{\partial g_N(u_N, x_N, \theta)}{\sigma_{EY_NY}} & \ldots & \frac{u_{11}}{\sigma_{EU_1}} & \frac{u_{12}}{\sigma_{EU_2}} & \ldots & \frac{u_{NN}}{\sigma_{EU_N}} \end{bmatrix} \quad (4-13)$$

In equation (4-13), $\sigma_{EY_j}$ and $\sigma_{EU_l}$ are the standard deviations for the $j^{th}$ uncertain model output and the $l^{th}$ uncertain input, respectively. Derivations for equations (4-11) and (4-12) and details about how to obtain the Jacobian matrix $g_{\text{aug},s}$ are provided in Appendix D. Equations (4-11) and (4-12) result in similar covariance matrices for the parameters.\textsuperscript{76} As a result, we recommend both methods and we suggest that modelers should make a decision about which to use, based on the optimizer utilized for their EVM parameter estimation. For example, if they are using ‘fmincon’ in MATLAB to minimize $J_{EVM}$, equation (4-11) is recommended because the Hessian matrix can be provided automatically as one of the outputs. Alternatively, if a modeler uses ‘lsqnonlin’ in MATLAB, equation (4-12) is recommended because ‘lsqnonlin’ provides the Jacobian information that can be used to obtain $J_{g_{\text{aug},s}}$.

To obtain a linearization-based covariance matrix for the parameters $\Sigma_{\widehat{\theta}}$, consider the parameter estimates with respect to the measurements of the uncertain inputs and outputs:\textsuperscript{77, 78}

$$\widehat{\theta} = h(y_{m1}, y_{m2}, \ldots, y_{mN}, u_{m1}, u_{m2}, \ldots, u_{mN}) \quad (4-14)$$

where $h \in \mathbb{R}^\theta$ is the result obtained by optimizing the objective function in equation (4-9). A Taylor-series expansion around the true values of uncertain model inputs and parameters can be used to propagate the uncertainties:

$$\widehat{\theta} \cong \theta + \sum_{i=1}^{N} \frac{\partial h}{\partial y_{mi}} \bigg|_{x_i, u_i, \theta} (y_{mi} - g(x_i, u_i, \theta)) + \sum_{i=1}^{N} \frac{\partial h}{\partial u_{mi}} \bigg|_{x_i, u_i, \theta} (u_{mi} - u_i) \quad (4-15)$$
where $\frac{\partial h}{\partial y_{m_i}}$ is a $N_\theta$-by-$N_f$ Jacobian matrix containing derivatives of the optimal parameter estimates with respect to the output measurements for the $i^{th}$ experimental run. Likewise, $\frac{\partial h}{\partial u_{m_i}}$ is a $N_\theta$-by-$N_U$ Jacobian matrix containing derivatives of the optimal parameter estimates with respect to the measured values of uncertain inputs. Taking variances of both sides of equation (4-15) and evaluating the derivatives at the estimated values of parameters $\hat{\theta}$ and inputs $\hat{u}_i$ gives:

$$
\Sigma\hat{\theta} = \sum_{i=1}^{N} \frac{\partial h}{\partial y_{m_i}} \Sigma_Y \left( \frac{\partial h}{\partial y_{m_i}} \right)^T + \sum_{i=1}^{N} \frac{\partial h}{\partial u_{m_i}} \Sigma_U \left( \frac{\partial h}{\partial u_{m_i}} \right)^T
$$

(4-16)

We do not recommend use of equation (4-16), especially when the parameter estimation is computationally intensive, because repeated parameter estimation is required to evaluate the elements of each column of Jacobian matrices $\frac{\partial h}{\partial y_{m_i}}$ and $\frac{\partial h}{\partial u_{m_i}}$ based on finite difference approximations. As a result, use of equation (4-16) is much more computationally demanding than use of equations (4-11) and (4-12).

Assuming that output and input measurement variances are perfectly known, the covariance matrices obtained from equations (4-11), (4-12), or (4-16) can be used to obtain 95% joint confidence regions for the parameters:

$$
(\theta - \hat{\theta})^T \Sigma\hat{\theta} (\theta - \hat{\theta}) \leq \chi^2(N_\theta, 0.95)
$$

(4-17)

where $\chi^2(N_\theta, 0.95)$ is the abscissa of a chi-squared distribution with $N_\theta$ degrees of freedom and 95% of the probability to the left. Diagonal elements in $\Sigma\hat{\theta}$ can be used to construct 95% confidence intervals for individual parameters. For example, for the $p^{th}$ parameter, the confidence interval is:

$$
\left[ \hat{\theta}_p \pm z_{0.025} \hat{\sigma}\hat{\theta}_p \right]
$$

(4-18)

where $\hat{\sigma}\hat{\theta}_p$ is the $p^{th}$ diagonal element of the estimated covariance matrix $\Sigma\hat{\theta}$. In situations where the output measurement variances are not known, the 95% joint confidence region is of the form:
\[
(\theta - \hat{\theta})^T \Sigma_{\hat{\theta}} (\theta - \hat{\theta}) \leq N_\theta F(N_\theta, \nu_\theta, 0.95) \quad (4-19)
\]

where \( F(N_\theta, \nu_\theta, 0.95) \) is the abscissa of a \( F \) distribution with \( N_\theta \) numerator degrees of freedom, \( \nu_\theta \) denominator degrees of freedom, and 95% of the probability to the left. In this situation, the 95% confidence interval for the \( t \) parameter is:

\[
\left[ \hat{\theta}_p \pm t_{\nu_\theta, 0.025} \hat{\sigma}_{\hat{\theta}_p} \right] \quad (4-20)
\]

where \( t_{\nu_\theta, 0.025} \) is the abscissa of a Student \( t \)-distribution with \( \nu_\theta \) degrees of freedom and with 97.5% of the probability to the left. Unfortunately, estimation of the appropriate degrees of freedom \( \nu_\theta \) is not straightforward, and it seems to be an unsolved problem in the literature. Our research group recently developed methods to estimate output measurement variances from replicate or pseudo-replicate experiments involving uncertain inputs. This methodology is helpful for estimation the elements of \( \Sigma_Y^{-1} \) for use in objective function (4-9). However, we do not know how to estimate the degrees of freedom \( \nu_\theta \) arising from the uncertain elements of \( \Sigma_Y^{-1} \) and the perfectly known elements of \( \Sigma_U^{-1} \).

In summary, bootstrap uncertainty quantification can be used to estimate parameter variances in situations with perfectly known measurement variances and situations where the measurement variances \( \Sigma_Y \) require estimation. Linearization-based methods are less computationally intensive, but provide uncertainty quantification results that are less reliable if the model nonlinearity is important over the typical range of the uncertain inputs and parameters.
### Table 4-4 Summary of studies involving parameter uncertainty quantification when input measurements are uncertain

<table>
<thead>
<tr>
<th>Authors</th>
<th>System</th>
<th>Uncertainty quantification technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sutton and MacGregor</td>
<td>Vapour-liquid equilibrium for ethanol-isooctane and benzene- heptane systems</td>
<td>Linearization-based JCR Bootstrap-based histograms for the parameters</td>
</tr>
<tr>
<td>Patino-Leal and Reilly</td>
<td>Vapour-liquid equilibrium for water/propanol system</td>
<td>Linearization-based JCR Bootstrap joint confidence</td>
</tr>
<tr>
<td>Duever et al.</td>
<td>Vapor-liquid equilibrium</td>
<td>Linearization-based JCR Bootstrap joint confidence</td>
</tr>
<tr>
<td>Valko and Vajda</td>
<td>Biological oxygen consumption and chemical oxygen consumption relationship in a wastewater Vapour-liquid equilibrium for methanol and 1,2-dichloroethane</td>
<td>Linearization-based CIs</td>
</tr>
<tr>
<td>Dube et al.</td>
<td>Styrene/butyl acrylate copolymerization</td>
<td>Linearization-based JCR</td>
</tr>
<tr>
<td>Keeler and Reilly</td>
<td>Copolymer composition calibration between IR and NMR measurements</td>
<td>Linearization-base JCR</td>
</tr>
<tr>
<td>Kim et al.</td>
<td>Water-gas shift reaction</td>
<td>Linearization-based CIs</td>
</tr>
<tr>
<td>Brar et al.</td>
<td>Styrene/butyl acrylate copolymerization</td>
<td>Linearization-based CIs and JCR</td>
</tr>
<tr>
<td>Brar and Charan</td>
<td>Vinyl acetate/methyl acrylate copolymerization Vinyl acetate/ethyl methacrylate copolymerization</td>
<td>Linearization-based CIs</td>
</tr>
<tr>
<td>Schoonbrood et al.</td>
<td>Styrene/2-Hydroxyethyl methacrylate copolymerization</td>
<td>Linearization-based CIs and JCR</td>
</tr>
<tr>
<td>Brar et al.</td>
<td>Acrylonitrile/methacrylic acid copolymerization</td>
<td>Linearization-based CI and JCR</td>
</tr>
<tr>
<td>Brar and Malhorta</td>
<td>Vinilidene chloride/methyl acrylate copolymerization</td>
<td>Linearization-based JCR</td>
</tr>
<tr>
<td>McManus and Penlidis; Scott and Penlidis</td>
<td>Styrene/ethyl acrylate copolymeration</td>
<td>Linearization-based JCR</td>
</tr>
<tr>
<td>Suddaby et al.</td>
<td>Methyl methacrylate/n-butyl methacrylate copolymerization</td>
<td>Linearization-based JCR</td>
</tr>
<tr>
<td>Brar et al.</td>
<td>Acrylonitrile/glycidyl methacrylate copolymerization</td>
<td>Linearization-based JCR</td>
</tr>
<tr>
<td>Hakim et al.</td>
<td>Butyl acrylate/methyl methacrylate copolymerization</td>
<td>Linearization-based JCR</td>
</tr>
<tr>
<td>Baradie et al.</td>
<td>Tetrafluoroethylene/ vinyl acetate, chloror trifluoroethylene/ vinyl acetate, and vinylidene fluoride/ vinyl acetate copolymerization</td>
<td>Linearization-based JCR and CI</td>
</tr>
<tr>
<td>Faber et al.</td>
<td>Vapor-liquid equilibrium of methanol/ 1,2-dichloroethane</td>
<td>Linearization-based covariance matrix</td>
</tr>
<tr>
<td>Zavala and Biegler</td>
<td>Low-density polyethylene tubular reactor</td>
<td>Linearization-based JCR</td>
</tr>
<tr>
<td>Kazemi et al.</td>
<td>Acrylonitile/methyl acrylate copolymerization</td>
<td>Linearization-based JCR</td>
</tr>
<tr>
<td>Hajipour and Satyro</td>
<td>A correlation for predicting the critical properties of a variety of hydrocarbons</td>
<td>Linearization-based JCR</td>
</tr>
<tr>
<td>Zhang and Dube</td>
<td>N-Butyl Methacrylate/d-limonene copolymerization</td>
<td>Linearization-based JCR</td>
</tr>
</tbody>
</table>
4.5 Simultaneous Propagation of Input and Parameter Uncertainties into Prediction Uncertainties

Model predictions are sometimes obtained from uncertain parameters and uncertain inputs. Consider predictions for a multi-response model, which are desired at an arbitrary target condition \( k \), where data may or may not be available:

\[
\hat{y}_k = g(\hat{\theta}, \hat{u}_k, x_k) \tag{4-21}
\]

Uncertainties for the parameter estimates might be available in two different forms: i) they might be available as a cloud of plausible parameter estimates obtained from bootstrap simulations (e.g., obtained using the method in Table 4-3), or ii) they might be available from a probability density function (e.g., parameter estimates may belong to a multivariate normal distribution with a specified covariance matrix). Table 4-5 shows bootstrap steps to obtain prediction uncertainties in the first situation where a cloud of plausible parameter estimates is available and the uncertain inputs belong to a specified probability distribution. In steps 1 to 3, plausible values for uncertain parameters and inputs are sampled and corresponding model predictions are obtained. In step 4, the
corresponding 95% confidence intervals are constructed for the mean response of the model at the operating condition of interest. Alternatively, in the second situation where there is a specified mathematical distribution for the model parameters (e.g., parameter estimates belong to a multivariate normal distribution), random samples from this probability distribution are obtained in step 2 of the algorithm. The information obtained from steps 1 to 3 can be used to construct box plots or uncertainty histograms for the model predictions if this information is preferred compared to 95% confidence intervals.

The algorithm in Table 4-5 generates uncertainty information for predictions of the mean responses at the experimental condition of interest. If the modeler wants instead to obtain uncertainty information for a single new experiment that might be conducted at the condition of interest, equation (4-5) should be used to generate a cloud of plausible values for the future measurements.
Table 4-5 MC method for obtaining approximate 95% confidence intervals for mean model predictions at an experimental condition of interest

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Specify experimental settings for inputs of interest (i.e., values of ( \mathbf{x}_k ) and ( \hat{\mathbf{u}}_k )) and the elements of ( \Sigma_U ). Sample from the distribution of plausible true values ( \mathbf{u}<em>k ) for the uncertain inputs to obtain ( c</em>{max} ) candidate values for the uncertain input vector (i.e., ( \mathbf{u}_k^{(1)} ), ( \mathbf{u}<em>k^{(2)} ), ..., ( \mathbf{u}<em>k^{(c</em>{max})} )), where ( c</em>{max} ) is a large number.</td>
</tr>
<tr>
<td>2</td>
<td>Use the cloud of parameters from a previous bootstrap (Monte Carlo) uncertainty quantification (( \hat{\boldsymbol{\theta}}^{(1)}, \hat{\boldsymbol{\theta}}^{(2)}, ..., \hat{\boldsymbol{\theta}}^{(c_{max})} )). Alternatively, if a mathematical description of the parameter uncertainty distribution is available, sample ( c_{max} ) times from this distribution to obtain plausible values of parameters (i.e., ( \hat{\boldsymbol{\theta}}^{(1)}, \hat{\boldsymbol{\theta}}^{(2)}, ..., \hat{\boldsymbol{\theta}}^{(c_{max})} )).</td>
</tr>
<tr>
<td>3</td>
<td>For ( b = 1 ) to ( c_{max} ), calculate the corresponding predicted values for the model outputs: ( \hat{y}_{k,j}^{(c)} = g_j(\hat{\boldsymbol{\theta}}^{(c)}, \mathbf{u}_k^{(c)}, \mathbf{x}_k) ) (( j = 1, ..., N_Y )) (4-5.1)</td>
</tr>
<tr>
<td>4</td>
<td>Obtain the 2.5 and 97.5 percentile values of ( \hat{y}_{k,j}^{(c)} ). These values are approximate bounds for the 95% confidence intervals for the mean predicted response.</td>
</tr>
</tbody>
</table>

Table 4-6 shows a summary of previous chemical engineering studies involving prediction uncertainty quantification involving uncertain inputs and parameters. As shown in the final column, the majority of these studies relied on MC methods. The studies in Table 4-6 are concerned with a wide variety of processes such as distillation, benzene production, bioethanol production, and pharmaceutical production. In some of these studies, surrogate models such as polynomial expansion or Kriging were used instead of the original model \( g \) to decrease the computational intensities of the MC-based techniques.\(^{90-97}\)

In 6 out of 39 studies, linearization-based approaches were used. Using a Taylor-series expansion, the uncertainties in parameter estimates and uncertain model inputs can be propagated to obtain model prediction uncertainties:
\[ \hat{\sigma}_{y_{k,j}}^2 = \left( \frac{\partial g}{\partial \bar{u} \mid \hat{\theta}, \hat{u}_k, x_k} \right) \Sigma_U \left( \frac{\partial g}{\partial \bar{u} \mid \hat{\theta}, \hat{u}_k, x_k} \right)^T + \left( \frac{\partial g}{\partial \hat{\theta} \mid \hat{\theta}, \hat{u}_k, x_k} \right) \Sigma_{\hat{\theta}} \left( \frac{\partial g}{\partial \hat{\theta} \mid \hat{\theta}, \hat{u}_k, x_k} \right)^T \] (j = 1, ..., \( N_Y \)) \tag{4-22}

Assuming the elements in \( \Sigma_U \) are perfectly known and that parameter estimates are normally distributed, approximate symmetric 95\% confidence intervals for mean predictions are

\[ \left[ \hat{y}_{k,j} \left( \hat{\theta}, \hat{u}_k, x_k \right) \pm z_{0.025} \hat{\sigma}_{y_{k,j}} \right] . \]

Variance estimates for predictions of a new experimental data value \( \hat{y}_{sk,j} \) (j = 1, ..., \( N_Y \)) are obtained by adding the variance estimates in equation (4-22) and the corresponding random measurement variances (see equation (4-8)).

In summary, MC-based and linearization-based techniques have been used to obtain prediction uncertainties from uncertain parameter estimates and model inputs. Bootstrap-based techniques are more computationally demanding, but they can handle a variety of situations where either empirical distributions are available for the parameters or mathematical distributions are available. Linearization-based techniques require less computational effort than MC methods, but are only available for situations where the multivariate normal distributions for the inputs and parameters are assumed.

**Table 4-6** Summary of studies involving prediction uncertainty quantification from parameters and inputs uncertainties

<table>
<thead>
<tr>
<th>Authors</th>
<th>Process</th>
<th>Uncertainty quantification technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atherton et al. 98</td>
<td>A first-order reaction in a batch reactor</td>
<td>Linearization</td>
</tr>
<tr>
<td>Tørvi and Hertzberg 99</td>
<td>Batch distillation Column</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Turányi et al. 100</td>
<td>Methane combustion</td>
<td>Linearization-based</td>
</tr>
<tr>
<td>Zádor et al. 101</td>
<td>Methane combustion</td>
<td>Monte Carlo and linearization</td>
</tr>
<tr>
<td>Zsély et al. 102</td>
<td>H2/air and wet CO/air combustion</td>
<td>Monte Carlo and linearization</td>
</tr>
<tr>
<td>Zádor et al. 103</td>
<td>Methane combustion</td>
<td>Monte Carlo and linearization</td>
</tr>
<tr>
<td>Fischer et al. 104</td>
<td>Acetylsalicylic acid (aspirin) production</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Zsély et al. 105</td>
<td>NO production during methane combustion</td>
<td>Monte Carlo and linearization</td>
</tr>
<tr>
<td>Authors</td>
<td>Topic</td>
<td>Method</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Chen et al.</td>
<td>Benzene production from hydrodealkylation of toluene</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Hajipour and Satyro</td>
<td>A correlation for predicting the critical properties of a variety of hydrocarbons</td>
<td>Linearization</td>
</tr>
<tr>
<td>Morales-Rodriguez et al.</td>
<td>Bioethanol production</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Lane et al.</td>
<td>A fluidized bed containing horizontal tubes</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Rasoulian and Ricardez-Sandoval</td>
<td>Thin film deposition process</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Donato and Pitchumani</td>
<td>A three-dimensional circulating fluidized bed and a one-dimensional two-phase turbulent flow</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Weber et al.</td>
<td>Combustion of methylcyclohexane and propene</td>
<td>Linearization</td>
</tr>
<tr>
<td>Wang et al.</td>
<td>Heat conduction in a thermal plate</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Soepyan et al.</td>
<td>Hydroloc transports of solid particles in conduits</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Frutiger et al.</td>
<td>Rankine Cycle</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Xu et al.</td>
<td>Pilot-scale carbon capture system</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Fajraoui et al.</td>
<td>Natural convection in porous media</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Ahmadian Behrooz and Hosseini</td>
<td>Benzene-toluene distillation column</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Frutiger et al.</td>
<td>Heat pump</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Morgan et al.</td>
<td>Solvent-based CO₂ capture</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Mondejar et al.</td>
<td>Rankine cycle</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Jones et al.</td>
<td>Heat recovery of organic Rankine cycle</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Neba et al.</td>
<td>Biogas plants</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Fenila and Shastri</td>
<td>Enzymatic hydrolysis of lignocellulosic biomass</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Kimaev et al.</td>
<td>Chemical vapor deposition</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>Son and Du</td>
<td>Fed-batch penicillin manufacturing process</td>
<td>Monte Carlo</td>
</tr>
</tbody>
</table>
4.6 Conclusion and Discussion

A literature review is presented on methods that have been used for quantifying uncertainties in model predictions and parameter estimates in chemical process models where uncertainties in model inputs are important. Available algorithms are described for three different situations: i) when a model user wants to quantify prediction uncertainties that arise predominantly from model input uncertainties rather than parameter uncertainties; ii) when a model user wants to characterize parameter uncertainties that arise from a combination of uncertain input and output measurements, and iii) when a model user wants to quantify prediction uncertainties arising from uncertain inputs and uncertain parameters. Linearization-based techniques and MC-based techniques are described for all three situations and recommendations are made concerning which methods are best to use. In all situations, linearization-based techniques require less computational effort, but MC-based techniques have fewer simplifying assumptions (i.e., MC-based techniques are more accurate if enough simulations are performed).

We recommend modelers to choose between these two approaches based on the extent of the uncertainties in their uncertain model inputs and on the degree of nonlinearity in their model. In situations where input uncertainties are relatively small, the use of linearization-based uncertainty quantitation techniques is recommended. In situations where input uncertainties are larger and the model is highly nonlinear over the range of uncertain inputs, the use of linearization-based techniques may result in unrealistic quantification of uncertainties.

Further, MC-based techniques are attractive for assessing prediction and parameter uncertainties when a model user is uncertain about the appropriate degrees of freedom for the corresponding parameter estimates. In this situation, linearization-based techniques can only be used to obtain approximate parameter covariance matrices and approximate covariances for the model predictions. Lack of knowledge about degrees of freedom makes it difficult to construct accurate joint confidence regions and confidence intervals for parameters and model predictions.
Degrees of freedom are not an issue for MC-based techniques that construct confidence intervals based on a large number of plausible values for parameters or predictions. The only drawback of MC-based techniques is that computational effort can be excessive when model evaluations are time-consuming.

When the uncertainties in the model predictions are calculated, model users should determine the type of model predictions whose corresponding uncertainties are desired. The uncertainties involved in prediction of a single output measurement are larger than uncertainties in the mean prediction at the same conditions, because of random noise in future output measurements. In the chemical engineering literature, attention has been focused only on quantifying uncertainties in the mean response. We recommend that model users should carefully decide between these two types of prediction uncertainties, based on their needs and goals.

When uncertainties in the model predictions are quantified based on the uncertainties in parameters and inputs, the uncertainties in the parameters are often obtained from a previous parameter estimation study. In this situation, the modeler assumes that the estimated parameters are relevant to the new situation. However, in some situations a modeler is confronted with a data set based on uncertain inputs, where parameter uncertainties are unknown. In such situations, the modeler must estimate the parameters and their corresponding uncertainties for using the available model and data set. After the parameters and parameter uncertainties have been estimated, the modeler can propagate the parameter uncertainties and input uncertainties to obtain uncertainties in the model predictions. To our knowledge, only two research groups have considered this combined EVM parameter estimation and prediction uncertainty quantification problem. Both groups considered error propagation in simple models with a single response (dependent) variable. In a future study, we will extend the available methods so they can be used for characterizing prediction uncertainties for multi-response models.
4.7 References for Chapter 4


61. Minh, L. Q.; Duong, P. L. T.; Goncalves, J.; Kwok, E.; Lee, M., A two-stage approach of multiplicative dimensional reduction and polynomial chaos for global sensitivity analysis and uncertainty quantification with a large number of process
uncertainties. *Journal of the Taiwan Institute of Chemical Engineers* 2017, 78, 254-264.
66. Preacher, K. J.; Selig, J. P. J. C. M.; Measures, Advantages of Monte Carlo confidence intervals for indirect effects. 2012, 6, (2), 77-98.


Chapter 5

Parameter Estimation and Prediction Uncertainties for Multi-response Kinetic Models with Uncertain Inputs§

5.1 Abstract

Error-in-variables model (EVM) methods are used for parameter estimation when independent variables are uncertain. During EVM parameter estimation, output measurement variances are required as weighting factors in the objective function. These variances can be estimated based on data from replicate experiments. However, conducting replicates is complicated when independent variables are uncertain. Instead, pseudo-replicate runs may be performed where the target values of inputs for repeated runs are the same, but the true input values may be different. Here, we propose a method to estimate output-measurement variances for use in multivariate EVM estimation problems, based on pseudo-replicate data. We also propose a bootstrap technique for quantifying uncertainties in resulting parameter estimates and model predictions. The methods are illustrated using a case study involving n-hexane hydroisomerization in a well-mixed reactor. Case-study results reveal that assumptions about input uncertainties can have important influences on parameter estimates, model predictions and their confidence intervals.

§ This chapter has been submitted to AIChE Journal with B. Celse and K. B. McAuley as coauthors
5.2 Introduction

Fundamental kinetic models are important for chemical process development and improvement. These models usually contain kinetic parameters and activation energies, which are estimated using conventional parameter estimation techniques such as weighted least squares (WLS). In WLS parameter estimation, the independent variables (model inputs) are assumed to be perfectly known and measurement errors are considered for the dependent variables (model outputs). However, in some experimental situations, input errors are relatively large compared to output measurement errors, which can result in inaccurate parameter estimates. Error-in-variables model (EVM) parameter estimation methods have been developed to address this issue by accounting for input uncertainties during parameter estimation. EVM objective functions for multi-response models are typically of the form:

\[ J_{EVM} = \sum_{i=1}^{N} \sum_{k=1}^{N_{Y}} \left( \frac{(y_{m,ik} - g_{k}(x_{i},u_{i},\theta))}{\sigma_{y_{ik}}^2} \right)^2 + \sum_{l=1}^{N} \sum_{i=1}^{N_{U}} \left( \frac{(u_{m,il} - u_{il})}{\sigma_{u_{il}}^2} \right)^2 \]  

(5-1)

where \( y_{m,ik} \) is the \( k^{th} \) \((k = 1, ..., N_{Y})\) measured output response obtained using the \( i^{th} \) target condition, \( g_{k}(x_{i},u_{i},\theta) \) is the corresponding model solution obtained from the perfectly known model inputs \( x_{i} \in \mathbb{R}^{N_{x}} \), the imperfectly-known inputs \( u_{i} \in \mathbb{R}^{N_{U}} \) and the model parameters \( \theta \in \mathbb{R}^{N_{\theta}} \). In equation (5-1), \( \sigma_{y_{ik}}^2 \) is the measurement variance for the \( k^{th} \) model output and \( \sigma_{u_{il}}^2 \) is the measurement variance for the \( l^{th} \) \((l = 1, ..., N_{U})\) uncertain input. In the second term, \( u_{m,il} \) is the measured value of the \( l^{th} \) uncertain input at the \( i^{th} \) experimental setting and \( u_{il} \) is the \( l^{th} \) element in the corresponding vector \( u_{i} \) of true values for the uncertain inputs.

During EVM parameter estimation, parameters \( \theta \) and uncertain inputs \( u_{i} \) for each experimental run are estimated by minimizing \( J_{EVM} \). EVM parameter estimation is more complicated than WLS because EVM requires additional information about the uncertainties in inputs (i.e., values \( \sigma_{u_{il}}^2 \)) and EVM involves additional decision variables in the optimization problem. Variances \( \sigma_{u_{il}}^2 \) in process inputs are usually assumed known, based on the information that
experimentalists have regarding the reproducibility of inputs (e.g., based on variations involved in analytical devices). In the chemical engineering literature, output measurement variances are also usually assumed to be known a priori in most EVM parameter estimation studies. However, modelers can also use replicate data to estimate the output measurement variances $\sigma^2_k$. The resulting variance estimates should account for the measurement variability for the output responses as well as all other sources of experimental variations (e.g., ambient condition and sample collection) except variability associated with $u_i$.

Conducting replicate experiments in EVM situations is not straight-forward because uncertain inputs should be kept constant during different replicate runs. In some situations, it is possible to perform true replicate (TR) experiments (e.g., when the same uncertain stock feed solution is used to perform multiple experiments). In other situations involving pseudo replicate (PR) experiments, the true values of uncertain inputs cannot be enforced to be the same during attempted replicate runs. In TR situations, quantifying output measurement variances $\sigma^2_k$ is relatively straightforward (e.g., using pooled variance estimates from several different experimental conditions). Estimation of output measurement variances from pseudo-replicate data is more difficult because the contribution of uncertain inputs to the overall variability in the outputs should be taken into account. In a previous study, we devised a method to estimate $\sigma^2_k$ based on pseudo-replicates, but we only considered a simple single-response copolymerization model. One of the objectives in the current chapter is to extend this method for use in multi-response models. We use a n-hexane hydroisomerization model as a case study to illustrate the proposed methodology. Further, we extend a bootstrap technique developed for parameter uncertainty quantification in single-response models so that it can be used for multi-response models and we demonstrate the technique using the case study. We also show how to obtain model prediction uncertainties based on the bootstrap parameter uncertainties, input uncertainties and estimated output uncertainties.
5.3 Background

5.3.1 Using Pseudo-Replicate Experiments to Estimate Parameters and Obtain Output Measurement Variances

Consider a situation where an experimenter attempts to perform $n_i$ replicate experiments at the $i^{th}$ target run condition where the desired flow rate of one species is $5.0 \frac{mmol}{s}$ and the desired flow rate of a second species is $0.10 \frac{mmol}{s}$. Assume that both of the flow rates contain significant uncertainty and that the experimenter makes adjustments to valves until the corresponding measured flow rates are $u_{mi} = [5.0 \frac{mmol}{s}, 0.10 \frac{mmol}{s}]^T$. If attempted replicate experiments are conducted on different days, where the measured values are the same on each day, but the underlying true flow rates are different, then the experimenter performs pseudo replicates.

In general, for a multi-response system with several uncertain inputs, an error-in-variables model of the process can be expressed as:

\[ Y_{ij} = g(x_i, u_{ij}, \theta) + \varepsilon_{Yij} \tag{5-2} \]
\[ u_{mi} = u_{ij} + \varepsilon_{Uij} \tag{5-3} \]

where $Y_{ij} \in \mathbb{R}^{N_Y}$ is a vector of uncertain measurements for $N_Y$ dependent variables obtained from the $j^{th}$ ($j = 1, ..., n_i$) pseudo-replicate experiment conducted using the $i^{th}$ target condition. A capital $Y$ is used to indicate that $Y_{ij}$ is a random variable due to the measurement noise $\varepsilon_{Yij}$ on the right-hand side of equation (5-2). Model predictions in vector $g(x_i, u_{ij}, \theta)$ depend on the model parameters $\theta$, the true values of the uncertain inputs $u_{ij}$ and any perfectly-known inputs $x_i$. In equations (5-2) and (5-3), $\varepsilon_{Yij}$ and $\varepsilon_{Uij}$ are vectors of normally-distributed random noise for the output and input measurements, respectively. Note that $\varepsilon_{Yij}$ accounts mainly for measurement error in the $j^{th}$ dependent variable, as well as additional variability in the experiments which is not
attributed to \( \varepsilon_{U ij} \). For simplicity, throughout the remainder of this chapter we will refer to \( \varepsilon_{Y ij} \) as output measurement uncertainty.

Assuming that input and output measurement noise vectors in equations (5-2) and (5-3) are independent, an EVM objective function can be obtained from maximum-likelihood arguments for situations involving pseudo-replicate experiments:

\[
J_{EVM} = \sum_{i=1}^{N_i} \sum_{j=1}^{n_i} \left[ (y_{m ij} - g(x_i, u_{ij}, \theta))^T \Sigma_{Y}^{-1}(y_{m ij} - g(x_i, u_{ij}, \theta)) + (u_m - u_{ij})^T \Sigma_{U}^{-1}(u_m - u_{ij}) \right]
\]

where \( y_{m ij} \) is the vector of measured values for the dependent variables from the \( j \)th run conducted at the \( i \)th experimental target condition. In the current study, we assume that \( \Sigma_Y \) and \( \Sigma_U \) the covariance matrices for random errors in the dependent and independent variables, respectively, are diagonal. Diagonal elements in \( \Sigma_U \) (i.e., \( \sigma^2_{U l} \) where \( l = 1, \ldots, N_U \)) are variances associated with the measurement or implementation of the uncertain inputs, which are assumed to be known based on the experimenter’s experience with the experimental setup. The diagonal elements in \( \Sigma_Y \) (i.e., \( \sigma^2_{Y k} \) where \( k = 1, \ldots, N_Y \)) are unknown and need to be estimated from the pseudo-replicate data. In a previous study, we developed a linearization-based approach to obtain estimates of output measurement variances for each of the \( N_Y \) dependent variables based on the pseudo-replicate data obtained at the \( i \)th target condition. For the \( k \)th dependent variable:

\[
\hat{\sigma}^2_{Y ik} = \frac{\sum_{j=1}^{n_i} (y_{m ij} - \bar{y}_{m ik})^2}{n_i - 1} - \left( \frac{\partial g_k(x_i, u, \theta)}{\partial u}_{u_m, x_i, \theta} \right)^T \Sigma_U \left( \frac{\partial g_k(x_i, u, \theta)}{\partial u}_{u_m, x_i, \theta} \right)
\]

where \( y_{m ij} \) is the measured value of the \( k \)th dependent variable obtained from the \( j \)th pseudo-replicate run at the \( i \)th experimental condition and \( \bar{y}_{m ik} \) is the average of the \( n_i \) measured values for these runs. The second term on the right-hand side of equation (5-5) accounts for the contribution of uncertain inputs on the overall variability of the dependent variables. Sometimes, \( \hat{\sigma}^2_{Y ik} \) obtained from equation (5-5) can be negative, so a cut-off value \( \delta_k \) is used to obtain reliable variance estimates, based on the modeler’s prior knowledge:
\[ \hat{\sigma}^2_{ik} = \max \left( \frac{\sum_{j=1}^{n_i} (y_{m, jk} - \bar{y}_{m, ik})^2}{n_i - 1} \right) = \left( \frac{\partial g_k(x_i u \theta)}{\partial u} \right)_{u_m = x_i \hat{\theta}}^T \Sigma_u \left( \frac{\partial g_k(x_i u \theta)}{\partial u} \right)_{u_m = x_i \hat{\theta}} \delta_k \] (5-6)

The estimated variances obtained from equation (5-6) can be pooled together to obtain:

\[ \hat{\sigma}^2_{ik} = \frac{\sum_{i=1}^{N} (n_i - 1) \hat{\sigma}^2_{ik}}{\sum_{i=1}^{N} (n_i - 1)} \quad (k = 1, ..., N_Y) \] (5-7)

Note that, the parameter estimates \( \hat{\theta} \) are required in equation (5-6) to calculate the Jacobian matrix \( \frac{\partial g_k(x_i u \theta)}{\partial u} \) \( u_m = x_i \hat{\theta} \). So, an iterative approach was developed to update the elements of \( \frac{\partial g_k(x_i u \theta)}{\partial u} \) as improved parameter estimates become available (see algorithm in Table 5-1). This method is used to obtain an estimate of \( \Sigma_Y \) in the hexane hydroisomerization case study in section 5.5.2.

**Table 5-1 Procedure for EVM Parameter Estimation Using Pseudo-Replicate Data**

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Assign appropriate cut-off values ( \delta_k ) ( (k = 1, ..., N_Y) ) for each of the measured outputs variances. Set the step counter to ( s=0 ).</td>
</tr>
<tr>
<td>2</td>
<td>Obtain an initial guess for the output measurement variances by treating the attempted replicates as true replicates (i.e., using ( \hat{\sigma}^2_{ik} = \frac{\sum_{j=1}^{n_i} (y_{m, jk} - \bar{y}_{m, ik})^2}{n_i - 1} ) to calculate variance in each experimental setting) and using equation (5-7) to pool variance estimates.</td>
</tr>
<tr>
<td>3</td>
<td>Obtain initial parameter estimates ( \hat{\theta}^{(0)} ) using weighted-least squares parameter estimation.</td>
</tr>
<tr>
<td>4</td>
<td>For each PR target condition and measured output, calculate ( \hat{\sigma}^2_{ik} ) from Equation (5-6) using the most-recent parameter estimates ( \hat{\theta}^{(s)} ).</td>
</tr>
<tr>
<td>5</td>
<td>Use equation (5-7) to pool the variance estimates obtained in step 4.</td>
</tr>
<tr>
<td>6</td>
<td>Use ( J_{EVM} ) in equation (5-4) to perform an EVM parameter estimation, using the ( \hat{\sigma}^2_{ik} ) as diagonal elements of ( \Sigma_Y ), resulting in updated parameter estimates ( \hat{\theta}^{(s+1)} ).</td>
</tr>
<tr>
<td>7</td>
<td>Calculate the relative change in the parameter values ( e = \sqrt{\sum_{p=1}^{N_\theta} \left( \frac{\hat{\theta}_p^{(s+1)} - \hat{\theta}_p^{(s)}}{\hat{\theta}_p^{(s)}} \right)^2} ) where subscript ( p ) denotes the ( p^{th} ) element in ( \hat{\theta} ). If ( e ) is smaller than a tolerance set by the user, stop and report the parameter values. Otherwise, increase the value of ( s ) by one and return to step 4.</td>
</tr>
</tbody>
</table>
5.3.2 Parameter Uncertainty Quantification Using Bootstrapping for a Single-response Model with a Single Uncertain Input

In a previous study, we extended a parametric bootstrap technique for models with perfectly-known independent variables so that it can be used for parameter uncertainty quantification in EVM. Synthetic data sets are generated multiple times, using typical input and output uncertainties, and parameters are estimated for each synthetic data set. The resulting cloud of bootstrap parameter estimates is then used for quantifying uncertainties in parameter estimates. Previously, we used this bootstrap method to quantify uncertainties in estimated reactivity-ratio parameters in a single-response Mayo-Lewis model with a single uncertain input. In section 5.4.1, this bootstrapping methodology is extended for use in multi-response models with multiple uncertain inputs.

An alternative approach to bootstrapping would be to sample from a posterior distribution for parameter estimates obtained using an Bayesian EVM approach based on Markov-Chain Monte Carlo simulations. In the current chapter, we decided to extend the bootstrapping approach instead, because our hexane hydroisomerization case study does not involve prior information about parameter values that should be imposed during estimation. Also, we believe that the proposed bootstrapping approach will be simpler for model developers to use than Bayesian methods.

5.3.3 Quantifying Prediction Uncertainties based on Uncertainties in Inputs and Parameter Estimates

In a previous study, we reviewed methods for quantifying prediction uncertainties when inputs are uncertain. Linearization-based and Monte-Carlo-based (MC-based) techniques are typically used for this purpose. MC-based methods are usually more accurate, but they require computationally intensive simulations when the model equations are complex. Nevertheless, they are relatively straight-forward and can readily handle situations where a cloud of parameter estimates is available from bootstrap computations. Past researchers have sampled from the distribution of the parameters and inputs simultaneously to propagate the corresponding
uncertainties into the model predictions. Table 5-2 provides an algorithm for quantifying uncertainties in model predictions using a cloud of $b_{max}$ parameter estimates available from bootstrapping. This algorithm quantifies the uncertainties in a vector of model predictions $\hat{y}_c$ at an experimental condition of interest $c$. The results from Table 5-2 provide 95% confidence intervals for noise-free values of the dependent variables at the condition of interest. If the modeler is interested instead in a 95% confidence interval on noisy measured values that might be obtained from a future experiment, then additive measurement errors in the dependent variables must also be considered. Both types of prediction uncertainties are computed in the case study in section 5-5-3.

Table 5-2: MC method for obtaining approximate 95% confidence intervals for mean model predictions at an experimental condition of interest

1. Specify experimental settings for inputs of interest (i.e., values of $x_c$ and $u_{c,m}$) and the elements of $\Sigma_U$. Sample from the distribution of plausible true values $u_c$ for the uncertain inputs to obtain $b_{max}$ candidate values for the uncertain input vector (i.e., $u_c^{(1)}$, $u_c^{(2)}$ ... $u_c^{(b_{max})}$), where $b_{max}$ is a large number.

2. Use the cloud of parameters from a previous bootstrap (Monte Carlo) uncertainty quantification ($\hat{\theta}^{(1)}$, $\hat{\theta}^{(2)}$, ..., $\hat{\theta}^{(b_{max})}$).

3. For $b = 1$ to $b_{max}$, calculate the corresponding predicted values for the model outputs:

$$\hat{y}_{c,k}^{(b)} = g_k(\hat{\theta}^{(b)}, u_c^{(b)}, x_c) \quad (k = 1, ..., N_Y) \quad (5-2.1)$$

4. Obtain the 2.5 and 97.5 percentile values of $\hat{y}_{c,k}^{(b)}$. These values are approximate bounds for the 95% confidence intervals for the mean predicted responses.

5.4 Proposed Methodology

5.4.1 Bootstrap Method for Quantifying Parameter Uncertainties in Multi-response EVM

In the current study, we extend a bootstrap algorithm for quantifying parameter uncertainties in single-response models so that it can be used for multi-response models with several uncertain inputs. Steps required for implementing the extended algorithm are shown in
In the first step, parameters and elements of the output covariance matrix $\Sigma_Y$ are estimated using the algorithm in Table 5-1. In steps 2 and 3, synthetic data are generated and their corresponding bootstrap parameters are estimated. In step 4, uncertainty information is extracted from the cloud of bootstrap parameter estimates.

Table 5-3 Bootstrap steps to obtain parameter uncertainties for multi-response error-in-variable models with several uncertain inputs using pseudo-replicate experiments

1- Estimate the model parameters $\hat{\theta}$ and the diagonal output measurement covariance matrix $\hat{\Sigma}_Y$ using the algorithm in Table 5-1. Assign bootstrap counter $b=1$ and set the desired number of bootstrap runs $b_{max}$.

2- Generate synthetic bootstrap data using $\hat{\theta}$ and $\hat{\Sigma}_Y$:

$$u_{ij}^{(b)} = u_{m_i} - \epsilon_{u_{ij}}^{(b)}$$

for $i = 1, ..., N; j = 1, ..., n_i$ where $\epsilon_{u_{ij}}^{(b)} \sim N(0, \Sigma_U)$

$$y_{mij}^{(b)} = g(x_i, u_{ij}^{(b)}, \hat{\theta}) + \epsilon_{y_{ij}}^{(b)}$$

for $i = 1, ..., N; j = 1, ..., n_i$ where $\epsilon_{y_{ij}}^{(b)} \sim N(0, \Sigma_Y)$

3- Obtain bootstrap parameter estimates $\hat{\theta}(b)$ and covariance matrix $\hat{\Sigma}_Y(b)$ using the $b^{th}$ synthetic data set. If $b = b_{max}$ go to step 4; otherwise increase $b$ by one and go to step 2.

4- Use the bootstrap parameter estimates $\hat{\theta}(b) (b = 1, 2, ..., b_{max})$ to construct confidence intervals or box plots for the parameters.

5.4.2 Bootstrap Method for Quantifying Prediction Uncertainties from Error-in-Variables Models

After parameter uncertainties are obtained using the multivariate bootstrap technique in Table 5-3, a cloud of $b_{max}$ parameter estimates is available. Using this cloud of parameter estimates, it is straightforward to obtain the corresponding cloud of model predictions using the algorithm in Table 5-2. Note that the algorithm in Table 5-2 provides an estimate of uncertainty in the mean predicted responses at the target condition of interest. In section 5-5-3, we will demonstrate how to also obtain the uncertainty associated with measured responses for a single future experiment. Uncertainty associated with a possible measured value from a single new
experiment is larger than the corresponding uncertainty in the mean response due to the associated measurement error.

5.5 Case Study: Catalytic n-Hexane Hydroisomerization in a Well-mixed Reactor

5.5.1 Model Equations and Experimental System

Here we present a multi-response model developed by Toch et al. to describe hydroisomerization of hexane (nC₆) in a catalytic well-mixed continuous reactor. The reaction scheme for this process is shown in Table 5-4. Products 2-methyl pentane (2MP) and 3-methyl pentane (3MP) are produced along with propane (C₃) as an undesirable side product. The corresponding algebraic equations obtained from mole balances are shown in Table 5-5. Derivations are provided in Appendix E. In Table 5-5, $F^0_q$ and $F_q$ are the molar inflow and outflow rates for species $q$ ($q = H_2, nC_6, 2MP, 3MP$ and $C_3$), respectively, $W$ is the weight of catalyst (i.e., H-ZSM-5) and $K_{phys}$ is the equilibrium constant for physisorption of n-hexane on the catalyst. The rate coefficients $k_r$ ($r = 1, 2, 3$) and the equilibrium constant $K_{phys}$ are temperature-dependant and are calculated using the equations in Table 5-6 where $k_{ref, r}$ is the rate coefficient of the $r^{th}$ reaction at the reference temperature $T_{ref} = 531.48$ K and $K_{ref}$ is the corresponding equilibrium coefficient. In equations (5-6.1) and (5-6.2), $E_{ar}$ is the activation energy of the $r^{th}$ reaction and $\Delta H^0$ is the enthalpy of physisorption.

As shown in the second column of Table 5-6, there are 8 unknown parameters

$$\theta = \begin{bmatrix} k_{ref, 1}, k_{ref, 2}, k_{ref, 3}, K_{ref}, E_{a, 1}, E_{a, 2}, E_{a, 3}, \Delta H^0 \end{bmatrix}^T$$

requiring estimation. When the values of the parameters and the experimental inputs (i.e., reactor pressure $P$, reactor temperature $T$, catalyst weight $W$ and flow rates $F^0_{nC_6}$ and $F^0_{H_2}$) are given, the three equations in three unknowns in Table 5-5 can be solved numerically to predict $F_{2MP}$, $F_{3MP}$ and $F_{C3}$. After predictions for the outputs are obtained, yield can be calculated:
\[ \text{Yield} = \frac{(F_{2MP} + F_{3MP})}{F_{nC_6}} \quad (5-8) \]

and selectivity for different species can also be calculated:

\[ S_q = \frac{\alpha_q F_q}{F_{nC_6}} \quad \left( q = 2MP, 3MP \text{ and } C_3 \right) \quad (5-9) \]

where \( \alpha_q \) is the number of carbon atoms in the species molecules (i.e., \( \alpha_{2MP} = \alpha_{3MP} = 6 \) and \( \alpha_{C_3} = 3 \)).

**Table 5-4 Reaction Scheme Proposed for nC_6 Hydroisomerization**

<table>
<thead>
<tr>
<th>Reaction Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>nC_6 ( \rightarrow ) 2MP</td>
</tr>
<tr>
<td>nC_6 ( \rightarrow ) 3MP</td>
</tr>
<tr>
<td>2MP + H_2 ( \rightarrow ) 2 C_3</td>
</tr>
</tbody>
</table>

**Table 5-5 Algebraic equations based on the mole balance equations for different species**

<table>
<thead>
<tr>
<th>Species</th>
<th>Corresponding Mole Balance Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2MP</td>
<td>[ F_{2MP} - \frac{k_1 (\frac{F_{nC_6}}{F_{2MP}} - F_{3MP} - \frac{F_{C_3}}{2}) - k_3 F_{2MP}}{1 + K_{phys} P (\frac{F_{nC_6}}{F_{H_2}} - \frac{F_{C_3}}{2}) + F_{H_2} + F_{C_6}} ] ( W = 0 \quad (5-5.1) )</td>
</tr>
<tr>
<td>3MP</td>
<td>[ F_{3MP} - \frac{k_2 (\frac{F_{nC_6}}{F_{2MP}} - F_{3MP} - \frac{F_{C_3}}{2})}{1 + K_{phys} P (\frac{F_{nC_6}}{F_{H_2}} - \frac{F_{C_3}}{2}) + F_{H_2} + F_{C_6}} ] ( W = 0 \quad (5-5.2) )</td>
</tr>
<tr>
<td>C_3</td>
<td>[ F_{C_3} - \frac{2k_3 \frac{F_{2MP}}{F_{H_2}} - \frac{F_{C_3}}{2}}{1 + K_{phys} P (\frac{F_{nC_6}}{F_{H_2}} - \frac{F_{C_3}}{2}) + F_{H_2} + F_{C_6}} ] ( W = 0 \quad (5-5.3) )</td>
</tr>
</tbody>
</table>
Table 5-6 Kinetic rate and equilibrium constant equations for solving balance equations and the corresponding parameters requiring estimation

<table>
<thead>
<tr>
<th>Kinetic rate and equilibrium constant equation</th>
<th>Parameters Requiring Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_r = k_{ref_r} e^{-\frac{E_a r}{R} \left( \frac{1}{T_{ref}} - \frac{1}{T} \right)}$  (r = 1,2,3) [(5-6.1)]</td>
<td>$k_{ref_1}, k_{ref_2}, k_{ref_3}, E_{a_1}, E_{a_2}, E_{a_3}$</td>
</tr>
<tr>
<td>$K_{phys} = K_{ref} e^{-\frac{\Delta H^0}{R} \left( \frac{1}{T_{ref}} - \frac{1}{T} \right)}$</td>
<td>$K_{ref}$ and $\Delta H^0$</td>
</tr>
</tbody>
</table>

Toch et al. performed 36 experimental runs at a variety of conditions to estimate the model parameters, assuming that all inputs were perfectly known (see Table E.1 in Appendix E). In the current case study, we consider two uncertain inputs \(u = [F_{H_2}^0, F_{nC_6}^0]^T\) and three perfectly-known inputs \(x = [P, T, W]^T\). We assume that the replicate experiments reported by Toch et al. are pseudo-replicates because slightly different true values of inputs $F_{H_2}^0$ and $F_{nC_6}^0$ would have been used when experiments were repeated. For the 36 experimental runs, there are 25 unique experimental conditions and 11 repeated runs conducted at a variety of different conditions. During each run, Toch et al. measured three outputs (i.e., measured flow rates for 2MP, 3MP and $C_3$). As shown in Table 5-7, three different levels of input uncertainties are considered, which are 1% and 5% and 10% of the corresponding average measured input values in Table E.1. Corresponding measurement standard deviations are shown in the second and the third columns. In the current case study, we use the data in Table E.1 multiple times, considering different levels of uncertainties in independent variables $F_{H_2}^0$ and $F_{nC_6}^0$ that could have been encountered by Toch et al. during their experiments.
### Table 5-7: Different uncertainty level for inputs with measurement uncertainties

<table>
<thead>
<tr>
<th>Levels of input uncertainties</th>
<th>Standard deviation for measurements of $F_{H_2}^0$ ($\sigma_{F_{H_2}^0}$)</th>
<th>Standard deviation for measurements of $F_{nC_6}^0$ ($\sigma_{F_{nC_6}^0}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>$1.63 \times 10^{-5}$ mol/s</td>
<td>$2.56 \times 10^{-7}$ mol/s</td>
</tr>
<tr>
<td>5%</td>
<td>$8.16 \times 10^{-5}$ mol/s</td>
<td>$1.28 \times 10^{-6}$ mol/s</td>
</tr>
<tr>
<td>10%</td>
<td>$1.63 \times 10^{-4}$ mol/s</td>
<td>$2.56 \times 10^{-6}$ mol/s</td>
</tr>
</tbody>
</table>

#### 5.5.2 EVM Parameter and Variance Estimation

The EVM objective function and decision variables for estimating the case-study parameters are shown in the first row of Table 5-8. The vector of model predictions $g(x_i, u_{ij}, \theta)$ obtained from the $j^{th}$ pseudo-replicate run at the $i^{th}$ target experimental condition has three elements:

$$g(x_i, u_{ij}, \theta) = [F_{2MP}(x_i, u_{ij}, \theta), F_{3MP}(x_i, u_{ij}, \theta), F_{C3}(x_i, u_{ij}, \theta)]^T$$ (5-10)

where $F_{2MP}, F_{3MP}$ and $F_{C3}$ are predicted outflow rates obtained by solving the three equations in three unknowns in Table 5-5. As shown in Table 5-8, the vector $x_i$ contains the perfectly-known inputs (i.e., temperature, pressure and catalyst weight) and $u_{ij}$ contains the uncertain inputs (i.e., inflow rates of hydrogen and n-hexane). In equation (5-8.1), the vector $y_{m,ij} = [y_{2MP,ij}, y_{3MP,ij}, y_{C3,ij}]^T$ contains measured data for the dependent variables and the vector $u_{m,i} = [u_{mC_6,i}, u_{mH_2,i}]^T$ contains measured data for the uncertain independent variables.

In this study, we compare the results of the EVM parameter estimation, assuming several different levels of input uncertainties, with results obtained from WLS parameter estimation using objective function (5-8.2). The WLS parameter estimation problem is considerably simpler than the EVM problem because it involves fewer decision variables (i.e., 8 compared to 44).
Further, the estimation of the output measurement variances is simpler when using WLS. In WLS situations, the attempted replicates are assumed to be true replicates, from which the modeler can readily obtain pooled estimates of output measurement variances. In EVM situations considered in this case study, output measurement variances and parameters should be estimated simultaneously using the algorithm in Table 5-1. Note the that the vector of perfectly-known inputs $\mathbf{x}_i$ in the EVM case have three elements for each experimental setting, but $\mathbf{x}_i$ in the WLS situation has 5 elements because all the inputs are assumed to be perfectly known.

The results of the parameter estimation and the output measurement variance estimation for WLS and EVM situations are shown in Table 5-9. The variance estimates are provided in the $2^{nd}$ to the $5^{th}$ rows and the parameter estimates in the $6^{th}$ to $14^{th}$ rows. As shown in the Table, estimated values of the variances and parameters obtained using WLS are similar to corresponding values from EVM with 1% input uncertainties, as expected. As such, in situations where input uncertainties are small relative to output uncertainties, we recommend that modelers use WLS rather than EVM because: i) WLS objective functions are simpler than EVM objective functions; ii) WLS estimation involves fewer decision variables than EVM; and iii) output measurement variances are simpler to estimate using the WLS assumption of true replicates, compared to the EVM assumption of pseudo replicates. Table 5-9 shows that, when larger input uncertainties are assumed (i.e., 5% and 10% of typical input values) the estimated parameters and output measurement variances are noticeably different using EVM compared to WLS. As expected, the estimated output measurement variances become smaller (see the $2^{nd}$ to the $5^{th}$ rows in Table 5-9) for increasing levels of input uncertainties, because a higher amount of variability in measured outputs in pseudo-replicate runs is attributed to the input uncertainties. Note that, the output measurement variances reported in Table 5-9 were computed using equations (5-6) and (5-7). We used cut-off values of $\delta_{F_{2MP}} = 1.44 \times 10^{-17}$, $\delta_{F_{3MP}} = 6.84 \times 10^{-18}$ and $\delta_{F_{C3}} = 1.94 \times 10^{-19}$ in the first step of the algorithm in Table 5-1. These cut-off values are 0.01% of the pooled variance.
estimates for the dependent variables obtained from equation (5-7) when input uncertainties are neglected and all variability in the measurement values of the outputs in attempted replicate run is attributed to $\varepsilon_{Yij}$.

Figure 5-1 shows the contribution of input uncertainties $\varepsilon_{Uij}$ and output uncertainties $\varepsilon_{Yij}$ to the overall uncertainty in $F_{2MP}$. The heights of the red bars in Figure 5-1 were obtained from $\hat{\sigma}^2_{F_{2MP}U}/\hat{\sigma}^2_{F_{2MP}o}$ where the overall variance in the $F_{2MP}$ responses was estimated from the pseudo-replicate data:

$$\hat{\sigma}^2_{F_{2MP}o} = \frac{\sum_{i=1}^{N_i} \sum_{j=1}^{n_i} (y_{2MPij} - \bar{y}_{2MPi})^2}{\sum_{i=1}^{N_i}(n_i-1)}$$  \hspace{1cm} (5-11)

and the input contribution to this overall variance was estimated from:

$$\hat{\sigma}^2_{F_{2MP}U} = \frac{\sum_{i=1}^{N_i}(n_i-1) \left( \frac{\partial F_{2MP}(x_i u_i \theta)}{\partial u} \right) \sum_{i=1}^{N_i}(n_i-1) \left( \frac{\partial F_{2MP}(x_i u_i \theta)}{\partial u} \right)^T}{\sum_{i=1}^{N_i}(n_i-1)}$$ \hspace{1cm} (5-12)

In equation (5-11), $\bar{y}_{2MPi}$ is the average measured value for $F_{2MP}$ obtained from data collected at the $i^{th}$ target condition. As shown in Figure 5-1, assigning a typical 1% error in the uncertain inputs results in pseudo-replicate experiments where nearly all the variability is due to $\varepsilon_{Yij}$ rather than $\varepsilon_{Uij}$. Conversely, when 10% typical input uncertainty is assumed, most of the variability in the responses arises from the input uncertainties.
Table 5-8 The objective function and their corresponding decision variables for EVM and WLS objective functions

\[
J_{\text{EVM}} = \sum_{i=1}^{N} \sum_{j=1}^{n_i} \left[ (y_{m\ ij} - g(x_i, u_{ij}, \theta))^T \Sigma^{-1}_Y (y_{m\ ij} - g(x_i, u_{ij}, \theta)) + (u_{m\ i} - u_{ij})^T \Sigma^{-1}_U (u_{m\ i} - u_{ij}) \right]
\]

\[
= \sum_{i=1}^{25} \sum_{j=1}^{n_i} \left[ \frac{(y_{2\ MP\ ij} - F_{2\ MP\ i}(x_i, u_{ij}, \theta))^2}{\sigma^2_{2\ MP}} + \frac{(y_{2\ MP\ ij} - F_{3\ MP\ i}(x_i, u_{ij}, \theta))^2}{\sigma^2_{3\ MP}} + \frac{(y_{C3\ ij} - F_{C3\ i}(x_i, u_{ij}, \theta))^2}{\sigma^2_{C3}} + \frac{(u_{m\ i\ C6\ ij} - F_{C6\ i}(u_{ij}, \theta))^2}{\sigma^2_{C6}} + \frac{(u_{m\ H2\ ij} - F_{H2\ i}(u_{ij}, \theta))^2}{\sigma^2_{H2}} \right] (5-8.1)
\]

where the decision variables are:

\[
u_{ij} = \left[ F_{C6\ ij}^0, F_{H2\ i}^0 \right]^T \text{ and } \theta = [k_{\text{ref}1}, k_{\text{ref}2}, k_{\text{ref}3}, K_{\text{ref}}, E_a\ 1, E_a\ 2, E_a\ 3, \Delta H^0]^T
\]

and the perfectly-known inputs are \(x_i = [P_i, T_i, W]^T\).

\[
J_{\text{WLS}} = \sum_{i=1}^{25} \sum_{j=1}^{n_i} \left[ \frac{(y_{2\ MP\ ij} - F_{2\ MP\ i}(x_i, \theta))^2}{\sigma^2_{2\ MP}} + \frac{(y_{2\ MP\ ij} - F_{3\ MP\ i}(x_i, \theta))^2}{\sigma^2_{3\ MP}} + \frac{(y_{C3\ ij} - F_{C3\ i}(x_i, \theta))^2}{\sigma^2_{C3}} \right] (5-8.2)
\]

where the decision variables are \(\theta = [k_{\text{ref}1}, k_{\text{ref}2}, k_{\text{ref}3}, K_{\text{ref}}, E_a\ 1, E_a\ 2, E_a\ 3, \Delta H^0]^T\),

and the perfectly-known inputs are \(x_i = [P_i, T_i, W, F_{H2\ i}^0, F_{C6\ i}^0]^T\).
Table 5-9 Values of parameter estimates obtained by minimizing objective functions in Table 5-8 when different levels of uncertainties in model inputs

<table>
<thead>
<tr>
<th></th>
<th>WLS</th>
<th>EVM with 1% uncertainties in model inputs</th>
<th>EVM with 5% uncertainties in model inputs</th>
<th>EVM with 10% uncertainties in model inputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}<em>{F</em>{2,MP}}^2 (10^{-13} mol^2/s^2)$</td>
<td>1.44</td>
<td>1.43</td>
<td>1.26</td>
<td>0.96</td>
</tr>
<tr>
<td>$\hat{\sigma}<em>{F</em>{3,MP}}^2 (10^{-14} mol^2/s^2)$</td>
<td>6.84</td>
<td>6.78</td>
<td>5.46</td>
<td>1.85</td>
</tr>
<tr>
<td>$\hat{\sigma}<em>{F</em>{C3}}^2 (10^{-14} mol^2/s^2)$</td>
<td>1.94</td>
<td>1.86</td>
<td>0.8</td>
<td>0.16</td>
</tr>
<tr>
<td>$k_{ref1} (10^{-6} mol s^{-1} \ g_{cat}^{-1})$</td>
<td>202.3</td>
<td>200.7</td>
<td>195.7</td>
<td>183.3</td>
</tr>
<tr>
<td>$k_{ref2} (10^{-6} mol s^{-1} \ g_{cat}^{-1})$</td>
<td>115.6</td>
<td>114.7</td>
<td>112.0</td>
<td>105.1</td>
</tr>
<tr>
<td>$k_{ref3} (10^{-6} mol s^{-1} \ g_{cat}^{-1})$</td>
<td>14.8</td>
<td>14.6</td>
<td>13.8</td>
<td>11.8</td>
</tr>
<tr>
<td>$K_{phys} (10^{-6} Pa^{-1})$</td>
<td>10.6</td>
<td>10.0</td>
<td>7.3</td>
<td>3.5</td>
</tr>
<tr>
<td>$E_{a1} (kJ mol^{-1})$</td>
<td>54.1</td>
<td>54.3</td>
<td>54.0</td>
<td>58.2</td>
</tr>
<tr>
<td>$E_{a2} (kJ mol^{-1})$</td>
<td>62.2</td>
<td>62.3</td>
<td>62.0</td>
<td>66.2</td>
</tr>
<tr>
<td>$E_{a3} (kJ mol^{-1})$</td>
<td>68.3</td>
<td>68.7</td>
<td>70.7</td>
<td>77.5</td>
</tr>
<tr>
<td>$\Delta H^0 (kJ mol^{-1})$</td>
<td>−80.9</td>
<td>−83.5</td>
<td>−98.9</td>
<td>−119.8</td>
</tr>
</tbody>
</table>
5.5.3 Parameter and Prediction Uncertainty Quantification

Figure 5-2 shows boxplots for parameter uncertainties for two parameters (i.e., $k_{ref_1}$ and $\Delta H^0$) based on results from 200 bootstrap parameter estimations using the algorithm in Table 5-3. Similar boxplots for the other 6 parameters are provided in the Appendix E. As shown in Figure 5-2, when uncertainties in the inputs are small (i.e., 1% for both $F_{Hz,m}$ and $P_{nc6,m}$) the boxplots of the parameter estimates are very similar to those obtained using WLS parameter estimation, as expected. However, for higher input uncertainties (i.e., 5% and 10%), the confidence intervals for the parameter estimates (i.e., the distance between the ends of the whiskers) are noticeably larger than those obtained using WLS. It is interesting that the confidence interval for $k_{ref_1}$ when input uncertainties are 10% is slightly smaller than the corresponding confidence interval when the input uncertainties are 5%. This unanticipated trend is attributed to the decrease in output measurement variance estimates $\hat{\sigma}_F^2$, $\hat{\sigma}_F^2$ and $\hat{\sigma}_F^2$ when input uncertainties are assumed to be larger (see Table 5-9). Notice that the median of the bootstrap parameter estimates changes as the assumed

Figure 5-1 Contribution of input uncertainty and output uncertainty to the overall pooled estimates of variance $\hat{\sigma}_F^2$ when different levels of input uncertainties are assumed during EVM parameter estimation.
input uncertainties increase, revealing that input uncertainties also have important effects on point estimates for the parameters. Based on the results in Figure 5-2, we recommend that modelers should use WLS in situations where input uncertainties are small. However, when the influence of input uncertainties becomes noticeable compared to output measurement uncertainties, EVM parameter estimation is recommended.

Figure 5-2 Parameter uncertainty boxplots for a) \(k_{\text{ref}}\) and b) \(\Delta H^0\) constructed using 200 bootstrap parameter estimates when different levels of input uncertainties are considered for both \(F_{H_2}^0\) and \(F_{nC_6}^0\). The ends of the whiskers correspond to approximate 95% confidence intervals.

Figure 5-3 shows model-prediction uncertainty boxplots for \(F_{2MP}\) and \(F_{C_3}\) based on 200 MC simulations of the 15th experimental run (i.e., conducted at \(T = 513 K, P = 1 MPa, F_{H_2}^0 = 1.3 \times 10^{-3} \text{mol} / \text{s}\), and \(F_{nC_6}^0 = 2.56 \times 10^{-5} \text{mol} / \text{s}\)) for different assumed levels of input uncertainties.

The algorithm in Table 5-2 was used to obtain these results. Similar boxplots for the third output \(F_{3MP}\) are provided in the Appendix E. Figure 5-3 shows that, as expected, uncertainties in the mean predictions (see Figures 5-3a and 5-3c) when input uncertainties are small are similar to the corresponding uncertainties when WLS is used. When larger input uncertainties are assumed, the sizes of the confidence intervals increase. Figures 5-3b and 5-3d show boxplots for single future measurements of \(F_{2MP}\) and \(F_{C_3}\) that might be obtained using the experimental condition for the 15th experimental run. The cloud of 200 MC predictions for plausible single future measurements was obtained using the results of the 200 MC predictions for the mean response, with additional
random measurement error added. As expected, the uncertainties in the single future measurements are larger than the uncertainties in the corresponding mean predictions. It is interesting that Figure 5-3d shows that the single-prediction uncertainties of $F_{C3}$ are larger when the assumed input uncertainties are smaller. This trend occurs because the output measurement variances for $F_{C3}$ are larger when small input uncertainties are assumed than when larger input uncertainties are assumed (see Table 5-9). It is also interesting that the trend in Figure 5-3b is the opposite of the trend in Figure 5-3d. For $F_{2MP}$, assuming larger input uncertainties results in increasing uncertainty about what might happen in a single new experiment. It seems that large increase in uncertainty for the mean prediction of $F_{2MP}$ (see Figure 5-3a) for this output outweighs the decrease in $\hat{\sigma}^2_{F_{2MP}}$ shown in Table 5-9. In summary, the results in Figures 5-2 and 5-3 reveal that different assumptions about the size of input uncertainties can have important (and sometimes counter-intuitive) consequences for parameter and prediction uncertainties.
Figure 5-3 Prediction uncertainty boxplots for a) mean prediction of $F_{2MP}$, b) a single-future measurement of $F_{2MP}$, c) mean prediction of $F_{C3}$ and d) a single future measurement of $F_{C3}$ constructed based on 200 MC simulations. Predictions correspond to the experimental condition at the 15th run in Table E.1 ($T = 513 \, K$, $P = 1 \, MPa$, and $F_{H2}^0 = 1.3 \times 10^{-3} \, mol/s$, and $F_{nC6}^0 = 2.56 \times 10^{-5} \, mol/s$). The whisker ends correspond to approximate 95% confidence intervals.

Figure 5-4 shows the boxplots for 200 MC simulations for the predictions of $F_{2MP}$ and $F_{C3}$ for a new experimental condition ($T = 580 \, K$, $P = 1.0 \, MPa$, $F_{H2}^0 = 1.3 \times 10^{-3} \, mol/s$, and $F_{nC6}^0 = 2.56 \times 10^{-5} \, mol/s$). Additional results for the other output $F_{3MP}$ are provided in Appendix E. As shown, the confidence intervals for the mean predictions and single future predictions increase when higher input uncertainties are assumed. Also, the confidence intervals for the single future measurements of $F_{2MP}$ and $F_{C3}$ are noticeably larger than the
corresponding confidence intervals for the mean predictions when input uncertainties are neglected or small (i.e., when it is zero or 1%). This can be attributed to the larger estimates of the output measurement variances when the input uncertainties are small. The confidence intervals for the mean and a single future measurement prediction are similar for situations with higher input uncertainties. This can be attributed to higher values of measurement variance estimates in smaller input uncertainties (e.g., \( \hat{\sigma}^2_{F_C^3} = 1.86 \times 10^{-14} \text{mol}^2/\text{s}^2 \) when there are 1% input uncertainties and \( \hat{\sigma}^2_{F_C^3} = 0.16 \times 10^{-14} \text{mol}^2/\text{s}^2 \) when there are 5% input uncertainties).

**Figure 5-4** Prediction uncertainty boxplots for a) mean prediction of \( F_{2MP} \), b) a single future measurement of \( F_{2MP} \), c) mean prediction of \( F_{C^3} \) and d) a single future measurement of \( F_{C^3} \) constructed based on 200 MC simulations. The model predictions are obtained for a new experimental condition (\( T = 580 \, K, P = 1.0 \, MPa, F^0_{H_2} = 1.3 \times 10^{-3} \, \text{mol}/\text{s}, \) and \( F^0_{nC_6} = 2.56 \times 10^{-5} \, \text{mol}/\text{s} \)). The whisker ends correspond to approximate 95% confidence intervals.
5.5.4 Uncertainties in Predicted Yield and Selectivities

Figure 5-5 shows boxplots for yield predictions when different levels of input uncertainties are assumed. These results were obtained by substitution of points from the cloud of MC-based output predictions (i.e., for $F_{2MP}, F_{3MP}$ and $F_{C3}$) into equation (5-8). Figure 5-5a shows yield prediction uncertainty boxplots for the 15th experimental run and Figure 5-5b shows yield prediction uncertainty boxplots for a new experimental condition that was not considered by Toch et al. during their experiments (i.e., $T = 580 \, K$, $P = 1.0 \, MPa$, $F_{H_2 m}^0 = 1.3 \times 10^{-3} \, mol/s$, and $F_{nC_6 m}^0 = 2.56 \times 10^{-5} \, mol/s$). As shown in Figure 5-5, the uncertainties in the yield predictions increase when a higher input uncertainty is assumed, as expected.

Figure 5-6a shows boxplots for the selectivity of 3MP. Additional boxplots for selectivities of other species and at different experimental conditions are provided in Appendix E. It is interesting that in Figure 5-6a, when input uncertainties increase the confidence intervals for the selectivity predictions become smaller. This unanticipated result is attributed to increasing correlation between the numerator ($F_{3MP}$) and the denominator ($F_{nC_6}^0 - F_{nC_6}$) for $S_{3MP}$ in equation (5-9) when input uncertainties become larger. As shown in Figure 5-6b, when input uncertainties are large, the uncertainties in $F_{3MP}$ and $F_{nC_6}^0 - F_{nC_6}$ both become larger, but high correlation between them makes their ratio less uncertain than when input uncertainties are assumed to be small. These results reveal that assumptions about the size of the input uncertainties can have noticeable influence on predictions of selectivities, yields and other quantities that are calculated from model predictions and parameter values. Also, the size of prediction uncertainties can be influenced in important ways by the input uncertainty assumptions that are made by the modeler.
**Figure 5-5** Boxplots for yield predictions constructed based on 200 MC simulations: a) for mean yield predictions at the target experimental condition with $T = 513 \, K$, $P = 1 \, MPa$, $F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nc_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$; b) for mean yield predictions at a new experimental condition with higher temperature where $T = 580 \, K$, $P = 1.0 \, MPa$, $F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nc_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).

**Figure 5-6** Selectivity predictions for 3MP constructed based on 200 MC simulations: a) boxplot for mean selectivity predictions at the experimental condition with $T = 513 \, K$, $P = 1 \, MPa$, $F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nc_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$ for two levels of input uncertainties; b) the corresponding scatter plot for computed values of $F_{3MP}^0$ vs. $F_{nc_6}^0 - F_{nc_6}$ to investigate the correlation between the numerator and denominator in the selectivity expression in equation 5-9. Ends of the whiskers correspond to the approximate 95% confidence intervals.
5.6 Conclusions

Conducting replicate experiments is not straightforward when inputs are uncertain, because it might not be possible to obtain the same true input values during repeated runs. Instead, experimenters can operate their equipment so that measured or target values of the inputs are the same, while acknowledging that the true values may be different. These types of attempted replicate experiments are called pseudo replicates. Modelers can use pseudo-replicate data to estimate the output measurement variances required in EVM objective functions for parameter estimation, if they are careful to properly account for influence of uncertain inputs. Otherwise, their output-measurement-variance weighting factors will be too large.

In this study, we extend a methodology for estimating output measurement variances so that it can be used for multi-response models with multiple uncertain inputs. A case study involving n-hexane hydroisomerization, which has two uncertain inputs (i.e., n-hexane feed rate and hydrogen feed rate) is used to illustrate the proposed method. Kinetic and thermodynamic parameters in this hydroisomerization model are estimated using literature data. Parameter estimates and model predictions are obtained using three different levels of input uncertainty so that results can be compared with weighted-least-squares estimates, which neglect input uncertainties.

Further, we extend a bootstrap method for assessing parameter uncertainties so that it can be applied for multi-response models with several uncertain inputs. The results of the parameter estimation and parameter uncertainty calculations confirm that, when small input uncertainties are assumed (i.e., 1% of the typical flow rates), results from the EVM parameter estimation are similar to those obtained using WLS. However, when larger input uncertainties are assumed (i.e., 5 % or 10 % of the typical feed rates), the parameter estimates and their uncertainties are noticeably different compared to WLS results.
Increasing the input uncertainties resulted in several different trends for parameter confidence intervals. For four of the model parameters (i.e., $E_{a1}, E_{a1}, E_{a1}$ and $\Delta H^0$) increasing input uncertainties resulted in increased width of the confidence intervals. However, for the other four parameters (i.e., $k_{ref1}, k_{ref2}, k_{ref3}$ and $K_{ref}$), increasing the input uncertainties led to an initial increase in the width of the parameter confidence intervals followed by a decrease when input uncertainties increased from 5% to 10%. We attribute this complicated trend to the tradeoff between the influences of the larger input uncertainties and the corresponding smaller estimates for the output measurement variances.

In this study, we obtained prediction uncertainties for the mean response that would be obtained at several different target settings (e.g., $T = 513 \text{ K}, P = 1\text{ MPa}, F_{H_2}^0 = 1.3 \times 10^{-3} \text{ mol s}^{-1}$ and $F_{H_2}^0 m = 2.56 \times 10^{-5} \text{ mol s}^{-1}$). We also obtained larger prediction uncertainties for a single-future experiment conducted at the same target. As input uncertainties increased, the confidence intervals for the predicted mean product outflow rates became wider for all three responses (i.e., $F_{2\text{MP}}, F_{3\text{MP}}$ and $F_{c_2}$). However, confidence intervals for the corresponding single-future measured responses increased for two responses (i.e., $F_{2\text{MP}}$ and $F_{3\text{MP}}$) but decreased for the third. This result is complicated but not surprising due to the tradeoff between larger assumed input uncertainties and smaller estimated output measurement variances.

We also obtained uncertainties in the predicted yield and selectivities for 2MP and 3MP which are calculated from the predicted product outflow rates. These results showed that increasing the input uncertainties led to increasing width of the confidence interval for the yield. However, more complicated results were obtained for selectivity confidence intervals. For example, the selectivity confidence interval for 3MP became narrower when increasing input uncertainty was assumed. These results reveal that assumptions about input uncertainties can have important and sometimes complicated implications for parameter estimates, model predictions and other predicted quantities of interest to modelers and model users. We recommend that modelers should
use weighted-least-squares estimation when they believe that uncertainties in experimental inputs have minimal influence on the overall reproducibility of their experiments. However, when input uncertainties are an important issue of concern, EVM parameter estimation should be used. The methods proposed in the current chapter should be helpful for assessing the resulting uncertainties in model parameters and model predictions.
5.7 References for Chapter 5


Chapter 6

Summary, Conclusions and Recommendations

In chapter 2, a mathematical model is proposed for photopolymerization of 1,6-hexanediol diacylate with a bifunctional initiator BAPO. This mathematical model accounts for a variety of phenomena including initiation, propagation, branching (crosslinking), backbiting, cyclization, and diffusion-dependent reactions. The corresponding model includes 40 parameters requiring estimation. To do so, a systematic parameter subset selection and parameter estimation method was used to estimate a subset of the model parameters while avoiding overfitting of the industrial data. The results revealed that 30 parameters are estimable. These parameters are related to initiation, propagation, termination, branching (crosslinking), backbiting, and cyclization reactions. Further, the parameter estimability analysis showed that, initiation efficiency, propagation reactions and reactions involving large molecules (i.e., termination reactions involving two macroradicals and branching reactions involving a pendant vinyl group and a macroradical) become diffusionally-dependent. Furthermore, the estimability analysis reveals that diffusional effects related to thirteen reactions involving small molecules can be neglected (i.e., initiation reactions involving small molecules, cyclization reaction, backbiting reactions and termination reactions involving two small molecules). The resulting model fit to data that were used for parameter estimation is very good (i.e., within 1.61 % for overall vinyl group conversion). Model predictions for two dynamic data sets saved for model validation are excellent (i.e., within 1.34 % for overall vinyl group conversion). The current model only accounts for situations where the oxygen concentration is negligible in the reaction atmosphere. I have the following recommendations related to the study in Chapter 2:
i. Use the model for situations where the oxygen concentration is negligible to predict the conversion of vinyl groups and monomer.

ii. Include more experimental data for different quantities such as more monomer concentration data to estimate parameters more reliably.

iii. Investigate whether NMR, electron spin resonance (ESR) or other methods can be used to quantify additional details about branch points, cyclic groups and radicals that could provide valuable data for parameter estimation.

iv. Improve the model to account for oxygen contamination within the polymerizing film.

v. Improve the model to account for concentration gradients within the film that can arise due to oxygen contamination and spatial gradients in light intensity.

vi. Modify the model to account for shrinkage of the polymer film as polymerization proceeds.

vii. Account for additional species during polymerization such as comonomers, inhibitors and pigments.

viii. Account for input uncertainties when determining the parameter estimates and the optimal subset of parameters that can be estimated from the available data.

Note that I recently coauthored a follow-up study related to recommendation iv. This contribution is included as Appendix B. The model proposed in Appendix B is accurate for situations where the film thickness is small and light intensity is low so that uniform oxygen concentration within the film can be assumed. In our research group, we are now working on models that can account for concentration gradients and film shrinkage during polymerization.

In Chapter 3, methods are proposed for estimation of output measurement variances used as weighting factors in EVM objective functions. When model inputs are uncertain, conducting replicate experiments could be challenging, because it may not be possible to have similar true values of uncertain inputs during different attempted replicate runs. When the true values of uncertain inputs are thought to be different during different repeated runs, but their measured values
are the same, I categorize the repeated runs as “pseudo-replicates”. In situations where true values are the same for several replicate runs, the attempted replicate runs are called “true replicates”. In Chapter 3, I provide methodologies for calculating estimated output measurement variances for both situations. A new linearization-based approach is proposed to obtain output measurement variances when data from pseudo-replicate experiments are available. When true replicates are available, output measurement variances can be computed using traditional techniques. Further, a bootstrap technique is extended to obtain information about parameter uncertainties. A mathematical model for n-butyl methacrylate/n-butyl acrylate copolymerization, along with corresponding literature data, is used in Chapter 3 to illustrate the proposed methodologies. Reactivity ratio values and their uncertainties were obtained for the copolymerization model when three different assumptions are made about the output measurement variances: i) that they are known, ii) that they are unknown and can be estimated from true replicate experiments and iii) that they are unknown and can be estimated from pseudo-replicate experiments. The results of parameter estimation revealed that the estimated reactivity ratio values are similar when different assumptions about repeated experiments are made. However, parameter-estimate uncertainties change noticeably when different assumptions are made about output measurement variances and the types of replicates. For example, the uncertainties of the parameters are very large when the output measurement standard deviations are assumed to be 5 % of the measured values. When smaller output measurement variances were estimated from the replicate experiments, parameter uncertainties were smaller. When pseudo replicates were assumed using the same data set, the calculated uncertainties in the parameter estimates were smaller than when true replicates were assumed. This result makes sense because smaller estimates for output measurement variances were obtained when pseudo replicates were assumed. Chapter 3 provides guidelines for experimentalists who want to conduct replicate runs when some of their inputs are uncertain. For example, I explain how they can decide whether their repeated experiments should be treated as true replicates or
pseudo replicates. As a result, modelers will be able to choose appropriate statistical methods for parameter estimation and parameter uncertainty quantification.

In Chapter 3, I used a linearization-based approach to account for the contribution of input uncertainties to the calculated output uncertainties for pseudo-replicate runs. However, the proposed linearization-based approach may not be accurate in situations where input uncertainties are large and models are highly nonlinear in the inputs and the parameters. For the future studies, I propose the following recommendation to develop a more-reliable (but more computationally expensive) methodology compared to the linearization-based approach proposed in Chapter 3:

1. Compute output variance estimates using a Monte-Carlo-based approach that accounts for model nonlinearity when input uncertainties are large and repeated experiments are pseudo replicates.

Furthermore, the research in Chapter 3 has made me realize that determining appropriate values of the degrees of freedom for EVM parameter estimates is a complicated issue that has not been discussed in the literature. I recommend that:

2. Someone should explore the difficult problem of determining the degrees of freedom associated with variance estimates from pseudo replicates.

3. Someone should explore the difficult problem of assessing the degrees of freedom for parameter estimates which are computed using estimated output variances and perfectly known input variances.

In Chapter 4, a literature review study is provided for uncertainty quantification in situations where input uncertainties are assumed to be important. I provide a summary of previous studies along with recommendations for three important situations: i) where a model user wants to obtain prediction uncertainties based on known input uncertainties, when uncertainties in model parameters are relatively small so that they can be neglected; ii) where a model user wants to obtain parameter uncertainties based on data from experiments with uncertainties in inputs and outputs;
iii) where a model user wants to obtain prediction uncertainties based on uncertain EVM parameter estimates and uncertain inputs. For all three situations, linearization-based and Monte Carlo methods were reviewed. Linearization-based methods require less computation effort than MC methods and are recommended for situations where the nonlinearity of the model and the extent of uncertainties are modest. Monte Carlo methods, which require less-restrictive assumptions, are recommended for situations with larger uncertainties and model nonlinearities. Further, Monte Carlo methods are appropriate for prediction-uncertainty quantification in situations where the size of the data set is relatively small and information about the degrees of freedom for parameter estimates is difficult to determine. In this situation, Monte Carlo simulations are recommended.

The main drawbacks of Monte Carlo methods for assessing EVM parameter and prediction uncertainties are that: i) they are computationally intensive, especially when the corresponding mathematical model is complicated and ii) they require the assumption that the point estimates of parameters are close enough to the true values so that the resulting “clouds” of Monte-Carlo parameter estimates and Monte-Carlo predictions will give reliable uncertainty information.

The information provided in Chapter 4 will be useful in future EVM studies involving uncertainty quantification of parameters and model predictions. Based on the study in Chapter 4, I make the following recommendations for future research:

i. Modelers can use linearization-based and Monte Carlo methods for uncertainty quantification in EVM models. They should choose between the available methods based on their concerns about model nonlinearity, the size of input uncertainties and the computational costs of solving the model equations.

ii. During prediction uncertainty quantification, modelers should be careful about the type of prediction uncertainties they are interested to obtain (i.e., they should think about whether they want to obtain uncertainties for mean predictions at a certain operating condition or
they want to obtain the uncertainty associated with the outcome of a single future experiment).

In Chapter 5, a hexane hydrosimerization model from the literature is used to illustrate output-measurement-variance estimation and parameter estimation in a situation where data from pseudo-replicate experiments are available. The hexane hydroisomerization model used in this study contains several model outputs (i.e., outflow rates of products) and two uncertain inputs (i.e., feed flow rates of hydrogen and n-hexane). Three levels of input uncertainties are considered for the uncertain inputs and the parameter estimation and prediction uncertainty quantification results are compared for four situations (i.e., for 0 %, 1%, 5 % and 10 % typical errors, respectively, in the feed flowrates). When 0 % uncertainty input uncertainty is assumed weighted-least squares (WLS) parameter estimation is used. When larger input uncertainties are assumed EVM parameter estimation is used based on output uncertainties estimated from pseudo-replicate data. The results of variance estimation calculations and the associated parameter estimations reveal that when input uncertainties are small (i.e., 1 % of the typical values of uncertain feed rates) the resulting variance estimates and parameter estimates are very similar to the corresponding values obtained from WLS parameter estimation. However, when higher input uncertainties are assumed for the uncertain inputs, noticeably different measurement variance values and parameter estimates are obtained. Further, in Chapter 5, the parameter uncertainties and prediction uncertainties are obtained for the parameter estimates and model predictions. A bootstrap technique from Chapter 3 is extended so that parameter uncertainty boxplots can be obtained. The prediction uncertainty results confirm that assuming 1 % uncertainties in inputs leads to similar results compared to prediction uncertainties obtained from WLS parameter estimates. For larger input uncertainties, however, noticeably different prediction uncertainties are obtained. As expected, prediction uncertainties for single
future experiments are larger than prediction uncertainties for mean process behaviour, which was confirmed using two conditions of interest.

Unanticipated results were obtained for some of the prediction uncertainties. For example, in some situations where the assumed input uncertainties increased, the resulting prediction uncertainties decreased. This trend is attributed to a tradeoff between an increase in the influence of the input uncertainty and the associated decrease in the estimated output measurement variances. The increasing input uncertainty, by itself, increases the prediction uncertainty whereas the decreasing output uncertainty estimate decreases the prediction uncertainty.

The recommendations following from Chapter 5 are:

i. Modelers should use the proposed output-measurement-variance-quantification methodology for EVM parameter estimation and output-measurement-variance estimation when pseudo-replicate data are available.

ii. Future researchers may want to develop Monte-Carlo-based methods for output-measurement-variance estimation to replace the linearization-based approach developed in this thesis. A Monte Carlo approach may lead to more accurate and reliable results in some situations where models are highly nonlinear and input uncertainties are large.
Appendix A

Supplementary Information for Chapter 2

The change in the values of the objective function vs. the number of estimated parameters for Case I is shown in Figure A.1. The estimated parameters and their corresponding ranks are shown in Table A.1. The estimated parameters in Case I were used to obtain simulation results in Figures A.2 to A.6. Representative simulation results for concentrations of initiator molecule, the remaining initiator fragments containing phosphorous-carbon weak bonds, monomer, and pendant vinyl groups are shown in Figure A.2. In Figure A.3, representative simulation results are shown for different initiator radical types (i.e., $I^\ast$, $\tilde{I}^\ast$, and $I_p^\ast$). Figure A.4 shows the simulation results for large macromolecules (i.e., $R^\ast$ and $B^\ast$). The representative simulations are chosen for only some of the experiments with different light intensities and exposure times. In Figure A.5, the conversion vs. time predictions is compared with their corresponding measured values. In Figure A.6, model predictions of the end monomer conversions are compared with their corresponding measured values for the short-time-exposure experiments.

![Figure A.1](image)

**Figure A.1** Influence of the number of estimated parameters from the ranked list on the value of objective function in Case I parameter estimation
Table A.1 The list of ranked parameters and their corresponding estimates for Case I parameter estimation

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parameter ranks</th>
<th>Estimated</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_{f,0}$</td>
<td>3</td>
<td>0.0736</td>
<td>-</td>
</tr>
<tr>
<td>$v_{f,1}/v_{f,0}$</td>
<td>13</td>
<td>0.433</td>
<td>-</td>
</tr>
<tr>
<td>$\phi$</td>
<td>8</td>
<td>0.800</td>
<td>-</td>
</tr>
<tr>
<td>$f_0$</td>
<td>4</td>
<td>0.507</td>
<td>-</td>
</tr>
<tr>
<td>$k_d/k_d$</td>
<td>10</td>
<td>0.102</td>
<td>-</td>
</tr>
<tr>
<td>$f/f$</td>
<td>14</td>
<td>0.509</td>
<td>-</td>
</tr>
<tr>
<td>$k_{in,IMO}/k_{in,MO}$</td>
<td>20</td>
<td>0.01</td>
<td>-</td>
</tr>
<tr>
<td>$k_{in,IMO}$</td>
<td>19</td>
<td>$5 \times 10^7$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>$k_{in,IVP_0}/k_{in,IMO} = k_{in,IVP_0}/k_{in,IMO}$</td>
<td>15</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>$k_{p,R0}$</td>
<td>1</td>
<td>$4.93 \times 10^4$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>$k_{p,BO} = k_{b,BO}$</td>
<td>6</td>
<td>0.0748</td>
<td>-</td>
</tr>
<tr>
<td>$k_{p,R0} = k_{b,R0}$</td>
<td>16</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>$A_{BR}$</td>
<td>11</td>
<td>0.996</td>
<td>-</td>
</tr>
<tr>
<td>$v_{FCBR}$</td>
<td>2</td>
<td>0.496</td>
<td>-</td>
</tr>
<tr>
<td>$v_{FCR}$</td>
<td>18</td>
<td>$5 \times 10^7$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>$k_{in,I}$</td>
<td>1</td>
<td>Not Estimable</td>
<td>-</td>
</tr>
<tr>
<td>$k_{LBR,0}$</td>
<td>9</td>
<td>$1.60 \times 10^7$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>$A_{LR}$</td>
<td>12</td>
<td>0.502</td>
<td>-</td>
</tr>
<tr>
<td>$v_{FCBR}$</td>
<td>7</td>
<td>0.078</td>
<td>-</td>
</tr>
<tr>
<td>$R_{rd}$</td>
<td>5</td>
<td>0.109</td>
<td>L mol$^{-1}$</td>
</tr>
<tr>
<td>$k_{LBB0}$</td>
<td>17</td>
<td>0.500</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure A. 2 Simulation results corresponding to Case I for a) Initiator molecules $I$, b) initiator fragments attached to polymer chains $\tilde{I}$, c) Monomer, and d) pendant vinyl groups $V_p$. 
Figure A. 3 Simulation results obtained from parameter estimates in Case I for concentration of a) carbon-centered initiator radicals $I^\bullet$, b) small phosphorous-centered radicals $I_\sim^\bullet$, and c) large molecules with phosphorous-centered radicals $I_P^\bullet$.
Figure A. 4 Numerical simulation corresponding to Case I for a) Regular radical ends $R^*$, and b) Radicals at the branch points $B^*$

Figure A. 5 Comparison between model predictions (curves) from Case I and their corresponding measurements (shown with ×, ×, ×, ×, and ×) for different light intensities. The same colors used for the prediction curves and the corresponding measurement data.
Figure A. 6 Comparison between final monomer conversion predictions and their corresponding measured values obtained using estimated parameters from Case I for short-time-exposure experiments with light intensity of 20 W/m². Straight lines between the model predictions are shown to guide the eye.

In Figure A.7, the change in objective function values vs. the number of estimated parameters is shown for parameter estimation in Case II. In Table A.2, the results of parameter ranking and the estimated parameters corresponding to the Case II are shown. Parameter estimates in Table A.2 were used to obtain simulation results in Figures A.8 to A.12. In Figure A.8, representative simulation results are shown for concentrations of some species (i.e., $M$, $I$, $\tilde{I}$, and $V_p$). In Figure A.9, representative simulation results are shown for concentrations of carbon-centered initiator radicals $I^*$ and phosphorous-centered initiator radicals $\tilde{I}^*$. Similarly, simulation results are shown in Figure A.10 for concentrations of large macroradicals (i.e., $R^*$, $B^*$, and $T^*$). In Figure A.11, the model predictions for overall vinyl group conversions are compared with their corresponding measured values for representative data sets. In Figure A.11, the final monomer conversion predictions are compared with their corresponding measured values for the short-time-exposure experiments.
Figure A. 7 Influence of the number of estimated parameters from the parameter ranks list on the value of objective function in Case II parameter estimation
Table A. 2 The list of ranked parameters and their corresponding estimates for Case II parameter estimation

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parameter ranks</th>
<th>Estimated</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_{f,0}$</td>
<td>Not estimable</td>
<td>0.0749</td>
<td>-</td>
</tr>
<tr>
<td>$v_{f,1}/v_{f,0}$</td>
<td>3</td>
<td>0.431</td>
<td>-</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Not estimable</td>
<td>0.800</td>
<td>-</td>
</tr>
<tr>
<td>$f_0$</td>
<td>Not estimable</td>
<td>0.507</td>
<td>-</td>
</tr>
<tr>
<td>$k_d/k_d$</td>
<td>Not estimable</td>
<td>0.102</td>
<td>-</td>
</tr>
<tr>
<td>$\dot{f}/f$</td>
<td>Not estimable</td>
<td>0.509</td>
<td>-</td>
</tr>
<tr>
<td>$k_{in,IM}/k_{in,IM0}$</td>
<td>Not estimable</td>
<td>0.01</td>
<td>-</td>
</tr>
<tr>
<td>$k_{in,IM0}$</td>
<td>Not estimable</td>
<td>$5 \times 10^7$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>$k_{in,IVp}/k_{in,IM0} = k_{in,IVp}/k_{in,IM0}$</td>
<td>Not estimable</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>$k_{P,R0}$</td>
<td>Not estimable</td>
<td>$6.08 \times 10^4$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>$k_{p,R0}/k_{p,R0} = k_{b,R0}/k_{b,R0}$</td>
<td>20</td>
<td>0.0774</td>
<td>-</td>
</tr>
<tr>
<td>$k_{p,T0}/k_{p,R0} = k_{b,T0}/k_{b,R0}$</td>
<td>4</td>
<td>$2.18 \times 10^{-2}$</td>
<td>-</td>
</tr>
<tr>
<td>$k_{BR,0}/k_{PR,0}$</td>
<td>Not estimable</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>$A_{BR}$</td>
<td>Not estimable</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>$v_{f,IBR}/v_{f,II}$</td>
<td>2</td>
<td>0.502</td>
<td>-</td>
</tr>
<tr>
<td>$k_{b,R0}/k_{p,R0}$</td>
<td>Not estimable</td>
<td>$8.00 \times 10^{-3}$</td>
<td>mol L$^{-1}$</td>
</tr>
<tr>
<td>$k_{t_{in,II}}$</td>
<td>Not estimable</td>
<td>$5.0 \times 10^7$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>$k_{t_{in,III}}/k_{t_{in,II}}$</td>
<td>Not estimable</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>$k_{LRR,0}$</td>
<td>1</td>
<td>$1.534 \times 10^7$</td>
<td>L mol$^{-1}$ s$^{-1}$</td>
</tr>
<tr>
<td>$A_{LRR}$</td>
<td>Not estimable</td>
<td>0.501</td>
<td>-</td>
</tr>
<tr>
<td>$v_{f,IRR}$</td>
<td>6</td>
<td>0.076</td>
<td>-</td>
</tr>
<tr>
<td>$R_{bd}$</td>
<td>Not estimable</td>
<td>0.109</td>
<td>L mol$^{-1}$</td>
</tr>
<tr>
<td>$k_{LBD0}/k_{LRR0}$</td>
<td>Not estimable</td>
<td>0.500</td>
<td>-</td>
</tr>
<tr>
<td>$k_{L,TT0}/k_{LRR0}$</td>
<td>Not estimable</td>
<td>0.050</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure A. 8 Simulation results corresponding to parameter estimation in Case II for concentration of a) unreacted initiators $I$, b) Initiator fragments attached to other molecules $\tilde{I}$, c) monomer $M$, and d) pendant vinyl groups $V_P$.
Figure A. 9 Simulation results corresponding to the Case II parameter estimation for concentrations of a) carbon-centered initiator radicals $I^\bullet$, b) phosphorous-centered initiator radicals $\tilde{I}^\bullet$, and c) large molecules with phosphorous-centered radicals $I_p^\bullet$.
Figure A. 10 Simulation results associated with the Case II parameter estimation for concentration of a) Regular radical ends $R^\ast$, and b) radical at the branch points $B^\ast$, and c) tertiary radicals $T^\ast$
Figure A. 11 Comparison between model predictions and their corresponding measurement values (shown with ×××× and ×) obtained using parameter estimates in Case II for some of the conversion vs. time data at some of the light intensities. Similar colors used for the prediction curves and the corresponding measurement data.
Figure A. 12 Comparison between model predictions and the corresponding measurement data obtained using the estimated parameters in Case II for different pulse experiments at $I_{I_0}=20$ W/m$^2$. Straight lines between the model predictions are shown to guide the eye.

In Figure A.13, the values of objective function vs. the number of estimated parameters are shown for Case III. The values of estimated parameters and their ranks in Case III are shown in Table A.3. The parameter estimates in Table A.3 were used to obtain simulation results in Figures A.14 to A.18. The simulation results for concentrations of $M$, $I$, $\tilde{I}$, and $V_p$ are shown in Figure A.14 for representative experiments. In Figure A.15, the results of simulation results are shown for concentrations of initiator radicals including carbon-centered $I^*$ radicals, phosphorous-centered initiator radicals $\tilde{I}^*$, and phosphorous-centered radicals attached to large molecules $I_p^*$. In Figure A.16, simulation results are provided for large macroradicals (i.e., $R^*$, $B^*$, $C^*$, and $T^*$). The results of model predictions and their associated measured values are provided in Figure A.17 for overall vinyl group conversions for conversion data sets corresponding to some of the light intensities used for parameter estimation. In Figure A.18, the results of model predictions and the corresponding measured values are provided for final monomer conversions at the end of short-time-exposure experiments.
Figure A. 13 Influence of the number of estimated parameters from the parameter ranks list on the value of objective function in Case III parameter estimation.
Table A. 3 The list of ranked parameters and their corresponding estimates for Case III parameter estimation

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parameter ranks</th>
<th>Estimated</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>(v_{f,0})</td>
<td>9</td>
<td>0.058</td>
<td>-</td>
</tr>
<tr>
<td>(v_{f,1}/v_{f,0})</td>
<td>4</td>
<td>0.50</td>
<td>-</td>
</tr>
<tr>
<td>(f_0)</td>
<td>13</td>
<td>0.90</td>
<td>-</td>
</tr>
<tr>
<td>(k_d/k_d)</td>
<td>10</td>
<td>0.199</td>
<td>-</td>
</tr>
<tr>
<td>(\dot{f}/f)</td>
<td>18</td>
<td>0.565</td>
<td>-</td>
</tr>
<tr>
<td>(k_{in,IMO}/k_{in,MO})</td>
<td>22</td>
<td>0.0051</td>
<td>-</td>
</tr>
<tr>
<td>(k_{in,IMO})</td>
<td>24</td>
<td>1.0 \times 10^6</td>
<td>L mol(^{-1})s(^{-1})</td>
</tr>
<tr>
<td>(k_{in,IVp0}/k_{in,MO} = k_{in,IVp0}/k_{in,IMO})</td>
<td>13</td>
<td>0.41</td>
<td>-</td>
</tr>
<tr>
<td>(k_{p,R0})</td>
<td>11</td>
<td>1.479 \times 10^4</td>
<td>L mol(^{-1})s(^{-1})</td>
</tr>
<tr>
<td>(k_{p,CO}/k_{p,R0} = k_{b,CO}/k_{b,R0})</td>
<td>19</td>
<td>0.50</td>
<td>-</td>
</tr>
<tr>
<td>(k_{p,B0}/k_{p,R0} = k_{b,B0}/k_{b,R0})</td>
<td>3</td>
<td>0.097</td>
<td>-</td>
</tr>
<tr>
<td>(k_{p,T0}/k_{p,R0} = k_{b,T0}/k_{b,R0})</td>
<td>17</td>
<td>6.49 \times 10^{-4}</td>
<td>-</td>
</tr>
<tr>
<td>(k_{b,R0}/k_{p,R0})</td>
<td>15</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>(A_p)</td>
<td>12</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>(v_{f,c})/(v_{f,c,RR})</td>
<td>2</td>
<td>0.492</td>
<td>-</td>
</tr>
<tr>
<td>(k_{CO}/k_{p,R0})</td>
<td>6</td>
<td>3.39</td>
<td>mol L(^{-1})</td>
</tr>
<tr>
<td>(k_{BB0}/k_{B,R0})</td>
<td>25</td>
<td>0.01</td>
<td>mol L(^{-1})</td>
</tr>
<tr>
<td>(k_{t,t0})</td>
<td>19</td>
<td>9.995 \times 10^8</td>
<td>L mol(^{-1})s(^{-1})</td>
</tr>
<tr>
<td>(k_{in,III}/k_{in,III})</td>
<td>26</td>
<td>1.35</td>
<td>-</td>
</tr>
<tr>
<td>(k_{ERR})</td>
<td>1</td>
<td>1.138 \times 10^7</td>
<td>L mol(^{-1})s(^{-1})</td>
</tr>
<tr>
<td>(A_{ERR})</td>
<td>8</td>
<td>0.50</td>
<td>-</td>
</tr>
<tr>
<td>(v_{f,ERR})</td>
<td>5</td>
<td>0.0776</td>
<td>-</td>
</tr>
<tr>
<td>(R_t)</td>
<td>7</td>
<td>5.17</td>
<td>L mol(^{-1})</td>
</tr>
<tr>
<td>(k_{L,CC0}/k_{L,RR0})</td>
<td>27</td>
<td>1.031</td>
<td>-</td>
</tr>
<tr>
<td>(k_{L,RR0}/k_{L,RR0})</td>
<td>16</td>
<td>0.0102</td>
<td>-</td>
</tr>
<tr>
<td>(k_{L,TT0}/k_{L,RR0})</td>
<td>21</td>
<td>0.085</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure A. 14 Simulation results corresponding to parameter estimation in Case III for concentration of a) unreacted initiators $I$, b) Initiator fragments attached to other molecules $\tilde{I}$, c) monomer $M$, d) pendant vinyl groups $V_p$
Figure A. 15 Simulation results corresponding to the Case III parameter estimation for concentrations of a) carbon-centered initiator radicals $I^*$, b) small phosphorous-centered initiator radicals $\tilde{I}^*$, and c) large phosphorous-centered radical $I_p^*$. 
Figure A. 16 Simulation results corresponding to the Case III parameter estimation for concentrations of a) regular radical ends $R^\bullet$, b) radicals at the branch points $B^\bullet$, c) cyclized radical ends $C^\bullet$, and d) tertiary radicals $T^\bullet$. 
Figure A. 17 Experimental data (shown with ×, ., and •) and corresponding model predictions obtained using parameter estimates obtained in Case III for some of the conversion data at some of the light intensities. Similar colors used for the prediction curves and the corresponding measurement data.

Figure A. 18 Comparison between final monomer conversion predictions and their corresponding measured values obtained using estimated parameters from Case III for short-time-exposure experiments with light intensity of 20 W/m². Straight lines between the model predictions are shown to guide the eye.
In Figure A.19, the values of objective function value vs. the number of estimated parameters are shown for parameter estimation in Case IV. In Table A.4, the values of the estimated parameters and parameters ranks are shown for Case IV. The estimated parameters in Table A.4 were used to obtain simulation results in Figures A.20 to A.24. In Figure A.20, the simulation results are shown for $I$, $I$, $M$, and $V_p$ at representative experimental settings. In Figure A.21, simulation results are shown for initiator radicals $I^*$ and $I^*$ and $I_p^*$. In Figure A.22, the results of simulations are shown for large macromolecules (i.e., $R^*$, $B^*$, $C^*$, and $T^*$). The results of model predictions for vinyl group conversions are compared with the corresponding measured values in Figure A.23. In Figure A.24, the model predictions for the final monomer conversions are compared with the corresponding measured values for the short-time-exposure experiments.
Figure A.19 Influence of the number of estimated parameters from the ranked list on the value of objective function in Case IV.

Table A.4 The list of ranked parameters and their corresponding estimates for Case IV parameter estimation

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parameter ranks</th>
<th>Estimated</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>(v_{f,0})</td>
<td>8</td>
<td>0.064</td>
<td>-</td>
</tr>
<tr>
<td>(v_{f,1}/v_{f,0})</td>
<td>10</td>
<td>0.50</td>
<td>-</td>
</tr>
<tr>
<td>(\phi)</td>
<td>16</td>
<td>0.90</td>
<td>-</td>
</tr>
<tr>
<td>(f_0)</td>
<td>6</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>(A_f)</td>
<td>11</td>
<td>0.50</td>
<td>-</td>
</tr>
<tr>
<td>(k_d/k_d)</td>
<td>14</td>
<td>0.57</td>
<td>-</td>
</tr>
<tr>
<td>(\bar{f}/f)</td>
<td>30</td>
<td>1.20</td>
<td>-</td>
</tr>
<tr>
<td>(k_{in,IMO}/k_{in,IMO})</td>
<td>19</td>
<td>0.005</td>
<td>-</td>
</tr>
<tr>
<td>(k_{in,IMO})</td>
<td>3</td>
<td>1.0 \times 10^6</td>
<td>L mol^{-1} s^{-1}</td>
</tr>
<tr>
<td>(A_{in,IM})</td>
<td>Not estimable</td>
<td>0.7</td>
<td>-</td>
</tr>
<tr>
<td>(v_{fc,iniM}/v_{fc,IRR})</td>
<td>33</td>
<td>0.121</td>
<td>-</td>
</tr>
<tr>
<td>(k_{in,IVp0}/k_{in,IMO})</td>
<td>24</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>(A_{in,IVp})</td>
<td>32</td>
<td>0.63</td>
<td>-</td>
</tr>
<tr>
<td>(v_{fc,iniIVp}/v_{fc,IRR})</td>
<td>26</td>
<td>0.22</td>
<td>-</td>
</tr>
<tr>
<td>(k_{P,R0})</td>
<td>7</td>
<td>(1.00 \times 10^4)</td>
<td>L mol(^{-1}) s(^{-1})</td>
</tr>
<tr>
<td>(A_p/A_b)</td>
<td>36</td>
<td>0.945</td>
<td>-</td>
</tr>
<tr>
<td>(v_{TCP}/v_{FC.D})</td>
<td>12</td>
<td>0.87</td>
<td>-</td>
</tr>
<tr>
<td>(k_{P,CO}/k_{P,R0} = k_{D,CO}/k_{D,R0})</td>
<td>21</td>
<td>0.500</td>
<td>-</td>
</tr>
<tr>
<td>(k_{P,R0}/k_{P,R0} = k_{B,R0}/k_{B,R0})</td>
<td>15</td>
<td>0.048</td>
<td>-</td>
</tr>
<tr>
<td>(k_{P,T0}/k_{P,R0} = k_{D,T0}/k_{D,R0})</td>
<td>2</td>
<td>(5.06 \times 10^{-4})</td>
<td>-</td>
</tr>
<tr>
<td>(k_{D,R0}/k_{P,R0})</td>
<td>20</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>(A_{bb})</td>
<td>37</td>
<td>0.70</td>
<td>-</td>
</tr>
<tr>
<td>(v_{TCP}/v_{FC.BB})</td>
<td>5</td>
<td>0.53</td>
<td>-</td>
</tr>
<tr>
<td>(k_{C,0}/k_{P,R0})</td>
<td>9</td>
<td>3.35</td>
<td>mol L(^{-1})</td>
</tr>
<tr>
<td>(A_{bb})</td>
<td>29</td>
<td>0.63</td>
<td>-</td>
</tr>
<tr>
<td>(v_{TCP}/v_{TCP})</td>
<td>28</td>
<td>0.25</td>
<td>-</td>
</tr>
<tr>
<td>(k_{bb}/k_{P,R0})</td>
<td>25</td>
<td>(1.00 \times 10^{-3})</td>
<td>mol L(^{-1})</td>
</tr>
<tr>
<td>(A_{bb})</td>
<td>37</td>
<td>0.70</td>
<td>-</td>
</tr>
<tr>
<td>(v_{TCP}/v_{TCP})</td>
<td>31</td>
<td>0.145</td>
<td>-</td>
</tr>
<tr>
<td>(k_{in/10})</td>
<td>39</td>
<td>(6.98 \times 10^8)</td>
<td>L mol(^{-1}) s(^{-1})</td>
</tr>
<tr>
<td>(A_{11})</td>
<td>34</td>
<td>0.545</td>
<td>-</td>
</tr>
<tr>
<td>(v_{TCP}/v_{TCP})</td>
<td>35</td>
<td>0.15</td>
<td>-</td>
</tr>
<tr>
<td>(k_{in/11}/k_{in/11})</td>
<td>38</td>
<td>1.15</td>
<td>-</td>
</tr>
<tr>
<td>(k_{L,RR0})</td>
<td>1</td>
<td>(1.69 \times 10^8)</td>
<td>L mol(^{-1}) s(^{-1})</td>
</tr>
<tr>
<td>(A_{L,RR})</td>
<td>18</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>(v_{TCP})</td>
<td>13</td>
<td>0.086</td>
<td>-</td>
</tr>
<tr>
<td>(R_{rd})</td>
<td>17</td>
<td>7.31</td>
<td>L mol(^{-1})</td>
</tr>
<tr>
<td>(k_{C,CO}/k_{L,RR0})</td>
<td>27</td>
<td>1.50</td>
<td>-</td>
</tr>
<tr>
<td>(k_{L,RR0}/k_{L,RR0})</td>
<td>4</td>
<td>0.01</td>
<td>-</td>
</tr>
<tr>
<td>(k_{L,RR0}/k_{L,RR0})</td>
<td>23</td>
<td>0.088</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure A.20 Simulation results corresponding to parameter estimation in Case IV for concentration of a) unreacted initiators $I$, b) Initiator fragments attached to other molecules $\tilde{I}$, c) monomer $M$, and d) pendant vinyl groups $V_p$. 
Figure A. 21 Simulation results corresponding to the Case IV parameter estimation for concentrations of a) carbon-centered initiator radicals $I^*$, b) phosphorous-centered initiator radicals $\tilde{I}^*$, and c) large molecules with phosphorous-centered radicals $I_p$. 
Figure A. 22 Simulation results corresponding to the Case IV parameter estimation for concentrations of a) regular radical ends $R^*$, b) radicals at the branch points $B^*$, c) cyclized radical ends $C^*$, and d) tertiary radicals $T^*$
Figure A. 23 Experimental data (shown with ×,×,×, and ×) and corresponding model predictions obtained using parameter estimates obtained in Case IV for some of the conversion data at some of the light intensities. Similar colors used for the prediction curves and the corresponding measurement data.

Figure A. 24 Comparison between final monomer conversion predictions and their corresponding measured values obtained using estimated parameters from Case IV for short-time-exposure experiments with light intensity of 20 W/m². Straight lines between the model predictions are shown to guide the eye.

The change in the values of the objective function vs. the number of estimated parameters in the final parameter-estimation step is shown in Figure A.25. The estimated parameters in the final step (Table
2.9 in section 2.5) were used to obtain simulation results in Figures A.26 to A.28. In Figure A.26, the simulation results are shown for $I, \bar{I}, M, \text{ and } V_p$ for representative experiments. In Figure A.27, simulation results are shown for initiator radicals $I^*, \bar{I}^*$, and phosphorous-centered radicals attached to large molecules $I_p^*$. In

![Graph showing the influence of number of estimated parameters on the optimal value of objective function.](image)

**Figure A. 25** Influence of number of estimated parameters from the ranking list on the optimal value of objective function in the final parameter estimation step
Figure A. 26 Simulation results corresponding to the final parameter estimation step for concentration of a) unreacted initiators $I$, b) Initiator fragments attached to other molecules $Ĩ$, c) monomer $M$, d) pendant vinyl groups $V_p$
Figure A. 27 Simulation results corresponding to the final parameter estimation step for concentrations of a) carbon-centered initiator radicals \( I^* \), b) phosphorous-centered initiator radicals \( \tilde{I}^* \), and c) phosphorous-centered radicals attached to large molecules \( I_P^* \)
Figure A. 28 Simulation results corresponding to the final parameter estimation step for concentrations of a) regular radical ends $R^*$, b) radicals at the branch points $B^*$, c) cyclized radical ends $C^*$, and d) tertiary radicals $T^*$

Standard deviations for the estimated parameters were calculated based on a linearization approach (see Table A.5). The seven parameter estimates, which are significant based on their 95% confidence intervals, are shown in bold. The corresponding correlation matrix for the parameter estimates is shown in Table A.6. For the sake of brevity, we only show those elements of the correlation matrix that are ≥ 0.8 or ≤ -0.8. The values 0.8 and -0.8 are chosen arbitrarily.
Table A. 5 List of final parameter estimates and their approximate 95% confidence intervals

<table>
<thead>
<tr>
<th>#</th>
<th>Parameters</th>
<th>Estimate</th>
<th>Standard deviation</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$v_{f,0}$</td>
<td>0.064</td>
<td>0.005</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>$v_{f,1}/v_{f,0}$</td>
<td>0.50</td>
<td>0.02</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>$\phi$</td>
<td>0.90</td>
<td>1.26</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>$f_0$</td>
<td>1.0</td>
<td>0.024</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>$A_f$</td>
<td>0.50</td>
<td>0.36</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>$k_d/k_d$</td>
<td>0.57</td>
<td>1.96</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>$\tilde{f}/f$</td>
<td>1.20</td>
<td>2.94</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>$k_{in,IMO}/k_{in,MO}$</td>
<td>0.005</td>
<td>0.12</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>$k_{in,MO}$</td>
<td>1.0 $\times 10^6$</td>
<td>3.2 $\times 10^6$</td>
<td>L mol^{-1} s^{-1}</td>
</tr>
<tr>
<td>12</td>
<td>$k_{in,IV}/k_{in,MO}$</td>
<td>1.0</td>
<td>7.43</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>$k_{p,R0}$</td>
<td>1.00 $\times 10^4$</td>
<td>7.3 $\times 10^4$</td>
<td>L mol^{-1} s^{-1}</td>
</tr>
<tr>
<td>16</td>
<td>$A_p/A_b$</td>
<td>0.95</td>
<td>0.41</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>$v_{fc,p}/v_{fc,b}$</td>
<td>0.87</td>
<td>0.14</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>$k_{p,C0}/k_{p,R0} = k_{D,C0}/k_{D,R0}$</td>
<td>0.50</td>
<td>4.21</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>$k_{p,R0}/k_{p,R0} = k_{D,R0}/k_{D,R0}$</td>
<td>0.0483</td>
<td>0.31</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>$k_{p,T0}/k_{p,R0} = k_{D,T0}/k_{D,R0}$</td>
<td>5.00 $\times 10^{-4}$</td>
<td>4.6 $\times 10^{-3}$</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>$k_{DR,0}/k_{PR,0}$</td>
<td>1.0</td>
<td>2.03</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>$A_p$</td>
<td>0.59</td>
<td>0.38</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td>$v_{fc,b}/v_{fc,RR}$</td>
<td>0.53</td>
<td>0.13</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>$k_{C,0}/k_{p,R0}$</td>
<td>3.35</td>
<td>3.18</td>
<td>mol L^{-1}</td>
</tr>
<tr>
<td>27</td>
<td>$k_{bb,b}/k_{p,R0}$</td>
<td>1.00 $\times 10^{-3}$</td>
<td>8.65 $\times 10^{-3}$</td>
<td>mol L^{-1}</td>
</tr>
<tr>
<td>30</td>
<td>$k_{in,II}/k_{in,II}$</td>
<td>6.98 $\times 10^8$</td>
<td>3.21 $\times 10^{10}$</td>
<td>L mol^{-1} s^{-1}</td>
</tr>
<tr>
<td>33</td>
<td>$k_{in,II}/k_{in,II}$</td>
<td>1.15</td>
<td>150.6</td>
<td>-</td>
</tr>
<tr>
<td>34</td>
<td>$k_{C,RR,0}$</td>
<td>1.69 $\times 10^8$</td>
<td>2.2 $\times 10^9$</td>
<td>L mol^{-1} s^{-1}</td>
</tr>
<tr>
<td>35</td>
<td>$A_{t,RR}$</td>
<td>1.00</td>
<td>0.69</td>
<td>-</td>
</tr>
<tr>
<td>36</td>
<td>$v_{fc,RR}$</td>
<td>0.086</td>
<td>0.019</td>
<td>-</td>
</tr>
<tr>
<td>37</td>
<td>$R_{rd}$</td>
<td>7.31</td>
<td>79.1</td>
<td>L mol^{-1}</td>
</tr>
<tr>
<td>38</td>
<td>$k_{LCC0}/k_{LRR0}$</td>
<td>1.50</td>
<td>21.5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>( k_{t,BB0}/k_{t,RR0} )</td>
<td>0.01</td>
<td>0.27</td>
<td>-</td>
</tr>
<tr>
<td>---</td>
<td>-----------------------------</td>
<td>------</td>
<td>------</td>
<td>---</td>
</tr>
<tr>
<td>39</td>
<td>( k_{t,TT0}/k_{t,RR0} )</td>
<td>0.088</td>
<td>1.30</td>
<td>-</td>
</tr>
</tbody>
</table>
Table A. 6 Correlation matrix elements greater than or equal to arbitrary value of 0.8 or less than or equal to arbitrary value of -0.8

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\frac{\nu_{f,1}}{\nu_{f,0}}$</th>
<th>$A_f$</th>
<th>$\frac{\nu_{fc,p}}{\nu_{fc,b}}$</th>
<th>$\frac{k_{p,RO}}{k_{p,R0}}$</th>
<th>$\frac{k_{p,T0}}{k_{p,R0}}$</th>
<th>$A_b$</th>
<th>$\frac{\nu_{fc,b}}{\nu_{f,IRR}}$</th>
<th>$k_{C,0}$</th>
<th>$k_{B,0}$</th>
<th>$k_{t,IRR,0}$</th>
<th>$k_{L,RR0}$</th>
<th>$k_{L,BB0}$</th>
<th>$k_{L,T0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_{f,0}$</td>
<td>-0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_f$</td>
<td>-0.82</td>
<td>-0.80</td>
<td>-0.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{in,IM0}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{in,IM0}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{p,R0}$</td>
<td>-0.94</td>
<td>-0.96</td>
<td>-0.80</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\frac{A_p}{A_b}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\frac{\nu_{fc,p}}{\nu_{fc,b}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{p,CO}$</td>
<td>$k_{b,CO}$</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{p,R0}$</td>
<td>$k_{b,R0}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{p,BO}$</td>
<td>$k_{b,BO}$</td>
<td>0.86</td>
<td>-0.84</td>
<td>-0.91</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{p,T0}$</td>
<td>$k_{b,T0}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{p,R0}$</td>
<td>$k_{b,R0}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\frac{k_{B,0}}{k_{p,R0}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{t,RR}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_{t,RR}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_{td}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\frac{k_{L,CC0}}{k_{L,RR0}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{L,BB0}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_{L,T0}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

182
Appendix B

Mathematical Modeling for 1,6-Hexanediol Diacrylate

Photopolymerization in Presence of Oxygen**

B.1 Abstract

A dynamic model is proposed for photopolymerization of 1,6-hexanediol diacrylate (HDDA) with the bifunctional initiator bis-acylphosphine oxide (BAPO) in the presence of oxygen. The model tracks time-varying concentrations of monomer, oxygen, and different radical end groups using ordinary differential equations. An analytical expression is derived for mass-transfer of oxygen. Oxygen-related parameters are estimated using real-time FTIR vinyl conversion data, which were collected during polymerization of HDDA films with different thicknesses, initiator concentrations, and light intensities. The resulting model and parameter estimates provide good predictions for experiments involving thin films up to 12 µm, with BAPO levels ranging from 1 wt% to 4 wt% and relatively low light intensities (200 to 1000 W/m²).

** This chapter with coauthors Anh-Duong Dieu Vo, Kaveh Abdi, Jurre F.U. Tak, Luuk van der Velden, Robin X.E. Willemse, Marjolein N. van der Linden, Piet D. Iedema, Kimberley B. McAuley has been accepted for publication in the Canadian Journal of Chemical Engineering.
B.2 Introduction

Free-radical photopolymerization has been used in a variety of industrial applications due to its advantages in manufacturing cost, efficiency, and environmental friendliness.\textsuperscript{1,2} For example, photopolymerization has been used to produce films, coatings, and materials for optical, dental, and microelectronic applications.\textsuperscript{3,4} The current appendix focuses on the application of light-induced polymerization of the divinyl monomer 1,6-hexanediol diacrylate (HDDA), which is of interest for high-speed printing for large colour posters. The resulting polymer forms a crosslinked network that protects the coloured pigments and holds them in place.

Different mathematical models have been developed to provide understanding about the influence of polymerization kinetics, variable film thickness, and oxygen inhibition in photopolymerization processes.\textsuperscript{5-22} Prior mathematical models have been developed to describe the photopolymerization of HDDA to produce acrylate films.\textsuperscript{20-22} An early model developed by Iedema was proposed for photocuring of HDDA in the absence of oxygen.\textsuperscript{20} Experimental data obtained using real-time transmission Fourier-Transform Infrared Reflection (trans-FTIR)\textsuperscript{23} were used to tune the kinetic parameters used in the model, which predicts the conversion of vinyl groups. A second modeling work focused on the thermomechanical properties and topology of HDDA polymer networks using molecular dynamics and graph theory tools.\textsuperscript{24}

Recently, Abdi et al. developed a more detailed kinetic model for photopolymerization of HDDA with the bifunctional initiator bis-acylphosphine oxide (BAPO) in the absence of oxygen.\textsuperscript{22} This model accounts for: \textit{i}) multiple types of vinyl groups; \textit{ii}) branching, backbiting, and cyclization reactions, and \textit{iii}) different types of radicals that are formed by the decomposition of the BAPO initiator. A systematic parameter ranking and estimation methodology\textsuperscript{25,26} revealed the importance of branching, backbiting, and cyclization reactions and helped to determine which reactions should be treated as diffusionally-controlled. Experimental data for monomer conversion and overall vinyl group conversion, obtained using H-NMR and FTIR, respectively, were used to support parameter
estimation. Using the data from 15 experimental runs, the 30 most-important parameters were
selected and estimated using a weighted-least-squares objective function. Abdi’s model provides
good predictions for overall vinyl group conversion in experiments with light intensities up to 1350
W/m², but underpredicts the overall conversion for runs with low light intensities (e.g., 115 W/m²).
Abdi et al. suggest that this discrepancy may be due to oxygen contamination and the fact that
oxygen inhibition was neglected in their model.22

Oxygen is a well-known inhibitor for free-radical polymerization processes. Oxygen from
the air diffuses into the reacting mixture and consumes free radicals to form peroxidic radicals.8–
19,21 These radicals are relatively stable and tend not to propagate with other end groups, so that
polymerization rates decrease significantly. Furthermore, peroxidic radicals can terminate with
other radicals in the polymerization, resulting in lower polymerization rates.8,12,13,17,19 The resulting
polymeric film therefore has lower vinyl-group conversion than when no oxygen is present and
similar reaction conditions are used. Iedema et al. and O’Brien and Bowman developed models for
photopolymerization of a HDDA film in the presence of oxygen, accounting for oxygen inhibition
and for diffusion of oxygen and monomer.10,21 Their models account for spatial variations of oxygen
concentration, as well as of other species in the system. Iedema’s model also accounts for local
swelling/shrinkage of film during photopolymerization of HDDA. However, the effects of oxygen
inhibition were accounted for in a simplified manner (i.e., by assuming peroxidic macroradicals are
nonreactive). Their models also did not account for the two different types of vinyl groups involved
in HDDA polymerization (i.e., vinyl groups on unreacted monomer molecules and pendant vinyl
groups attached to the polymer network). Side reactions such as cyclization and backbiting were
also ignored. The parameter values used in these studies were obtained from a similar acrylate
photopolymerization process (i.e., photopolymerization of 2-hydroxyethyl methacrylate), or by
fitting to experimental data obtained via real-time FTIR spectroscopy.
Improved models are desired to better understand crosslinking HDDA photopolymerization during high-speed printing and to assist in designing experiments aimed at process improvements. In the current appendix, the objective is to develop and test an updated model for photopolymerization of HDDA with the bifunctional initiator BAPO that accounts for oxygen. A reaction mechanism is proposed to account for reactions between oxygen and different radical species, and for subsequent reactions involving peroxidic radicals (i.e., propagation reactions, termination reactions, and backbiting reactions). The proposed material balances in this model are then built on the oxygen-free model equations developed by Abdi et al.\textsuperscript{22}, and are updated to account for oxygen-related reactions. To keep the model relatively simple, we focus on the average concentrations of oxygen and other species over the depth of the film (i.e., concentration gradients of oxygen and other chemical species and end groups within the polymer film are ignored). An algebraic expression is derived to account for mass-transfer of oxygen. Another objective of the current study is to obtain estimates of the five oxygen-related kinetic and diffusion parameters that appear in the proposed model. These parameter estimates are obtained from experimental data using a systematic parameter ranking and estimation methodology.\textsuperscript{25,26} The resulting parameter estimates are important because they will be used as initial values in a future modelling study that accounts for spatial variations in concentration of oxygen and other chemical species.

The remainder of this appendix is organized as follows. First, experimental data are described, which are used to support parameter estimation and model testing. Then, the proposed reaction mechanism is described and updated dynamic material balances are proposed to account for the influence of oxygen. Parameter estimation results are then presented for five new parameters related to oxygen reaction kinetics and mass transfer. Fits to experimental data are shown and the predictive ability of the proposed model is assessed using additional data not used for parameter
estimation. Finally, recommendations are made for development of a future model that accounts explicitly for spatial variations (i.e., gradients in species concentrations).

**B.3 Experimental Data**

Vinyl-group conversion data were collected using real-time FTIR spectroscopy during polymerization of HDDA films with different thicknesses, using different initiator concentrations and light intensities. At the onset of polymerization, a film contains nearly pure 1,6-hexanediol diacrylate (HDDA) monomer (Sigma Aldrich) with added bis(acylphosphine oxide) (BAPO) initiator (BASF) and Tegorad 2250 surfactant (Evonik, 2 wt%). Dissolved oxygen from the surrounding air was also present. Photopolymerization was started by exposure of the film to light from a lamp with a peak wavelength at 405 nm (provided by a FireFly FF200 50x20 light source from Phoseon Technology). Real-time infrared spectra were recorded using a Bruker Vertex 80 FT-IR spectrometer with rapid scan option at a frequency of ten or twenty spectra per second and a spectral resolution of 4 cm\(^{-1}\). The extinction coefficient \(\varepsilon = 135.7 \text{ m}^3\text{mol}^{-1}\text{m}^{-1}\) for BAPO was determined by recording the light absorption spectrum of BAPO. Experimental details are provided elsewhere.\(^{21,23}\)

Data were collected from two sets of experiments. A series of 12 experimental runs was conducted for polymerization of HDDA films with thickness \(z_f\) of \(\sim 12 \mu m\) and initiator concentration of \([BAPO] = 4 \text{ wt}\%\). Light intensity \(I_{l,0}\) ranging between 200 and 6000 W/m\(^2\) were used in these experiments. The second set of experiments arises from a factorial design involving three operating factors: \(i\) film thickness with target values of 8 \(\mu m\) and 16 \(\mu m\), \(ii\) initiator concentration with target values of 1 wt% and 4 wt%, and \(iii\) light intensity with target values of 1000 W/m\(^2\) and 5000 W/m\(^2\). For the first set of experimental runs, two sets of replicate data are available for runs with light intensities of 1000 W/m\(^2\) and 2000 W/m\(^2\), respectively. For the factorial-designed experiments, 24 runs were conducted using 8 sets of operating conditions (i.e., 3 replicate runs at each of the 8 conditions). Details about the computation of vinyl-group conversions from the real-time FT-IR measurements are provided elsewhere.\(^{23}\)
Eight runs from the described dataset are used for parameter estimation in the current appendix. The selected runs include 3 runs with light intensities of 200, 500, and 750 W/m² from the first series of experiment, and 5 runs from the factorial series, whose targeted settings were \( z_f = 8 \mu m, I_{0,0} = 1000 \text{ W/m}^2 \), and \([BAPO]\) = 1 wt% or 4 wt%. These 8 runs were selected because of their relatively thin film thicknesses and low light intensities. The assumption of relatively uniform oxygen concentration (assumption B.1.24 in Table B.1) is expected to be better for these experiments than for experiments with thicker films and high light intensities.\(^{13,14,21}\) The other experimental runs are saved for model validation and assessment.

**B.4 Model Development**

**B.4.1 Reaction Mechanism**

Abdi et al. proposed an oxygen-free reaction scheme for photopolymerization of HDDA with the bifunction initiator BAPO.\(^{22}\) **Figure B.1a** shows the structure of BAPO (denoted as \( I \)), which has two weak phosphorous-carbon bonds shown in red. Under UV radiation, one of these bonds is broken, generating a carbon-centered initiator radical \( I^\cdot \) and a more-reactive phosphorous-centered initiator radical \( \tilde{I}^\cdot \). In a later reaction, as shown in Figure 1b, the remaining carbon-phosphorous bond can decompose to produce another phosphorous-centered initiator radical \( \tilde{I}_p^\cdot \) and a carbon-centered initiator radical \( I^\cdot \). We assume that this remaining bond only decomposes after the radical on \( \tilde{I}^\cdot \) is consumed (see assumption B.1.5 in Table B.1).
Figure B. 1 Decomposition of a) initiator \( I \) (BAPO) and b) the remaining carbon-phosphorous bond in \( \tilde{I} \)\textsuperscript{22}
The radicals $I^\ast$, $I^\tilde{\ast}$ and $I_P^\ast$ initiate polymerization by reacting with HDDA monomer. Each HDDA molecule has two vinyl end groups. As shown in Figure B.2a, when a monomer molecule reacts with a radical (i.e., via initiation or propagation reactions), one of these vinyl groups is consumed. In the proposed mechanism, the resulting radical and the remaining unreacted double-bond are grouped together and denoted by the end group $R^{=\ast}$ (see assumption B.1.8 in Table B.1). These radical end groups can undergo cyclization reactions and backbiting reactions to generate the cyclic radical $C^\ast$ and the tertiary radical $T^\ast$, respectively, as shown in Figure B.2b and B.2c. In the proposed mechanism, whenever an $R^{=\ast}$ end group is consumed via propagation, the vinyl group associated with $R^{=\ast}$ becomes a pendant vinyl group $V_P$. As shown in Figure B.2d, subsequent consumption of pendant vinyl groups results in branch-point radicals $B^\ast$. A complete list of oxygen-free reactions proposed by Abdi et al. is provided in the Supplementary Information and elsewhere. This overall reaction scheme accounts for decomposition of BAPO, initiation reactions involving $I^\ast$, $I^\tilde{\ast}$ and $I_P^\ast$, as well as propagation, branching, cyclization, backbiting, and termination reactions.
Table B. 1 Model assumptions

<table>
<thead>
<tr>
<th>B.1.1</th>
<th>The initial monomer is a film. This thin film layer has constant volume and is spatially uniform during polymerization.</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.1.2</td>
<td>Temperature is constant at 25 °C.</td>
</tr>
<tr>
<td>B.1.3</td>
<td>The light intensity is uniform within the small film thickness.</td>
</tr>
<tr>
<td>B.1.4</td>
<td>The initiator moieties remaining after initiator decomposition have the same absorbance properties as the initiator.</td>
</tr>
<tr>
<td>B.1.5</td>
<td>Radicals of type $\tilde{I}^\ast$ are consumed by initiation reactions and termination reactions before their weak carbon-phosphorous bond has an opportunity to break. This assumption is valid because the rate of consumption of the free radical on $\tilde{I}^\ast$ is typically $10^9$ times faster than the rate of dissociation of the weak carbon-phosphorous bond.</td>
</tr>
<tr>
<td>B.1.6</td>
<td>In this terminal model, reaction rates depend only on the type of end groups on a polymer chain. Penultimate effects are neglected.</td>
</tr>
<tr>
<td>B.1.7</td>
<td>The effect of stabilizer is neglected.</td>
</tr>
<tr>
<td>B.1.8</td>
<td>End groups of type $R^\ast$ contain a vinyl group and a free radical. We assume that the short-lived free radical is consumed much more rapidly than the corresponding vinyl group. When the radical on $R^\ast$ is consumed by propagation, branching, backbiting and termination reactions, the corresponding vinyl group on $R^\ast$ becomes a pendant vinyl group that can participate in subsequent reactions. This assumption is valid because the rate of consumption of a pendant vinyl group is $\sim 10^5$ times slower than the rate of consumption of the corresponding radical.</td>
</tr>
<tr>
<td>B.1.9</td>
<td>Chain-transfer to monomer and chain-transfer to polymer are neglected.</td>
</tr>
<tr>
<td>B.1.10</td>
<td>Carbon-centred and phosphorous-centred initiator radicals have similar relative propensities when reacting with pendant vinyl groups compared with monomeric vinyl groups. (i.e., $\frac{k_{in,I^Vp_0}}{k_{in,I^M_0}} = \frac{k_{in,I^Vp_0}}{k_{in,I^M_0}}$).</td>
</tr>
<tr>
<td>B.1.11</td>
<td>All phosphorous-centered radicals have the same kinetically-controlled reactivity, independent of whether they are generated by initiator decomposition or decomposition of initiator moieties $\tilde{I}$ attached to polymer molecules (i.e., $k_{in,I^M_0} = k_{in,I^pM_0}$ and $k_{in,I^Vp_0} = k_{in,I^pVp_0}$).</td>
</tr>
<tr>
<td>B.1.12</td>
<td>The kinetically-dependent rate coefficient for self-termination reactions involving two phosphorous-centered radicals $\tilde{I}^\ast$ and $I_\tilde{p}$ are the same (i.e., $k_{in,I^0} = k_{in,I^pI^p}$).</td>
</tr>
<tr>
<td>B.1.13</td>
<td>Self-termination reactions for initiator radicals (i.e., $I^\ast$, $\bar{I}^\ast$, and $I_p^\ast$) are ignored so that the SSH can be readily applied to compute the concentrations of these short-lived radicals. Cross-termination reactions between $I^\ast$ and $\bar{I}^\ast$ and between $I^\ast$ and $I_p^\ast$ are ignored in the balance on $I^\ast$.</td>
</tr>
<tr>
<td>B.1.14</td>
<td>All macroradicals have the same relative propensities for branching and propagation reactions. (i.e., $k_{b,R0} = k_{b,D0} = k_{b,C0} = k_{b,T0} = k_{b,O0}$).</td>
</tr>
<tr>
<td>B.1.15</td>
<td>Cyclization reactions may occur between a radical end $R=C^\ast$ and its vinyl group, but not with other vinyl groups on the polymer.</td>
</tr>
<tr>
<td>B.1.16</td>
<td>The only radical ends that participate in backbiting reactions are $R=C^\ast$, $C^\ast$, and $O^\ast$. Backbiting generates tertiary radicals $T^\ast$.</td>
</tr>
<tr>
<td>B.1.17</td>
<td>Termination reactions involving two secondary macroradicals are termination-by-combination reactions.</td>
</tr>
<tr>
<td>B.1.18</td>
<td>Termination reactions involving tertiary radical $T^\ast$ and other radicals are termination-by-disproportionation reactions.</td>
</tr>
<tr>
<td>B.1.19</td>
<td>The reactivity of terminal vinyl groups resulting from termination by disproportionation is the same as the reactivity of pendant vinyl groups.</td>
</tr>
<tr>
<td>B.1.20</td>
<td>Initiation reactions involving $I^\ast$, $\bar{I}^\ast$ and $I_p^\ast$ and reactions between radicals and oxygen are kinetically-controlled rather than diffusionally-controlled. Backbiting and cyclization are also kinetically-controlled. Propagation and branching reactions, and termination reactions between macromolecules are diffusionally-controlled.</td>
</tr>
<tr>
<td>B.1.21</td>
<td>Reaction diffusion, which influences termination rates, occurs predominantly by consumption of monomer vinyl groups rather than pendant vinyl groups.</td>
</tr>
<tr>
<td>B.1.22</td>
<td>The free-volume parameters within the following classes of reactions are equal: i) propagation reactions involving a small monomer molecule and a free-radical on a large molecule, ii) branching reactions involving two large molecules, and iii) termination reactions involving two large molecules. The same free-volume parameters are used within each class of reaction to limit the number of parameters that appear in the model. Further, the free-volume parameters for initiation decompositions $f$ and $\bar{f}$ are the same.</td>
</tr>
<tr>
<td>B.1.23</td>
<td>The free volume fraction of the film $v_f$ is assumed to change linearly with monomer conversion $X_M$.</td>
</tr>
<tr>
<td>B.1.24</td>
<td>Concentration gradients of oxygen and other chemical species and end groups within the polymer film are ignored.</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>B.1.25</td>
<td>All radicals have the same relative propensities for oxygen incorporation compared to their corresponding propensities for propagation or initiation with monomer (i.e., ( \frac{k_{O_2,R_0}}{k_{p,R_0}} = \frac{k_{O_2,B_0}}{k_{p,B_0}} = \frac{k_{O_2,C_0}}{k_{p,C_0}} = \frac{k_{O_2,T_0}}{k_{p,T_0}} = \frac{k_{O_2,I_0}}{k_{in,IM_0}} = \frac{k_{O_2,I_0}}{k_{in,IM_0}} = \frac{k_{O_2,I_0}}{k_{in,IM_0}} )).</td>
</tr>
<tr>
<td>B.1.26</td>
<td>Peroxidic radicals have the same relative propensities for backbiting and propagation reactions compared to the corresponding backbiting and propagation reactions of ( R^* ) (i.e., ( \frac{k_{bb,O_0}}{k_{p,O_0}} = \frac{k_{bb,R_0}}{k_{p,R_0}} )).</td>
</tr>
<tr>
<td>B.1.27</td>
<td>Hydroperoxide end groups, which are resulted from backbiting reactions of peroxidic radicals, do not propagate during the time interval of experiments.</td>
</tr>
</tbody>
</table>
Table B.2 summarizes the proposed additional reactions that are included in the current model to account for the effect of oxygen and peroxidic radicals. As shown in Figure B.2e and reactions R1 to R7 in Table B.1, all reactions between radical species and oxygen produce peroxidic radical end groups, denoted by $O^\bullet$. Propagation rates of these peroxidic radicals (see reactions R8 and R9) are expected to be very slow, compared to the corresponding propagation rates for other radicals. Reaction R10 shows that peroxidic radicals can abstract labile hydrogen atoms via backbiting to produce tertiary radicals $T^\bullet$. Reactions R11 to R18 accounts for termination reactions involving $O^\bullet$ end groups.
Figure B. 2 Formation of different radical end groups including: a) regular radical ends $R^{\ast\ast}$, b) cyclized radical ends $C^\ast$, c) tertiary radicals $T^\ast$, d) branch-point radicals $B^\ast$, and e) peroxidic radicals $O^\ast$. Ellipses are used to highlight the end group of interested.
**Table B. 2 - List of reactions accounting for effect of oxygen on photopolymerization of HDDA**

<table>
<thead>
<tr>
<th>Reaction Description</th>
<th>Reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oxygen incorporation</strong></td>
<td>$I^* + O_2 \xrightarrow{k_{O2I}} O^*$</td>
</tr>
<tr>
<td></td>
<td>$\bar{I}^* + O_2 \xrightarrow{k_{O2\bar{I}}} O^* + \bar{I}$</td>
</tr>
<tr>
<td></td>
<td>$I_p^* + O_2 \xrightarrow{k_{O2IP}} O^*$</td>
</tr>
<tr>
<td></td>
<td>$R^{=<em>} + O_2 \xrightarrow{k_{O2R}} O^</em> + V_P$</td>
</tr>
<tr>
<td></td>
<td>$B^* + O_2 \xrightarrow{k_{O2B}} O^*$</td>
</tr>
<tr>
<td></td>
<td>$C^* + O_2 \xrightarrow{k_{O2C}} O^*$</td>
</tr>
<tr>
<td></td>
<td>$T^* + O_2 \xrightarrow{k_{O2T}} O^*$</td>
</tr>
<tr>
<td><strong>Propagation of peroxidic radicals (slow)</strong></td>
<td>$O^* + M \xrightarrow{2kp,0} R^{=*}$</td>
</tr>
<tr>
<td></td>
<td>$O^* + V_P \xrightarrow{kb,0} B^*$</td>
</tr>
<tr>
<td><strong>Backbiting of peroxidic radicals</strong></td>
<td>$O^* \xrightarrow{kb_{b,0}} T^*$</td>
</tr>
<tr>
<td><strong>Termination of peroxidic radicals</strong></td>
<td>$O^* + O^* \xrightarrow{k_{t,00}} D + O_2$</td>
</tr>
<tr>
<td></td>
<td>$O^* + I^* \xrightarrow{k_{t,10}} D$</td>
</tr>
<tr>
<td></td>
<td>$O^* + \bar{I}^* \xrightarrow{k_{t,\bar{I}0}} D + \bar{I}$</td>
</tr>
<tr>
<td></td>
<td>$O^* + I_p^* \xrightarrow{k_{t,IP,0}} D$</td>
</tr>
<tr>
<td></td>
<td>$O^* + R^{=*} \xrightarrow{k_{t,R0}} D + V_P$</td>
</tr>
<tr>
<td></td>
<td>$O^* + B^* \xrightarrow{k_{t,B0}} D$</td>
</tr>
<tr>
<td></td>
<td>$O^* + C^* \xrightarrow{k_{t,C0}} D$</td>
</tr>
<tr>
<td></td>
<td>$O^* + T^* \xrightarrow{k_{t,T0}} D$</td>
</tr>
</tbody>
</table>
B.4.2 Model Equations

The proposed model assumes that the concentrations of all chemical species and reactive endgroups are uniform within the polymer film (see assumption B.1.24 in Table B.1). As a result, their time-varying concentrations are tracked using 13 ordinary differential equations (ODEs), which are shown in Table B.3. Equation (B.3.1) in Table B.3 is a dynamic material balance for initiator, along with its analytical solution.\textsuperscript{22} In equation (B.3.1), the factor 2 appears on the right-hand side because because each BAPO initiator molecule is bifunctional. Equations (B.3.2) to (B.3.11) are updated versions of ODEs derived by Abdi et al. with additional terms, which are highlighted in yellow, accounting for reactions involving peroxidic radicals. For example, equation (B.3.2) is a material balance on unreacted monomer. The first seven terms on the right-hand side appear in Abdi’s oxygen-free model, accounting for monomer consumption by initiation reactions (i.e., with $I^\cdot$, $\tilde{I}^\cdot$ and $I_P^\cdot$) and propagation reactions (i.e., with $R^\approx$, $B^\cdot$, $C^\cdot$, and $T^\cdot$). The updated version of equation (B.3.2) shown here has an additional term accounting for slow consumption of monomer by peroxidic radicals $O^\cdot$. ODEs (B.3.3) to (B.3.11) account for time-varying concentrations of initiator radicals, pendant vinyl groups, and free radical end groups.

Equations (B.3.12) and (B.3.13) are new material balances (highlighted), which account for time-varying concentrations of oxygen and peroxidic radicals. The first term on the right-hand side of equation (B.3.12) accounts for mass-transfer of oxygen where $A$ (m$^2$ kg$^{-1}$) is the surface area of the film per unit mass, $k_{m,O_2}$ (kg m$^{-2}$ s$^{-1}$) is the mass-transfer coefficient of $O_2$ into the film from the surrounding air, and $[O_2]^\ast = 1.17 \times 10^{-3}$ (mol L$^{-1}$)$^{32}$ is the concentration of $O_2$ in the film that would be in equilibrium with the air. The mass-transfer coefficient $k_{m,O_2}$ is computed from the diffusivity of oxygen $D_{O_2}$ in the polymer, the film density $\rho$ (kg m$^{-3}$), and the film thickness $z_f$. The expression for $k_{m,O_2}$ shown in Table B.3 is a new expression, which is derived in Appendix S4 in the Supplementary Information. This derivation is based on the analytical solution to a partial differential equation (PDE) describing heat conduction or diffusion in a flat plate.\textsuperscript{33} The diffusivity
of oxygen in the polymer $D_{O_2}$ (m$^2$ s$^{-1}$) is computed using

$$D_{O_2} = D_{O_2-HDDA} \exp \left[ -A_{O_2} \left( \frac{1}{v_f} - \frac{1}{v_{f0}} \right) \right],$$

where $D_{O_2-HDDA}$ is the diffusivity of oxygen in pure HDDA monomer, and $A_{O_2}$ is the free-volume parameter for $D_{O_2}$. Both $D_{O_2-HDDA}$ and $A_{O_2}$ are model parameters that require estimation in the current study. The next seven terms on the right-hand side of equation (B.3.12) account for depletion of oxygen as it reacts with free radicals in the polymer. The last term on the right-hand side accounts for oxygen regeneration from self-termination reactions of peroxidic radicals (see reaction R11).

Equation (B.3.13) describes the time-varying concentration of peroxidic radicals within the polymer film. As shown in the first seven terms on the right-hand side of equation (B.3.13), all radical species produce the end group $O^\bullet$ when they react with oxygen (see assumption B.1.23 in Table B.1). The next three terms on the right-hand side of equation (B.3.13) show that peroxidic radicals can participate (slowly) in propagation, branching, and backbiting reactions. The last eight terms on the right-hand side of equation (B.3.13) account for termination of peroxidic radicals with other radicals.
### Table B. 3 Material balances for chemical species and functional groups

| B.3.1 | $\frac{d[I]}{dt} = -2k_d[I]$ which is solved to give: $[I] = [I]_0 \exp (-2k_d t)$ |
| B.3.2 | $\frac{d[M]}{dt} = -2k_{in,IM} [I^*][M] - 2k_{in,JM} [\tilde{I}^*][M] - 2k_{in,JpM} [I^*][M] - 2k_{p,R} [R^=*][M] - 2k_{p,B} [B^*][M] - 2k_{p,C} [C^*][M] - 2k_{p,T} [T^*][M] - 2k_{p,O} [O^*][M]$ |
| B.3.3 | $\frac{d[I^*]}{dt} = 2f k_d [I] + \bar{f} k_{\tilde{d}} [\tilde{I}] - 2k_{in,IM} [I^*][M] - k_{in,IVp} [I^*][V_p] - k_{t,IR} [I^*][R^=*] - k_{t,IB} [I^*][B^*] - k_{t,IC} [I^*][C^*] - k_{t,IT} [I^*][T^*] - k_{t,IO} [I^*][O^*] - k_{O_2,\bar{I}^*} [O_2]$ which becomes, after applying SSH: $[I^*] = \frac{2f k_d [I] + \bar{f} k_{\tilde{d}} [\tilde{I}]}{(2k_{in,IM} [M] + k_{in,IVp} [V_p] + k_{t,IR} [R^=*] + k_{t,IB} [B^*] + k_{t,IC} [C^*] + k_{t,IT} [T^*] + k_{t,IO} [O^*] + k_{O_2,\bar{I}^*} [O_2])}$ |
| B.3.4 | $\frac{d[\tilde{I}^*]}{dt} = 2f k_d [I] - 2k_{in,IM} [\tilde{I}^*][M] - k_{in,IVp} [\tilde{I}^*][V_p] - k_{t,IR} [\tilde{I}^*][R^=*] - k_{t,IB} [\tilde{I}^*][B^*] - k_{t,IC} [\tilde{I}^*][C^*] - k_{t,IT} [\tilde{I}^*][T^*] - k_{t,IO} [\tilde{I}^*][O^*] - k_{O_2,\bar{I}^*} [O_2]$ which becomes, after applying SSH: |
\[ [\tilde{I}^*] = \frac{2f k_d [I]}{+ 2k_{inJM} [M] + k_{in,lv_P} [V_P]} + k_{t,IR} [R^=\star] + k_{t,IB} [B^\star] + k_{t,JC} [C^\star] + k_{t,IT} [T^\star] + k_{t,IR} [I^\star] + k_{t,IO} [O^\star] + k_{O_2,I} [O_2] \]

\[
\frac{d[I^*]}{dt} = \tilde{f} k_d [\tilde{I}] - 2k_{in,JPM} [I^*][M] - k_{in,Jv_P} [I^*][V_P] - k_{t,JpR} [I^*][R^=\star] - k_{t,JpB} [I^*][B^\star] - k_{t,JpC} [I^*][C^\star] - k_{t,JpT} [I^*][T^\star] - k_{t,JlP} [I^*][I^\star] - k_{t,JpO} [I^*][O^\star] - k_{O_2,Jp} [I^*][O_2]
\]

which becomes, after applying SSH:

\[
[I^*] = \frac{\tilde{f} k_d [\tilde{I}]}{+ 2k_{in,JPM} [M] + k_{in,Jv_P} [V_P] + k_{t,JpR} [R^=\star] + k_{t,JpB} [B^\star] + k_{t,JpC} [C^\star] + k_{t,JpT} [T^\star] + k_{t,JlP} [I^*] + k_{t,JpO} [O^\star] + k_{O_2,Jp} [O_2]}
\]

\[
\frac{d[\tilde{I}]}{dt} = -k_d [\tilde{I}] + 2k_{in,JM} [I^*][M] + k_{in,lv_P} [I^*][V_P] + k_{t,IR} [I^*][R^=\star] + k_{t,IB} [I^*][B^\star] + k_{t,JC} [I^*][C^\star] + k_{t,IT} [I^*][T^\star] + k_{t,II} [I^*][I^\star] + k_{t,JpO} [I^*][O^\star] + k_{O_2,I} [I^*][O_2]
\]
**B.3.7**

\[
\frac{d[V_p]}{dt} = 2k_{p,R}[R^=][M] + k_{bb,R}[R^=] + k_{t,RR}[R^=][R^=] + k_{t,RB}[R^=][B^*]
\]

\[
+ k_{t,RC}[R^=][C^*] + 2k_{t,RT}[R^=][T^*] + k_{t,BT}[B^*][T^*]
\]

\[
+ k_{t,CT}[C^*][T^*] + k_{t,TT}[T^*][T^*] + k_{t,IR}[I^*][R^=]
\]

\[
+ k_{t,TR}[I^*][T^*] + k_{t,IPR}[I^*][R^=] + k_{t,IT}[I^*][T^*]
\]

\[
+ k_{t,IT}[I^*][T^*] + k_{t,IT}[I^*][T^*] - k_{in,IVP}[I^*][V_p]
\]

\[
- k_{in,IVP}[I^*][V_p] - k_{in,IVP}[I^*][V_p] - k_{b,b}[B^*][V_p]
\]

\[
- k_{b,c}[C^*][V_p] - k_{b,T}[T^*][V_p] - k_{b,o}[O^*][V_p]
\]

\[
+ k_{O_2,R}[R^=][O_2] + k_{t,RO}[R^=][O^*] + k_{t,TO}[T^*][O^*]
\]

**B.3.8**

\[
\frac{d[R^=]}{dt} = 2k_{in,IM}[I^*][M] + 2k_{in,IM}[I^*][M] + 2k_{in,IPM}[I^*][M]
\]

\[
+ 2k_{p,B}[B^*][M] + 2k_{p,c}[C^*][M] + 2k_{p,T}[T^*][M]
\]

\[
+ 2k_{p,o}[O^*][M] - k_{b, R}[R^=][V_p] - k_{c}[R^=] - k_{b,b, R}[R^=]
\]

\[
- k_{t,RR}[R^=][R^=] - k_{t, RB}[R^=][B^*] - k_{t, RC}[R^=][C^*]
\]

\[
- k_{t,RT}[R^=][T^*] - k_{t, IR}[I^*][R^=] - k_{t, TR}[I^*][R^=]
\]

\[
- k_{t, IP R}[I^*][R^=] - k_{t, RO}[R^=][O^*] - k_{O_2, R}[R^=][O_2]
\]

**B.3.9**

\[
\frac{d[B^*]}{dt} = k_{in,IVP}[I^*][V_p] + k_{in,IVP}[I^*][V_p] + k_{in,IPV P}[I^*][V_p] + k_{b,b}[R^=][V_p]
\]

\[
+ k_{b,c}[C^*][V_p] + k_{b,T}[T^*][V_p] + k_{b,o}[O^*][V_p]
\]

\[
- 2k_{p,B}[B^*][M] - k_{t, RB}[R^=][B^*] - k_{t, BB}[B^*][B^*]
\]

\[
- k_{t, BC}[B^*][C^*] - k_{t, BT}[B^*][T^*] - k_{t, JB}[I^*][B^*] - k_{t, JB}[I^*][B^*]
\]

\[
- k_{t, IP B}[I^*][B^*] - k_{t, BO}[B^*][O^*] - k_{O_2, B}[B^*][O_2]
\]
\[
\frac{d[C^*]}{dt} = k_c[R^=\cdot] - 2k_{p,c}[C^*][M] - k_{b,c}[C^*][V_F] - k_{b,b,r}[C^*]
- k_{t,RC}[R^=\cdot][C^*] - k_{t,BC}[B^\cdot][C^*] - k_{t,CC}[C^*][C^*]
- k_{t,CT}[C^*][T^*] - k_{t,JC}[I^*][C^*] - k_{t,IC}[\bar{I}^*][C^*]
- k_{t,lpC}[I_P^*][C^*] - k_{t,CO}[C^*][O^*] - k_{O_2,C}[C^*][O_2]
\]

\[
\frac{d[T^*]}{dt} = k_{b,b,r}[R^=\cdot] + k_{b,b,r}[C^*] + k_{b,b,O}[O^*] - 2k_{p,T}[T^*][M] - k_{b,T}[T^*][V_F]
- k_{t,RT}[R^=\cdot][T^*] - k_{t,BT}[B^\cdot][T^*] - k_{t,CT}[C^*][T^*]
- k_{t,TT}[T^*][T^*] - k_{t,IT}[I^*][T^*] - k_{t,IR}[\bar{I}^*][T^*]
- k_{t,lpT}[I_P^*][T^*] - k_{t,TO}[T^*][O^*] - k_{O_2,T}[T^*][O_2]
\]

\[
\frac{d[O_2]}{dt} = A_{m,O_2}([O_2]^* - [O_2]) - k_{O_2,R}[R^=\cdot][O_2] - k_{O_2,B}[B^\cdot][O_2]
- k_{O_2,C}[C^*][O_2] - k_{O_2,T}[T^*][O_2] - k_{O_2,I}[I^*][O_2] - k_{O_2,\bar{I}^*}[O_2]
- k_{O_2,lp}[I_P^*][O_2] + \frac{1}{2}k_{t,oo}[O^*][O^*]
\]

where \( k_{m,O_2} = 1.2 \frac{D_{O_2} \rho \pi^2}{z_f} \) and \( D_{O_2} = D_{O_2-HDDA} \exp \left[ -A_{O_2} \left( \frac{1}{v_f} - \frac{1}{v_{f0}} \right) \right] \)

\[
\frac{d[O^*]}{dt} = + k_{O_2,I}[I^*][O_2] + k_{O_2,\bar{I}^*}[\bar{I}^*][O_2] + k_{O_2,lp}[I_P^*][O_2] + k_{O_2,R}[R^=\cdot][O_2]
+ k_{O_2,B}[B^\cdot][O_2] + k_{O_2,C}[C^*][O_2] + k_{O_2,T}[T^*][O_2]
- 2k_{p,o}[O^*][M] - k_{b,o}[O^*][V_F] - k_{b,b,O}[O^*] - k_{t,RO}[R^=\cdot][O^*]
- k_{t,B0}[B^\cdot][O^*] - k_{t,CO}[C^*][O^*] - k_{t,TO}[T^*][O^*]
- k_{t,IO}[I^*][O^*] - k_{t,\bar{I}^*}[\bar{I}^*][O^*] - k_{t,lpCO}[I_P^*][O^*]
- k_{t,oo}[O^*][O^*]
\]
**B.5 Parameter Estimation and Simulation Results**

**B.5.3 Parameter Estimation**

The proposed model requires a total of 35 parameter values for performing simulations. Among these parameters, values for 30 parameter estimates can be obtained from Abdi’s oxygen-free modelling study (see Table B.4). Table B.5 shows the additional 5 parameters that are introduced in the proposed model to account for oxygen-related phenomena. The initial guesses and bounds shown in Table B.5 were chosen from the literature or based on engineering judgement. Simplifying assumptions were made to reduce the number of parameters requiring estimation. For example, it is assumed (assumption B.1.25 in Table B.1) that all radicals have the same relative propensities for oxygen incorporation compared to their corresponding propensities for propagation or initiation reactions with monomer (i.e., \( \frac{k_{O_2,R_0}}{k_{p,R_0}} = \frac{k_{O_2,B_0}}{k_{p,B_0}} = \frac{k_{O_2,C_0}}{k_{p,C_0}} = \frac{k_{O_2,T_0}}{k_{p,T_0}} \)). It is also assumed (assumption B.1.14 in Table B.1) that peroxidic radicals have the same relative propensities for branching and propagation reactions compared with other radicals (i.e., \( \frac{k_{b,0_0}}{k_{p,0_0}} = \frac{k_{b,RO}}{k_{p,RO}} = \frac{k_{b,BO}}{k_{p,BO}} = \frac{k_{b,CO}}{k_{p,CO}} = \frac{k_{b,TO}}{k_{p,TO}} \)). In addition, it is assumed (assumption B.1.26) that peroxidic radicals have the same relative propensities for backbiting and propagation reactions compared to the corresponding backbiting and propagation reactions involving \( R_0^* \) (i.e., \( \frac{k_{bb,0_0}}{k_{p,0_0}} = \frac{k_{bb,RO}}{k_{p,RO}} \)).

The proposed model uses the following least-squares objective function for parameter estimation:

\[
J = \sum_{i=1}^{N_{X_V}} (X_{V,i} - X_{V_{meas},i})^2
\]  

(\text{B.1})

where \( X_{V,i} \) is the model prediction of the \( i^{th} \) measured value of the overall vinyl conversion \( X_{V_{meas},i} \). The total number of vinyl conversion measurements used for parameter estimation is \( N_{X_V} = 172 \), arising from 8 experimental runs used for parameter estimation. During parameter estimation, we focus more on fitting the data from the first 5 seconds of each run, because most of
the polymerization happens during this time period. Therefore, all of the data from the first 5 s in each run are used in the objective function, while only a portion of the longer-time data is used (i.e., longer-time data are sampled less frequently so that the number of long-time data points used in the objective function matches the number of data points before 5 s). In this way, the optimizer is prevented from focusing too much on the longer-time data compared to the shorter-time dynamic data.\textsuperscript{21} Prior to parameter estimation, a sensitivity-based estimability analysis was conducted to confirm that all 5 parameters are estimable from the available data.\textsuperscript{25,26}
Table B. 4 List of parameters and their estimates obtained from Abdi’s model\textsuperscript{22}

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Units</th>
<th>Parameter</th>
<th>Estimate</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>(v_{f,0})</td>
<td>(6.366 \times 10^{-2})</td>
<td>-</td>
<td>(\frac{k_{b,R0}}{k_{p,R0}})</td>
<td>(10.000 \times 10^{-1})</td>
<td>-</td>
</tr>
<tr>
<td>(\frac{v_{f,1}}{v_{f,0}})</td>
<td>(4.999 \times 10^{-1})</td>
<td>-</td>
<td>(\frac{k_{c0}}{k_{p,R0}})</td>
<td>(3.347)</td>
<td>-</td>
</tr>
<tr>
<td>(f_0)</td>
<td>(9.959 \times 10^{-1})</td>
<td>-</td>
<td>(\frac{k_{bb0}}{k_{p,R0}})</td>
<td>(1.001 \times 10^{-3})</td>
<td>-</td>
</tr>
<tr>
<td>(A_f)</td>
<td>(5.000 \times 10^{-1})</td>
<td>-</td>
<td>(\frac{k_{p,B0}}{k_{p,R0}} = \frac{k_{b,B0}}{k_{b,R0}})</td>
<td>(4.834 \times 10^{-2})</td>
<td>-</td>
</tr>
<tr>
<td>(\frac{\tilde{f}}{f})</td>
<td>(1.200)</td>
<td>-</td>
<td>(\frac{k_{p,C0}}{k_{p,R0}} = \frac{k_{b,C0}}{k_{b,R0}})</td>
<td>(5.008 \times 10^{-1})</td>
<td>-</td>
</tr>
<tr>
<td>(\phi)</td>
<td>(8.995 \times 10^{-1})</td>
<td>-</td>
<td>(\frac{k_{p,T0}}{k_{p,R0}} = \frac{k_{b,T0}}{k_{b,R0}})</td>
<td>(5.027 \times 10^{-4})</td>
<td>-</td>
</tr>
<tr>
<td>(\frac{k_{d}}{k_d})</td>
<td>(5.724 \times 10^{-1})</td>
<td>-</td>
<td>(k_{t_{in}II0})</td>
<td>(6.983 \times 10^{8})</td>
<td>L mol(^{-1})s(^{-1})</td>
</tr>
<tr>
<td>(k_{in,\overline{I}M0})</td>
<td>(1.005 \times 10^{6})</td>
<td>L mol(^{-1})s(^{-1})</td>
<td>(\frac{k_{t_{in}II0}}{k_{t_{in}II0}})</td>
<td>(1.149)</td>
<td>-</td>
</tr>
<tr>
<td>(\frac{k_{in,\overline{I}M0}}{k_{in,\overline{I}M0}})</td>
<td>(5.027 \times 10^{-3})</td>
<td>-</td>
<td>(R_{rd})</td>
<td>(7.308)</td>
<td>L mol(^{-1})</td>
</tr>
<tr>
<td>(\frac{k_{in,IVp0}}{k_{in,II0}})</td>
<td>(9.996 \times 10^{-1})</td>
<td>-</td>
<td>(A_{t,RR})</td>
<td>(10.000 \times 10^{-1})</td>
<td>-</td>
</tr>
<tr>
<td>(\frac{k_{in,IVp0}}{k_{in,II0}}) &amp; (\frac{k_{in,IVp0}}{k_{in,II0}})</td>
<td>9.996 \times 10^{-1}</td>
<td>(A_{t,RR})</td>
<td>10.000 \times 10^{-1}</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>(A_{b,R})</td>
<td>(5.911 \times 10^{-1})</td>
<td>-</td>
<td>(v_{f_c,RR})</td>
<td>(8.553 \times 10^{-2})</td>
<td>-</td>
</tr>
<tr>
<td>Symbol/Expression</td>
<td>Value</td>
<td>Units</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A_{p,R} \quad A_{b,R}$</td>
<td>$9.451 \times 10^{-1}$</td>
<td>-</td>
<td>$k_{t,RR0}$</td>
<td>$1.695 \times 10^{8}$</td>
<td>L mol$^{-1}$s$^{-1}$</td>
</tr>
<tr>
<td>$v_{fc,bR} \quad v_{fc,tRR}$</td>
<td>$5.292 \times 10^{-1}$</td>
<td>-</td>
<td>$k_{t,BB0} \quad k_{t,RR0}$</td>
<td>$1.001 \times 10^{2}$</td>
<td>-</td>
</tr>
<tr>
<td>$v_{fc,pR} \quad v_{fc,bR}$</td>
<td>$8.713 \times 10^{-1}$</td>
<td>-</td>
<td>$k_{t,CC0} \quad k_{t,RR0}$</td>
<td>1.500</td>
<td>-</td>
</tr>
<tr>
<td>$k_{p,R0}$</td>
<td>$1.001 \times 10^{4}$</td>
<td>L mol$^{-1}$s$^{-1}$</td>
<td>$k_{t,TT0} \quad k_{t,RR0}$</td>
<td>$8.795 \times 10^{2}$</td>
<td>-</td>
</tr>
</tbody>
</table>
Table B. 5 List of parameters accounting for effects of oxygen inhibition

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>Initial guess</th>
<th>Estimate</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>(D_{O_2} )</td>
<td>(1 \times 10^{-10})</td>
<td>(1 \times 10^{-8})</td>
<td>(3 \times 10^{-10})</td>
<td>(5.703 \times 10^{-10})</td>
<td>(\text{m}^2 \text{s}^{-1})</td>
</tr>
<tr>
<td>(A_{O_2} )</td>
<td>0</td>
<td>0.5</td>
<td>0.25</td>
<td>(8.477 \times 10^{-10})</td>
<td>-</td>
</tr>
<tr>
<td>(k_{O_2,R} )</td>
<td>(1 \times 10^7)</td>
<td>(1 \times 10^{10})</td>
<td>(1 \times 10^8)</td>
<td>(4.940 \times 10^7)</td>
<td>(\text{L mol}^{-1} \text{s}^{-1})</td>
</tr>
<tr>
<td>(\frac{k_{p,OO}}{k_{p,R0}} = \frac{k_{b,OO}}{k_{b,R0}} )</td>
<td>0</td>
<td>(1 \times 10^{-2})</td>
<td>(2 \times 10^{-3})</td>
<td>(3.530 \times 10^{-10})</td>
<td>-</td>
</tr>
<tr>
<td>(\frac{k_{l,OOO}}{k_{l,RR0}} )</td>
<td>0</td>
<td>1.5</td>
<td>(2 \times 10^{-3})</td>
<td>1.500</td>
<td>-</td>
</tr>
</tbody>
</table>
B.5.4 Simulation Results and Model Assessment

Figure B.3 shows fits to the experimental data that were used for parameter estimation. As shown in Figure 3a and 3b, the model provides relatively good fits to runs with targeted thickness of 8 µm and light intensity of 1000 W/m², using different BAPO concentrations (i.e., [BAPO] = 1 wt% = 0.0256 mol/L and [BAPO] = 4 wt% = 0.103 mol/L). Note that the target values of light intensity (i.e., 1000 W/m² and 5000 W/m²) were determined with a sensor. A correction factor of 1.18 was then applied to these target values to obtain the actual value of light intensity used in each experiment. The corresponding actual values of film thickness and light intensity are reported in each figure. As shown in Figure 3c, for runs with thicker films (i.e., \( z_f = 11.8 \) µm) and lower light intensities (i.e., \( I_{l,0} = 200 \) W/m² and \( I_{l,0} = 500 \) W/m²), the model overpredicts the overall vinyl group conversion. Figure B.4 compares the differences in vinyl group conversion when polymerization was conducted with and without oxygen. For experiments involving films with thickness of 11.8 µm, [BAPO] = 4 wt% and \( I_{l,0} = 750 \) W/m², the final vinyl conversion when oxygen is not present is 86.6 %, whereas the corresponding value is only 22.7 % when oxygen is present, indicating that the inhibition effects of oxygen (from the air) on photopolymerization of HDDA are significant. Notice that the model provides an excellent match to the data for both types of runs.

Figure B.5 shows the simulations for additional runs that were not used during parameter estimation. These runs were conducted using thick films (i.e., 15.7 µm < \( z_f < 20.2 \) µm). As shown in Figure B.5a, the model correctly predicts qualitative trends for experiments conducted using a low BAPO concentration (i.e., [BAPO] = 1 wt%), fixed light intensity targeting 1000 W/m², and a variety of film thicknesses. However, the model underpredicts the vinyl conversions for these runs. As shown in Figure 5b, the model slightly overpredicts the conversions obtained for three replicate runs conducted with a higher BAPO concentration (i.e., [BAPO] = 4 wt%) and the same targeted light intensity. For very-high-light-intensity runs shown in Figure 5c (i.e., \( I_{l,0} \) targets 5000 W/m²),
the model provides good predictions for vinyl conversions at long times, but the predicted rates at early times are too high.

In summary, the proposed model and parameter estimates provide good predictions for data obtained from experiments involving thin films (up to 12 μm) with a variety of BAPO levels (1 wt % to 4 wt %) and relatively low light intensities (200 to 1000 W/m²). Worse predictions are obtained for experiments with thicker films. This result is not surprising because thicker films are expected to have steeper oxygen concentration gradients than thinner films.13,14,21 These gradients are ignored in the current model (assumption B.1.24). In future, we will develop a more comprehensive PDE model that accounts for gradients in the concentration of oxygen and other species. The parameters that we estimated in the current study will be useful starting values for simulations with this more comprehensive model.
Figure B. 3 Experimental data (△, ○, ×) and corresponding model predictions shown using curves with the same colors for experiments with oxygen at air level with a) [BAPO] = 1 wt%, $z_f$ targets 8 µm and $I_{l,0}$ targets 1000 W/m², b) [BAPO] = 4 wt%, $z_f$ targets 8 µm and $I_{l,0}$ targets 1000 W/m², and c) [BAPO] = 4 wt%, $z_f$ = 11.8 µm, and light intensities $I_{l,0} = 200$ W/m² and $I_{l,0} = 500$ W/m².
Figure B. 4 Model predictions (curved lines) and the corresponding FTIR measured values for experiments in the absence (X) and in the presence of oxygen (Δ) using [BAPO] = 4 wt%, \( z_f \sim 12 \mu m \), and \( I_{l,0} \sim 750 \text{ W/m}^2 \).
Figure B. 5 Experimental data (△, ○, ⋄) and corresponding model predictions shown using curves with the same colors for experiments with oxygen at air level with a) \( [BAPO] = 1 \text{ wt\%} \), \( Z_f \) targets 16 \( \mu \text{m} \), and \( I_{L0} \) targets 1000 W/m\(^2\), b) \( [BAPO] = 4 \text{ wt\%} \), \( Z_f \) targets 16 \( \mu \text{m} \), and \( I_{L0} \) targets 1000 W/m\(^2\), and c) \( [BAPO] = 4 \text{ wt\%} \), \( Z_f \) targets 16 \( \mu \text{m} \), and \( I_{L0} \) targets 5000 W/m\(^2\).
B.6 Conclusions

A fundamental model was developed for photopolymerization of 1,6-hexanediol diacrylate (HDDA) with the bifunctional initiator bis-acylphosphine oxide (BAPO) to account for inhibition effects of oxygen. Real-time FTIR vinyl-group conversion data were used to support parameter estimation and model testing. A sensitivity-based estimability analysis confirmed that all of the oxygen-related parameters were estimable using data obtained from 8 dynamic experiments. The resulting model and parameter estimates provide good predictions for experiments involving thin films up to 12 \( \mu m \), with a variety of BAPO levels (1 wt% to 4 wt%) and relatively low light intensities (200 to 1000 W/m\(^2\)).

Because spatial variances in oxygen concentration and other species were neglected in the current model, it does not do an adequate job of predicting conversion vs. time behaviour in thick films using high light intensities. In the future, a more comprehensive model will be developed (i.e., using partial differential equations that account for oxygen and monomer diffusion) to account for gradients in species concentrations. The parameters estimates obtained in the current study will be used as starting values for this more comprehensive model, which should provide more accurate predictions for thicker films.

B.7 Acknowledgements

Financial support from MITACS is gratefully acknowledged by A-D. Vo, K. Abdi and K.B. McAuley. This project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 876362. The JU receives support from the European Union’s Horizon 2020 research and innovation programme and Austria, Belgium, Czech Republic, Finland, Germany, Italy, Latvia, Netherlands, Poland, and Switzerland.

We thank Hennie Boonen for performing part of the FT-IR measurements.
B.8 References for appendix B


Appendix C

Supplementary Information for Chapter 3

C.1 Extending Hamilton’s Method to Obtain Linearization-based Joint Confidence Region for Reactivity-Ratio Estimates

The linearization method proposed by Hamilton is used in nonlinear least-squares regression (with perfectly known inputs) to obtain the JCR for an important subset of parameters. Hamilton categorized the overall set of parameters \( \theta \) as primary parameters \( \theta_p \) and nuisance parameters \( \theta_n \). He showed that the resulting covariance matrix for the parameter estimates \( \hat{\theta}_p \) in a single-output model:

\[
Y = g(\theta, x) + \varepsilon \quad \text{(C-1)}
\]

is:

\[
\Sigma_{\theta_p} = \frac{\hat{J}_{\theta_p}^T \hat{Q} \hat{J}_{\theta_p}}{\hat{\sigma}_\varepsilon^2} \quad \text{(C-2)}
\]

where \( \hat{J}_{\theta_p} \) is the Jacobian matrix of the model predictions with respect to the primary parameters evaluated using the parameter estimates:

\[
\hat{J}_{\theta_p} = \\
\begin{bmatrix}
\frac{\partial g}{\partial \theta_{p,1}}|_{\hat{\theta}_p, \hat{\theta}_n, x_1} & \frac{\partial g}{\partial \theta_{p,1}}|_{\hat{\theta}_p, \hat{\theta}_n, x_1} & \cdots & \frac{\partial g}{\partial \theta_{p,N}}|_{\hat{\theta}_p, \hat{\theta}_n, x_1} \\
\frac{\partial g}{\partial \theta_{p,1}}|_{\hat{\theta}_p, \hat{\theta}_n, x_2} & \cdots & \cdots & \frac{\partial g}{\partial \theta_{p,N}}|_{\hat{\theta}_p, \hat{\theta}_n, x_2} \\
\vdots & \cdots & \cdots & \vdots \\
\frac{\partial g}{\partial \theta_{p,1}}|_{\hat{\theta}_p, \hat{\theta}_n, x_N} & \cdots & \cdots & \frac{\partial g}{\partial \theta_{p,N}}|_{\hat{\theta}_p, \hat{\theta}_n, x_N}
\end{bmatrix} \quad \text{(C-3)}
\]

In Equation (C-2), \( \hat{Q} \) is a weighting matrix that accounts for the effects of uncertainties in \( \hat{\theta}_n \) on the estimates of the primary parameters:

\[
\hat{Q} = I_N - \hat{J}_{\theta_n}(\hat{J}_{\theta_n}^T \hat{J}_{\theta_n})^{-1} \hat{J}_{\theta_n}^T \quad \text{(C-4)}
\]
where $I_N$ is a $N$-by-$N$ identity matrix and $J_{\theta_n}$ is the Jacobian matrix of the model predictions with respect to nuisance parameters $\theta_n$. 

First, we propose an extension to Hamilton’s method for our EVM reactivity-ratio case study where measurement variances in inputs and outputs are assumed known. To do so, we define an augmented parameter vector $\theta_{aug}$ which contains the reactivity ratios as the primary parameters and the unknown uncertain inputs as the nuisance parameters:

$$\theta_{aug} = [r_1, r_2, \ln f_{1,1}, \ln f_{1,2}, ..., ln f_{1,17}]^T$$ (C-5)

We also define a scaled augmented prediction vector:

$$g_{aug,s} = \begin{bmatrix} \frac{ln F_{1,1}}{\sigma_{ln F_{1,1}}} & \frac{ln F_{1,2}}{\sigma_{ln F_{1,2}}} & ... & \frac{ln F_{1,17}}{\sigma_{ln F_{1,17}}} \\ \frac{\ln f_{1,1}}{\sigma_{ln f_{1,1}}} & \frac{\ln f_{1,2}}{\sigma_{ln f_{1,2}}} & ... & \frac{\ln f_{1,17}}{\sigma_{ln f_{1,17}}} \end{bmatrix}^T$$ (C-6)

where $F_{1,i} (i = 1, ..., 17)$ is the predicted copolymer composition corresponding to the $i^{th}$ experimental condition. In Equation (C-6), the subscript $s$ in $g_{aug,s}$ is used to indicate that the output and uncertain inputs have been scaled. Scaling the model predictions and uncertain inputs in this way transforms the model in Equations (3-8.1) and (3-8.2) into an equivalent model where corresponding scaled measurement errors have unit variances. The extended Hamilton equation for estimating the covariance matrix of the reactivity ratios becomes:

$$\Sigma_{\theta_p} = \hat{J}_{\theta_p}^T \tilde{Q} \hat{J}_{\theta_p}$$ (C-7)

where the weighting matrix $\tilde{Q}$ can be calculated from:

$$\tilde{Q} = I_{34} - \hat{J}_{\theta_n}(\hat{J}_{\theta_n}^T \hat{J}_{\theta_n})^{-1} \hat{J}_{\theta_n}^T$$ (C-8)

Note that we use a $34 \times 34$ identity matrix because the vector $g_{aug,s}$ contains 34 elements. The Jacobian matrix $\hat{J}_{\theta_p}$ in Equation (C-7) contains the derivatives of the model predictions with respect to the reactivity ratios.
\[ \mathbf{J}_{\theta_p} = \begin{bmatrix} \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_1} & \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_2} \\ \sigma_{ln F_1} & \mathbf{f}_1 \end{bmatrix} \begin{bmatrix} \mathbf{f}_2 \mathbf{f}_1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_1} & \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_2} \\ \sigma_{ln F_1} & \mathbf{f}_1 \end{bmatrix} \]  

(C-9)

In \( \mathbf{J}_{\theta_p} \) the derivatives of the uncertain inputs with respect to the reactivity ratios are zero.

As a result, all elements in the 18\(^{th}\) to 34\(^{th}\) rows in \( \mathbf{J}_{\theta_p} \) are zero, so that:

\[ \mathbf{J}_{\theta_p} = \begin{bmatrix} \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_1} & \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_2} \\ \sigma_{ln F_1} & \mathbf{f}_1 \end{bmatrix} \begin{bmatrix} \mathbf{f}_2 \mathbf{f}_1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_1} & \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_2} \\ \sigma_{ln F_1} & \mathbf{f}_1 \end{bmatrix} \]  

(C-10)

To compute \( \hat{\mathbf{Q}} \) in Equation (C-7), we need the elements of \( \mathbf{J}_{\theta_n} \):

\[ \mathbf{J}_{\theta_n} = \begin{bmatrix} \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_1} & \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_2} \\ \sigma_{ln F_1} & \mathbf{f}_1 \end{bmatrix} \begin{bmatrix} \mathbf{f}_2 \mathbf{f}_1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_1} & \frac{1}{\sigma_{ln F_1}} \frac{\partial \ln F_1}{\partial r_2} \\ \sigma_{ln F_1} & \mathbf{f}_1 \end{bmatrix} \]  

(C-11)

Derivatives of the natural logarithms for the output predictions (\( \ln F_{i,1} \), where \( i = 1, \ldots, 17 \)) with respect to the natural logarithms for the uncertain inputs (\( \ln (f_{1,l}) \), where \( l=1,\ldots,17 \)) are non-
zero when \( i = l \) (e.g., \( \frac{\partial \ln f_{1,l}}{\partial n f_{1,l}} = \frac{r_2 - 2r_2 f_{1,l} + r_2 f_{1,l}^2 - 2r_2 r_2 f_{1,l} + 2r_2 r_2 f_{1,l}}{(r_1 f_{1,l} - f_{1,l} + 1)(2f_{1,l} + r_2 - 2r_2 f_{1,l} + r_1 f_{1,l} + r_2 f_{1,l}^2 - 2f_{1,l}^2)} \)). However, the corresponding derivatives when \( i \neq l \) are zero (e.g., \( \frac{\partial \ln f_{1,i}}{\partial n f_{1,i}} = 0 \)). Finally, derivatives of the natural logarithm for the inputs (\( \ln f_{1,i} \), where \( i = 1, \ldots, 17 \)) with respect to the natural logarithm of the uncertain inputs (\( \ln f_{1,i} \), where \( i = 1, \ldots, 17 \)) are one when \( i = l \) (e.g., \( \frac{\partial \ln f_{1,l}}{\partial n f_{1,l}} = 1 \)). However, the corresponding derivatives are zero when \( i \neq l \) (e.g., \( \frac{\partial \ln f_{1,i}}{\partial n f_{1,i}} = 0 \)). As a result, the matrix \( \hat{J}_{\theta_n} \) is simplified as follows:

\[
\hat{J}_{\theta_n} = \begin{bmatrix}
1 & \frac{\partial \ln f_{1,1}}{\partial n f_{1,1}} & 0 & \cdots & 0 \\
\frac{\partial \ln f_{1,1}}{\partial n f_{1,1}} & 0 & \cdots & \cdots & 0 \\
0 & \vdots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\sigma_{\ln f_{1,17}}} & 0 \\
\frac{1}{\sigma_{\ln f_{1,17}}} & 0 & \cdots & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \frac{1}{\sigma_{\ln f_{1,17}}} & \frac{1}{\sigma_{\ln f_{1,17}}} \\
\end{bmatrix}
\]

(C-12)

Substituting Equations (C-10) and (C-12) in Equations (C-7) and (C-8), respectively, results in the following covariance matrix for the reactivity rations:

\[
\Sigma_{(\hat{r}_1, \hat{r}_2)} = \begin{bmatrix}
\sum_{i=1}^{17} \frac{\left( \frac{\partial \ln f_{1,i}}{\partial r_1} \right)^2}{\sigma_{\ln f_{1,i}}^2 + \left( \frac{f_{1,i} \partial F_{1,i}}{F_{1,i} \partial f_{1,i}} \right)^2 \sigma_{f_{1,i}}^2} \\
\sum_{i=1}^{17} \frac{\left( \frac{\partial \ln f_{1,i}}{\partial r_2} \right)^2}{\sigma_{\ln f_{1,i}}^2 + \left( \frac{f_{1,i} \partial F_{1,i}}{F_{1,i} \partial f_{1,i}} \right)^2 \sigma_{f_{1,i}}^2} \\
\sum_{i=1}^{17} \frac{\left( \frac{\partial \ln f_{1,i}}{\partial r_1} \right)^2}{\sigma_{\ln f_{1,i}}^2 + \left( \frac{f_{1,i} \partial F_{1,i}}{F_{1,i} \partial f_{1,i}} \right)^2 \sigma_{f_{1,i}}^2} \\
\sum_{i=1}^{17} \frac{\left( \frac{\partial \ln f_{1,i}}{\partial r_2} \right)^2}{\sigma_{\ln f_{1,i}}^2 + \left( \frac{f_{1,i} \partial F_{1,i}}{F_{1,i} \partial f_{1,i}} \right)^2 \sigma_{f_{1,i}}^2} \\
\end{bmatrix}^{-1}
\]

(C-13)

After the covariance matrix for the reactivity ratios is estimated from Equation (C-13), the linearization-based JCR can be obtained from:

\[
[r_1 - \hat{r}_1 \quad r_2 - \hat{r}_2] \Sigma_{(\hat{r}_1, \hat{r}_2)} \begin{bmatrix} r_1 - \hat{r}_1 \\ r_2 - \hat{r}_2 \end{bmatrix} \leq \rho
\]

(C-14)
where $\rho$ is the value of abscissa for the chi-squared distribution with 2 degrees of freedom and with 5% of the area to the right. Note that, we use the chi-squared distribution to obtain the right-hand side of the Equation (C-14), because the measurement variances for $LnF_1$ and $Lnf_1$ are assumed known. As a result, the degrees of freedom associated with the covariance matrix $\hat{\Sigma}(\hat{r}_1, \hat{r}_2)$ is infinity. Equation (C-14) corresponds to the red ellipse shown in Figure 3-1. Note that this proposed approach for extending Hamilton’s method can readily be used to construct the JCR for model parameters in other EVM parameter estimation problems where measurement variances are assumed known.

In situations where the output measurement variance is estimated from true replicates (i.e., using Equation (3-5)), an F distribution should be used to construct the linearization-based JCR for the model parameters. The reason that an F distribution is required is that the degrees of freedom associated with the estimated output measurement variance are finite. Note that, the degrees of freedom for the numerator in this the F distribution is two when two primary parameters (i.e., $r_1$ and $r_2$) are being estimated. The denominator degrees of freedom for the F distribution is the degrees of freedom associated with the computation of $\Sigma(\hat{r}_1, \hat{r}_2)$. Determining the denominator degrees of freedom is not straightforward.

When we estimate the parameters in the copolymerization case study assuming TR1 replicates, the augmented parameter vector is:

$$\theta_{aug} = [r_1, r_2, \ln f_{1,1}, \ln f_{1,2}, ..., \ln f_{1,11}]^T$$ (C-15)

and augmented scaled prediction vector is:

$$g_{aug,s} = \left[ \frac{\ln F_{1,1}}{\sigma_{lnF_1}}, \frac{\ln F_{1,2}}{\sigma_{lnF_1}}, ..., \frac{\ln F_{1,10}}{\sigma_{lnF_1}}, \frac{\ln F_{1,11}}{\sigma_{lnF_1}}, ..., \frac{\ln f_{1,1}}{\sigma_{lnf_1}}, \frac{\ln f_{1,2}}{\sigma_{lnf_1}}, ..., \frac{\ln f_{1,11}}{\sigma_{lnf_1}} \right]^T$$ (C-16)
In Equation (C-16), there are four repeated values of \( \frac{\ln F_{1,10}}{\sigma_{\ln F_1}} \) and four repeated values of \( \frac{\ln F_{1,11}}{\sigma_{\ln F_1}} \), corresponding to the output model predictions for replicate data at target run conditions 10 and 11, respectively.

The weighting matrix \( \mathbf{Q} \) in this situation is:

\[
\mathbf{Q} = I_{28} - \mathbf{J}_\theta (\mathbf{J}_\theta^T \mathbf{J}_\theta)^{-1} \mathbf{J}_\theta^T \tag{C-17}
\]

A \( 28 \times 28 \) identity matrix appears in Equation (C-17) because there are 28 elements in \( g_{\text{aug}} \) in Equation (C-13). There are 17 scaled output predictions (some are repeated) and there are 11 scaled uncertain inputs (corresponding to the 11 distinct run conditions).

The covariance matrix for the parameters in the TR1 situation is:

\[
\Sigma = \begin{bmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{21} & \Sigma_{22}
\end{bmatrix} \begin{bmatrix}
\frac{\partial \ln F_1}{\partial r_1} & \frac{\partial \ln F_1}{\partial r_2} \\
\frac{\partial \ln F_1}{\partial r_1} & \frac{\partial \ln F_1}{\partial r_2}
\end{bmatrix}^{-1} \begin{bmatrix}
\frac{\partial \ln F_1}{\partial r_1} & \frac{\partial \ln F_1}{\partial r_2} \\
\frac{\partial \ln F_1}{\partial r_1} & \frac{\partial \ln F_1}{\partial r_2}
\end{bmatrix}
\]

(C-18)

The denominator for the four elements of \( \Sigma_{(\hat{r}_1, \hat{r}_2)} \) is different than in Equation (C-13) because we have repeated elements in \( g_{\text{aug}} \) corresponding to the output measurements at each replicate target condition. However, the augmented prediction vector \( g_{\text{aug}} \) only had unique elements, resulted in a different form for the denominator of the elements in \( \Sigma_{(\hat{r}_1, \hat{r}_2)} \).

The appropriate degrees of freedom for \( \Sigma_{(\hat{r}_1, \hat{r}_2)} \) depends on the degrees of freedom for the denominator \( \frac{1}{n_i} \delta_{\ln F_i} + \left( \frac{f_i \theta_{F_i}}{f_i \theta_{F_i}} \right) \beta_{\ln F_i} \). There is not an easy approach to obtain the degrees of freedom for \( \frac{1}{n_i} \delta_{\ln F_i} + \left( \frac{f_i \theta_{F_i}}{f_i \theta_{F_i}} \right) \beta_{\ln F_i} \), which is the weighted sum of two variance estimates. The first term in this sum contains \( \delta_{\ln F_i} \), which has six degrees of freedom because it is
estimated from two sets of replicate data with a total of six degrees of freedom. The second variance term has infinite degrees of freedom because the variance of the random error in $\ln f_1$ is assumed to be perfectly known. As a result, we believe that the degrees of freedom is somewhere between six and infinity, but we have been unable to determine an appropriate value to use (even after considering the Welch–Satterthwaite method for approximating degrees of freedom for sums of chi-squared distributions).\(^2\)

Instead, we resorted to a parametric bootstrapping approach to obtain empirical parameter uncertainty information for situations involving true replicates (See section C.2).

The PR situation is even more complicated than the TR1 situation, with the corresponding sum of squares in the denominators of the elements of $\Sigma_{(\hat{r}_1, \hat{r}_2)}$ becoming $\sigma^2_{\ln f_1} + (\frac{\partial \hat{f}_1}{\partial f_1} \frac{\partial \hat{f}_1}{\partial f_2} \frac{\partial f_1}{\partial f_2} \frac{\partial f_1}{\partial f_1})^2 \sigma^2_{\ln f_1}$. In this situation, we do not know the degrees of freedom for the variance estimate $\sigma^2_{\ln f_1}$ obtained from the pseudo replicates using Equation (3-8). As a result, we use a bootstrapping technique for the PR situation as well. Detailed information about a proposed bootstrapping method to obtain empirical joint confidence information for TR and PR situations is provided in section C.2.

### C.2 Proposed Bootstrapping Approach to Obtain Empirical Joint Confidence Information for Reactivity-Ratio Estimates

Parametric bootstrapping has been widely used to obtain joint confidence information for model parameters in nonlinear regression problems involving perfectly-known inputs.\(^3\) An extension to this method is proposed here for EVM situations. This extended method was used to quantify the uncertainties in reactivity ratio estimates as shown by the green dots in Figures 3-1a to 3-1c. Table C.1 summarizes the Steps used obtain uncertainty information for the reactivity ratios in the situation where the output measurement variance $\sigma^2_{\ln f_1}$ is assumed known (i.e., to generate the green dots in Figure 1a)). In Step one, the reactivity ratios are estimated. In Step 2, synthetic bootstrap data are generated for the true values of the uncertain inputs. In Step 3, synthetic
output measurement data are generated. In Step 4, bootstrap reactivity ratios \( \hat{r}_1^{(b)} \) and \( \hat{r}_2^{(b)} \) are estimated by using the synthetic data from Step 3 and the objective function \( J_{AK} \). A conceptual difference between the bootstrapping technique used here and other related parametric bootstrapping techniques involving perfectly known inputs\(^3\) is that we must generate plausible true values for the uncertain inputs while fixing the uncertain input measurements (in Step 2).

Table C. 1 Proposed bootstrap steps to obtain uncertainties of reactivity-ratio parameters estimated from data in Table 7 using EVM when output measurement variances are assumed known

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Estimate the parameters ( r_1 ) and ( r_2 ) and the true values of the uncertain monomer-composition inputs ( \ln f_{1m,i} ) (where ( i = 1, ..., 17 )) using ( J_{AK} ) objective function in Equation (3-9.1). Assign bootstrap counter ( b = 1 ) and the number of desired bootstrap parameter estimates ( b_{max} ).</td>
</tr>
</tbody>
</table>
| 2.   | Generate synthetic true values for the uncertain inputs by randomly perturbing each of the corresponding input measurements \( \ln f_{1m,i} \) (where \( i = 1, ..., 17 \)):
\[
\ln (f_{i,i}^{(b)}) = \ln f_{1m,i} - \varepsilon_{Ln f_{1,i}}^{(b)} \quad \text{where} \quad \varepsilon_{Ln f_{1,i}}^{(b)} \sim N(0, \sigma_{Ln f_{1,i}}^2) \quad (C.1.1)
\]
| 3.   | Generate synthetic bootstrap data by randomly perturbing the corresponding model predictions that would be obtained using the parameter estimates in Step 1 and true values from Step 2:
\[
\ln F_{1m,i}^{(b)} = \ln \left( \frac{\hat{r}_1(f_{i,i}^{(b)})^2 + f_{i,i}^{(b)}(1 - f_{i,i}^{(b)})}{\hat{r}_1(f_{i,i}^{(b)})^2 + 2f_{i,i}^{(b)}(1 - f_{i,i}^{(b)}) + \hat{r}_2(1 - f_{i,i}^{(b)})} \right) + \varepsilon_{Ln F_{1,i}}^{(b)} \quad (C.1.2)
\]
where \( \varepsilon_{Ln F_{1,i}}^{(b)} \sim N(0, \sigma_{Ln F_{1,i}}^2) \)
| 4.   | Use the data generated in Step 3 in the \( J_{AK} \) objective function:
\[
J_{AK}^{(b)} = \sum_{i=1}^{17} \left[ \frac{(\ln F_{1m,i}^{(b)} - \ln(F_{1m,i}))^2}{\sigma_{Ln F_{1,i}}^2} + \frac{(\ln f_{1m,i}^{(b)} - \ln(f_{1m,i}))^2}{\sigma_{Ln f_{1,i}}^2} \right] \quad (C.1.4)
\]
to estimate plausible reactivity ratio values \( \hat{r}_1^{(b)} \) and \( \hat{r}_2^{(b)} \).
| 5.   | If \( b = b_{max} \), go to step 6. Otherwise, increase \( b \) by one and return to the step two.
| 6.   | Plot \( \hat{r}_2^{(b)} \) vs. \( \hat{r}_1^{(b)} \) (\( b = 1, ..., b_{max} \)) to obtain the cloud of plausible parameter estimates in Figure 1a). |

The required steps for obtaining reactivity-ratio uncertainties in situation TR1 are provided in Table C.2. The main difference compared to the algorithm in Table C.1 is that we must estimate the measurement variance \( \sigma_{Ln F_{1}}^2 \) in Step 1, and we must re-estimate it in Step 5. The green points shown in Figure 3-1b were obtained using the bootstrapping technique in Table C.2.
Table C.2 Bootstrap steps to obtain uncertainties of parameter estimated from EVM in TR1 situation

1- Estimate output measurement variance $\sigma_{LnF_1}^2$ from true replicates using Equation (3-5).
2- Estimate the parameters $r_1$ and $r_2$ and the true values of uncertain inputs $f_{1,i}$ (where $i = 1, ..., 11$) using objective function $J_{TR1}$ (i.e., Equation (3-9.2) in Table 3-9). Assign bootstrap counter $b=1$ and set the number of desired bootstrap parameter estimates $b_{max}$.
3- Generate synthetic true values for the uncertain inputs ($i = 1, ..., 11$) by randomly perturbing the $i^{th}$ measured input:
   \[ \ln (f_{1,i}^{(b)}) = \ln f_{1m,i} - \varepsilon_{Lnf_{1,i}}^{(b)} \] (C.2.1)
   where $\varepsilon_{Lnf_{1,i}}^{(b)} \sim N(0, \sigma_{Lnf_{1,i}}^2)$
4- Generate synthetic bootstrap data by randomly perturbing the corresponding model predictions that would be obtained using the reactivity ratio estimates in Step 2 and the plausible true values from Step 3:
   \[ \ln F_{1m,ij}^{(b)} = \ln \left( \frac{\hat{r}_1 f_{1,i}^{(b)} + f_{1,i}^{(b)} (1-f_{1,i}^{(b)})}{\hat{r}_1 f_{1,i}^{(b)} + 2 f_{1,i}^{(b)} (1-f_{1,i}^{(b)}) + \hat{r}_2 (1-f_{1,i}^{(b)})} \right) + \varepsilon_{LnF_{1,ij}}^{(b)} \] (C.2.2)
   where $\varepsilon_{LnF_{1,ij}}^{(b)} \sim N(0, \sigma_{LnF_{1,i}}^2)$
5- Calculate the bootstrap output measurement variance $\sigma_{LnF_1}^2 (b)$ using Equation (5), based on the synthetic data from Step 4.
6- Estimate the reactivity ratio parameters using the synthetic data in Step 4, the estimated measurement variance $\hat{\sigma}_{LnF_1}^2 (b)$, and the $J_{TR1}$ objective function:
   \[ J_{TR1}^{(b)} = \sum_{i=1}^{11} \sum_{j=1}^{n_i} \left( \frac{\ln F_{1m,ij}^{(b)} - \ln (F_{1,i})}{\hat{\sigma}_{LnF_1}^2 (b)} \right)^2 + \sum_{i=1}^{11} \left( \frac{\ln f_{1m,i} - \ln (f_{1,i})}{\sigma_{Lnf_{1,i}}^2} \right)^2 \]
   to obtain the bootstrap reactivity-ratio estimates $\hat{r}_1^{(b)}$ and $\hat{r}_2^{(b)}$.
7- If $b = b_{max}$, go to Step 8. Otherwise, increase $b$ by one and return to Step 3.
8- Plot $\hat{r}_2^{(b)}$ vs. $\hat{r}_1^{(b)}$ ($b = 1, ..., b_{max}$) to obtain the cloud of plausible parameter estimates in Figure 1b).

The proposed steps for quantifying reactivity-ratio uncertainties in Figure 1c (i.e., the PR situation) using bootstrapping are provided in Table C.3. The steps are similar to the steps in Table C.2. The main difference is that, in the PR situation, at each replicate setting $n_i$ different plausible true values for $\ln f_1$ are synthesized, whereas in the TR1 situation only one plausible true value should be generated using each target input condition.
Table C. 3 Bootstrap steps to obtain uncertainties in parameter estimates from EVM in PR situation

1- Estimate the reactivity ratio values and output measurement variance $\sigma_{\ln F_1}^2$ using the algorithm in Table 3-8. Assign bootstrap counter b=1 and set the number of desired bootstrap parameter estimates $b_{\text{max}}$.

2- Generate synthetic true values for the $j^{th}$ ($j = 1, ..., n_i$) uncertain inputs for the $i^{th}$ ($i = 1, ..., 11$) experimental condition by randomly perturbing the $i^{th}$ measured input:

$$\ln (f_{1,i,j}^{(b)}) = \ln f_{1m,i} - \varepsilon_{\ln f_{1,i,j}}^{(b)} \quad (C.3.1)$$

where $\varepsilon_{\ln f_{1,i,j}}^{(b)} \sim N(0, \sigma_{\ln f_{1,i,j}}^2)$

3- Generate synthetic bootstrap data by randomly perturbing the corresponding model predictions that would be obtained using the plausible reactivity ratio values from Step 1 and the plausible true values of the inputs from Step 2:

$$\ln F_{1m,i,j}^{(b)} = \ln \left( \frac{\hat{r}_1 f_{1,j}^{(b)} + f_{1,i,j}^{(b)} (1 - f_{1,j}^{(b)})}{\hat{r}_1 f_{1,j}^{(b)} + 2 f_{1,i,j}^{(b)} (1 - f_{1,j}^{(b)}) + \varepsilon_{\ln F_{1,i,j}}^{(b)}} \right) + \varepsilon_{\ln F_{1,i,j}}^{(b)} \quad (C.3.2)$$

where $\varepsilon_{\ln F_{1,i,j}}^{(b)} \sim N(0, \tilde{\sigma}_{\ln F_1}^2)$

4- Obtain bootstrap measurement variance estimates $\tilde{\sigma}_{\ln F_1}^2(b)$ and reactivity-ratio estimates $\hat{r}_1^{(b)}$ and $\hat{r}_2^{(b)}$ using the synthetic data from Step 3 and the algorithm in Table 8.

5- If $b = b_{\text{max}}$, go to Step 6. Otherwise, increase b by one and return to the Step 2.

6- Plot $\hat{r}_2^{(b)}$ vs. $\hat{r}_1^{(b)}$ ($b = 1, ..., b_{\text{max}}$) to obtain the cloud of plausible parameter estimates in Figure 1c).
C.3 References for Appendix C


Appendix D

Supplementary Information for Chapter 4

D.1 Derivation of Linearization-based Covariance Matrices in Equations (4-11) and (4-12) and Further Details about Linearization-based Covariance Matrices

Consider a single response model with perfectly known model inputs:

\[ Y_i = g(x_i, \theta) + \varepsilon_{Y_i} \quad (D-1) \]

where \( \varepsilon_{Y_i} \) is normally distributed measurement noise with variance \( \sigma_{\varepsilon_Y}^2 \). Using a maximum-likelihood argument, the ordinary least squared objective function corresponding to equation (D-1) is:

\[ J_{OLS} = \sum_{i=1}^{N} (y_{m,i} - g(x_i, \theta))^2 \quad (D-2) \]

The covariance matrix for the parameter estimates \( \hat{\theta} \) obtained by minimizing equation (D-2) can be quantified using the Hessian of the objective function in equation (D-2):\(^1,^2\)

\[ \Sigma_{\hat{\theta}} = (0.5 \hat{H}_{J_{OLS}})^{-1} \sigma_{Y}^2 \quad (D-3) \]

where \( \hat{H}_{J_{OLS}} \) is the Hessian matrix of the objective function in equation (D-2) with respect to the parameters, evaluated at \( \hat{\theta} \):

\[ \hat{H}_{J_{OLS}} = \begin{bmatrix} \frac{\partial^2 J_{OLS}}{\partial \theta_1^2} & \frac{\partial^2 J_{OLS}}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 J_{OLS}}{\partial \theta_1 \partial \theta_N} \\ \frac{\partial^2 J_{OLS}}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 J_{OLS}}{\partial \theta_2^2} & \cdots & \frac{\partial^2 J_{OLS}}{\partial \theta_2 \partial \theta_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 J_{OLS}}{\partial \theta_N \partial \theta_1} & \frac{\partial^2 J_{OLS}}{\partial \theta_N \partial \theta_2} & \cdots & \frac{\partial^2 J_{OLS}}{\partial \theta_N \partial \theta_N} \end{bmatrix} \quad (D-4) \]

The covariance matrix for the parameter estimates can also be estimated from another linearization-based approach:\(^2,^3\)
\[\Sigma_\theta = (\hat{J}_g^T \hat{J}_g)^{-1} \sigma_y^2 \]  \hspace{1cm} (D-5)

where \(\hat{J}_g\) is the Jacobian matrix of the model predictions \(g(x_i, \theta)\) with respect to the parameters evaluated using parameter estimates:

\[
\hat{J}_g = \begin{bmatrix}
\frac{\partial g(x_1, \theta)}{\partial \theta_1} & \frac{\partial g(x_1, \theta)}{\partial \theta_2} & \cdots & \frac{\partial g(x_1, \theta)}{\partial \theta_N} \\
\frac{\partial g(x_2, \theta)}{\partial \theta_1} & \frac{\partial g(x_2, \theta)}{\partial \theta_2} & \cdots & \frac{\partial g(x_2, \theta)}{\partial \theta_N} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial g(x_N, \theta)}{\partial \theta_1} & \vdots & \ddots & \frac{\partial g(x_N, \theta)}{\partial \theta_N}
\end{bmatrix} \hspace{1cm} (D-6)
\]

In the current study, we extend the covariance matrix expressions in equations (D-3) and (D-5) to obtain parameter covariance matrices for multi-input, multi-output models with input uncertainties. To do so, we define the scaled augmented measurement vector as the vector of random variables:

\[YU_{aug,s} = \begin{bmatrix}
Y_{1,1} & Y_{1,2} & \cdots & Y_{1,N_Y} \\
\sigma_{Y_1} & \sigma_{Y_2} & \cdots & \sigma_{Y_N_Y} \\
Y_{2,1} & Y_{2,2} & \cdots & Y_{N,Y_N_Y} \\
\sigma_{eY_1} & \sigma_{eY_2} & \cdots & \sigma_{eY_{N_Y}} \\
U_{1,1} & U_{1,2} & \cdots & U_{1,N_{U}} \\
\sigma_{eU_1} & \sigma_{eU_2} & \cdots & \sigma_{eU_{N_U}} \\
U_{2,1} & U_{2,2} & \cdots & U_{N_{U},N_{U}} \\
\sigma_{eU_{N_U}} & \sigma_{eU_{N_U}} & \cdots & \sigma_{eU_{N_U}}
\end{bmatrix}^T \hspace{1cm} (D-7)
\]

The elements in \(YU_{aug,s}\) are shown with upper case letters because they are random variables associated with the measurements of uncertain outputs and inputs. The \(q^{th}\) element in \(YU_{aug,s}\) can be expressed using the model:

\[YU_{aug,s,q} = g_{aug,s,q}(x_q, \theta_{aug}) + \varepsilon_{YU_{aug,s,q}} \hspace{1cm} (q = 1, \ldots, N_{N_Y} + N_{N_Y}) \hspace{1cm} (D-8)\]

where \(\varepsilon_{YU_{aug,s,q}}\) are normally distributed random variables with unit variance due to the scaling of elements in equation (D-7). Recall from equation (4-10) that \(\theta_{aug}\) is a vector containing the model parameters and unknown inputs.

The first \(N_{N_Y}\) elements of \(g_{aug,s}\) (i.e., \(g_{aug,s,q}\) when \(1 \leq q \leq N_{N_Y}\)) are scaled output model predictions, and the next \(N_{N_U}\) elements of \(g_{aug,s}\) (i.e., \(g_{aug,s,q}\) when \(N_{N_Y} + 1 \leq q \leq\)

228


\( NN_U + NN_Y \) are the scaled uncertain inputs. Similarly, measured values for the uncertain outputs and inputs can be scaled and collected in an augmented vector:

\[
y_{m,\text{aug},s} = \left[ y_{m1,1} \ y_{m1,2} \ ... \ y_{m1,NN} \ y_{m2,1} \ y_{m2,2} \ ... \ y_{m2,NN} \ y_{mN,1} \ y_{mN,2} \ ... \ y_{mN,NN} \ u_{1,1} \ u_{1,2} \ ... \ u_{1,U1} \ u_{2,1} \ u_{2,2} \ ... \ u_{2,U2} \ ... \ u_{U1,U2} \ ... \ u_{U1,NN} \right]^T
\]  

(D-9)

Using a maximum-likelihood argument, we can obtain an EVM objective function which is structurally similar to the OLS objective function in equation (D-2):

\[
J_{EVM} = \sum_{q=1}^{NN_U + NN_Y} (y_{m,\text{aug},s,q} - g_{\text{aug},s,q}(x_q, \theta_{\text{aug}}))^2
\]  

(D-10)

Similar to equation (D-3), the covariance matrix for estimated values in \( \theta_{\text{aug}} \) can be obtained using a Hessian matrix:

\[
\Sigma_{\theta_{\text{aug}}} = \left( \frac{1}{2} \hat{H}_{EVM} \right)^{-1}
\]  

(4-11)

wherein the Hessian matrix \( \hat{H}_{EVM} \) contains second derivatives of \( J_{EVM} \) with respect to the parameters and uncertain inputs (i.e., the elements of \( \theta_{\text{aug}} \):

\[
\hat{H}_{EVM} =
\]

(D-11)

The Hessian matrix in equation (D-11) can be generated automatically when the optimization function ‘fmincon.m’ is used in MATLAB to minimize \( J_{EVM} \). In situations where the model developer uses a different optimizer for EVM parameter estimation, the second derivatives in \( \hat{H}_{EVM} \) can be obtained numerically based on finite difference approximations. For example, the upper left element can be approximated by: \(^4\)
\[
\frac{\partial^2 J_{EVM}}{\partial \theta_1^2} \bigg|_{\hat{\theta}, \hat{u}_1, \ldots, \hat{u}_N} \cong \frac{J_{EVM}|_{0.05 \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N, \hat{u}_1, \ldots, \hat{u}_N - 2 J_{EVM}|_{0.05 \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N, \hat{u}_1, \ldots, \hat{u}_N}}{(0.05 \hat{\theta}_1)^2} \quad (D-12)
\]

where \( J_{EVM}|_{0.05 \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N, \hat{u}_1, \ldots, \hat{u}_N} \) and \( J_{EVM}|_{0.95 \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N, \hat{u}_1, \ldots, \hat{u}_N} \) are values of the objective function obtained by perturbing \( \hat{\theta}_1 \) by 5\% in positive and negative directions, respectively.

The elements containing cross derivatives such as \( \frac{\partial^2 J_{EVM}}{\partial \theta_1 \partial \theta_2} \bigg|_{\hat{\theta}, \hat{u}_1, \ldots, \hat{u}_N} \) can be obtained as follows:

\[
\frac{\partial^2 J_{EVM}}{\partial \theta_1 \partial \theta_2} \bigg|_{\hat{\theta}, \hat{u}_1, \ldots, \hat{u}_N} \cong \frac{J_{EVM}|_{0.05 \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N, \hat{u}_1, \ldots, \hat{u}_N - J_{EVM}|_{0.95 \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N, \hat{u}_1, \ldots, \hat{u}_N}}{4(0.05 \hat{\theta}_1)(0.05 \hat{\theta}_2)} \quad (D-13)
\]

Obtaining the elements of \( \hat{H}_{EVM} \) using equations (D-12) and (D-13) can be cumbersome if the numbers of unknown parameters and uncertain inputs are large, especially for EVM parameter estimation with large data sets.

Alternatively, the Jacobian-based linearization approach in equation (4-12) can be used to obtain parameter uncertainty information:

\[
\Sigma_{\theta_{aug}} = (\hat{J}_{g_{aug,s}})^{-1} \quad (4-12)
\]

where \( \hat{J}_{g_{aug,s}} \) is the Jacobian matrix of \( g_{aug,s} \) with respect to the parameters and uncertain inputs:
can be generated from the built-in function `lsqnonlin.m` in MATLAB when it is used for minimizing $J_{EVM}$.

The elements in the first $N \times N_y$ rows of $\tilde{J}_{aug}$ correspond to derivatives of output model predictions with respect to the elements of $\theta_{aug}$, evaluated using $\widehat{\theta}$ and $\widehat{u}$. Further, the elements in the last $N \times N_U$ rows contain derivatives of uncertain inputs with respect to the elements of augmented parameter vector. Many of the elements of $\tilde{J}_{aug}$ are zero. For example, \[
\frac{\partial g_4(u_1, x_1, \theta)}{\partial u_{23}} |_{u_1, x_1, \theta} = 0 \]
because uncertain inputs used in the second run do not affect the model prediction from the first run. Also, for example, \[
\frac{\partial u_{11}}{\partial \theta_1} |_{u_1, x_1, \theta} = 0 \]
because model parameters do not influence the uncertain independent variables. Fortunately, the Jacobian matrix in equation (D-14) can be generated from the built-in function `lsqnonlin.m` in MATLAB when it is used for minimizing $J_{EVM}$.
D.2 References for Appendix D

E.1 Derivation of Material Balance Equations in Table 5-5

Assuming steady-state operation, mass balance equations for the species are of the form:

\[ F_q = F_q^0 + R_q W \quad (q = 2\text{MP}, 3\text{MP}, \text{C3}) \quad (E-1) \]

where the rate of generation \( R_q \) for species \( q \) is:

\[ R_{2\text{MP}} = \frac{k_1 \frac{P_{\text{nC6}}}{P_{\text{H2}}} - k_3 \frac{P_{2\text{MP}}}{P_{\text{H2}}}}{1 + K_{\text{phys}}(P_{\text{nC6}} + P_{2\text{MP}} + P_{3\text{MP}})} \quad (E-2) \]

\[ R_{3\text{MP}} = \frac{2k_2 \frac{P_{\text{nC6}}}{P_{\text{H2}}}}{1 + K_{\text{phys}}(P_{\text{nC6}} + P_{2\text{MP}} + P_{3\text{MP}})} \quad (E-3) \]

\[ R_{\text{C3}} = \frac{k_3 \frac{P_{2\text{MP}}}{P_{\text{H2}}}}{1 + K_{\text{phys}}(P_{\text{nC6}} + P_{2\text{MP}} + P_{3\text{MP}})} \quad (E-4) \]

where \( P_q \) (q=2MP, 3MP, C3, nC6) is the partial pressure for species \( q \). These partial pressures can be obtained from:

\[ P_{\text{H2}} = \frac{F_{\text{H2}}}{F_{\text{H2}}^0 + F_{\text{nC6}}^0} P \quad (E-5) \]

\[ P_{\text{nC6}} = \frac{F_{\text{nC6}}}{F_{\text{H2}}^0 + F_{\text{nC6}}^0} P \quad (E-6) \]

\[ P_{2\text{MP}} = \frac{F_{2\text{MP}}}{F_{\text{H2}}^0 + F_{\text{nC6}}^0} P \quad (E-7) \]

\[ P_{3\text{MP}} = \frac{F_{3\text{MP}}}{F_{\text{H2}}^0 + F_{\text{nC6}}^0} P \quad (E-8) \]

\[ P_{\text{C3}} = \frac{F_{\text{C3}}}{F_{\text{H2}}^0 + F_{\text{nC6}}^0} P \quad (E-9) \]
Substitution of equations (E-5) to (E-9) into equations (E-2) to (E-4) gives:

\[ R_{2MP} = \frac{k_1 F_{nC_6} - k_3 F_{2MP}}{1 + k_{phys} P \left( \frac{F_{nC_6} + F_{2MP} + F_{3MP}}{F_{H_2} + F_{C_6}} \right)} \]  
(E-10)

\[ R_{3MP} = \frac{2k_2 F_{nC_6}}{1 + k_{phys} P \left( \frac{F_{nC_6} + F_{2MP} + F_{3MP}}{F_{H_2} + F_{C_6}} \right)} \]  
(E-11)

\[ R_{C_3} = \frac{k_3 F_{2MP}}{1 + k_{phys} P \left( \frac{F_{nC_6} + F_{2MP} + F_{3MP}}{F_{H_2} + F_{C_6}} \right)} \]  
(E-12)

Substitution of equations (E-10) to (E-12) into equation (S1) for \( q = 2MP, 3MP \) and \( C_3 \), respectively, gives:

\[ F_{2MP} - \left( k_1 \left( F_{nC_6}^0 - F_{2MP} - F_{3MP} - \frac{F_{C_3}}{2} \right) - k_3 F_{2MP} \right) \frac{F_{H_2}}{1 + k_{phys} P \left( \frac{F_{nC_6} + F_{2MP} + F_{3MP}}{F_{H_2} + F_{C_6}} \right)} = 0 \]  
(E-13)

\[ F_{3MP} - \left( k_2 \left( F_{nC_6}^0 - F_{2MP} - F_{3MP} - \frac{F_{C_3}}{2} \right) \right) \frac{F_{H_2}}{1 + k_{phys} P \left( \frac{F_{nC_6} + F_{2MP} + F_{3MP}}{F_{H_2} + F_{C_6}} \right)} = 0 \]  
(E-14)

\[ F_{C_3} - \left( 2k_3 F_{2MP} \right) \frac{F_{C_3}}{1 + k_{phys} P \left( \frac{F_{nC_6} + F_{2MP} + F_{3MP}}{F_{H_2} + F_{C_6}} \right)} = 0 \]  
(E-15)

Equations (E-13) to (E-15) contain five unknowns (i.e., \( F_{2MP}, F_{3MP}, F_{C_3}, F_{H_2} \) and \( F_{nC_6} \)). The number of unknowns in equations (E-13) to (E-15) can be reduced to three using the following balance equations for carbon and hydrogen atoms:

\[ F_{H_2} = F_{H_2}^0 - \frac{F_{C_3}}{2} \]  
(E-16)

\[ F_{nC_6} = F_{nC_6}^0 - F_{2MP} - F_{3MP} - \frac{F_{C_3}}{2} \]  
(E-17)
Substitution of equations (E-16) and (E-17) into equations (E-13) to (E-15) results in balance equations (5-5.1) to (5-5.3) in Table 5-5.

### E.2 Experimental Data
The experimental data that are used in this study were obtained by Toch et al. and are shown in Table E-1.

**Table E. 1** Experimental data from Toch et al. for hydroisomerization of n-hexane in a gas-phase, well-mixed reactor

<table>
<thead>
<tr>
<th>Run #</th>
<th>T (K)</th>
<th>P (MPa)</th>
<th>$F_{nC_6}^0$ (10^{-5} \text{ mol s}^{-1})</th>
<th>$F_{H_2}^0$ (10^{-3} \text{ mol s}^{-1})</th>
<th>$F_{2MP}$ (10^{-6} \text{ mol s}^{-1})</th>
<th>$F_{3MP}$ (10^{-6} \text{ mol s}^{-1})</th>
<th>$F_{C_3}$ (10^{-6} \text{ mol s}^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>493</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>3.15</td>
<td>1.66</td>
<td>0.19</td>
</tr>
<tr>
<td>2</td>
<td>493</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>2.86</td>
<td>1.53</td>
<td>0.08</td>
</tr>
<tr>
<td>3</td>
<td>493</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>2.96</td>
<td>1.55</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>493</td>
<td>1.0</td>
<td>2.55</td>
<td>1.92</td>
<td>2.65</td>
<td>1.47</td>
<td>0.03</td>
</tr>
<tr>
<td>5</td>
<td>493</td>
<td>1.0</td>
<td>2.55</td>
<td>2.55</td>
<td>2.38</td>
<td>1.25</td>
<td>0.01</td>
</tr>
<tr>
<td>6</td>
<td>493</td>
<td>2.0</td>
<td>2.55</td>
<td>1.28</td>
<td>2.58</td>
<td>1.34</td>
<td>0.08</td>
</tr>
<tr>
<td>7</td>
<td>493</td>
<td>2.0</td>
<td>2.55</td>
<td>1.28</td>
<td>2.25</td>
<td>1.16</td>
<td>0.05</td>
</tr>
<tr>
<td>8</td>
<td>493</td>
<td>2.0</td>
<td>2.55</td>
<td>1.28</td>
<td>2.37</td>
<td>1.23</td>
<td>0.04</td>
</tr>
<tr>
<td>9</td>
<td>493</td>
<td>2.0</td>
<td>2.55</td>
<td>1.92</td>
<td>2.12</td>
<td>1.11</td>
<td>0.09</td>
</tr>
<tr>
<td>10</td>
<td>513</td>
<td>2.0</td>
<td>2.55</td>
<td>1.92</td>
<td>4.68</td>
<td>2.52</td>
<td>0.05</td>
</tr>
<tr>
<td>11</td>
<td>513</td>
<td>2.0</td>
<td>2.55</td>
<td>1.28</td>
<td>5.00</td>
<td>2.70</td>
<td>0.12</td>
</tr>
<tr>
<td>12</td>
<td>513</td>
<td>1.0</td>
<td>2.55</td>
<td>2.55</td>
<td>4.64</td>
<td>2.47</td>
<td>0.22</td>
</tr>
<tr>
<td>13</td>
<td>513</td>
<td>1.0</td>
<td>2.55</td>
<td>1.92</td>
<td>5.00</td>
<td>2.69</td>
<td>0.31</td>
</tr>
<tr>
<td>14</td>
<td>513</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>5.36</td>
<td>2.91</td>
<td>0.16</td>
</tr>
<tr>
<td>15</td>
<td>513</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>5.53</td>
<td>3.00</td>
<td>0.03</td>
</tr>
<tr>
<td>16</td>
<td>533</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>7.71</td>
<td>4.18</td>
<td>0.96</td>
</tr>
<tr>
<td>17</td>
<td>533</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>7.91</td>
<td>4.51</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>533</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>7.11</td>
<td>3.80</td>
<td>0.79</td>
</tr>
<tr>
<td>19</td>
<td>533</td>
<td>1.0</td>
<td>2.55</td>
<td>1.92</td>
<td>7.06</td>
<td>3.78</td>
<td>0.39</td>
</tr>
<tr>
<td>20</td>
<td>533</td>
<td>1.0</td>
<td>2.55</td>
<td>2.55</td>
<td>6.27</td>
<td>3.33</td>
<td>0.07</td>
</tr>
<tr>
<td>21</td>
<td>533</td>
<td>2.0</td>
<td>2.55</td>
<td>1.28</td>
<td>7.42</td>
<td>4.11</td>
<td>0.56</td>
</tr>
<tr>
<td>22</td>
<td>533</td>
<td>2.0</td>
<td>2.55</td>
<td>1.92</td>
<td>7.10</td>
<td>3.90</td>
<td>0.39</td>
</tr>
<tr>
<td>23</td>
<td>553</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>9.02</td>
<td>5.98</td>
<td>1.85</td>
</tr>
<tr>
<td>24</td>
<td>553</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>8.99</td>
<td>6.86</td>
<td>1.36</td>
</tr>
<tr>
<td>25</td>
<td>553</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>8.92</td>
<td>6.71</td>
<td>1.66</td>
</tr>
<tr>
<td>26</td>
<td>553</td>
<td>1.0</td>
<td>2.55</td>
<td>1.92</td>
<td>8.29</td>
<td>5.87</td>
<td>1.17</td>
</tr>
<tr>
<td>27</td>
<td>553</td>
<td>1.0</td>
<td>2.55</td>
<td>2.55</td>
<td>6.70</td>
<td>5.42</td>
<td>0.43</td>
</tr>
<tr>
<td>28</td>
<td>553</td>
<td>2.0</td>
<td>2.55</td>
<td>1.28</td>
<td>8.95</td>
<td>7.19</td>
<td>1.97</td>
</tr>
<tr>
<td>29</td>
<td>553</td>
<td>2.0</td>
<td>2.55</td>
<td>1.92</td>
<td>8.43</td>
<td>5.10</td>
<td>0.85</td>
</tr>
<tr>
<td>30</td>
<td>573</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>10.78</td>
<td>7.31</td>
<td>3.29</td>
</tr>
<tr>
<td>31</td>
<td>573</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>9.50</td>
<td>7.09</td>
<td>3.20</td>
</tr>
<tr>
<td>32</td>
<td>573</td>
<td>1.0</td>
<td>2.55</td>
<td>1.28</td>
<td>10.81</td>
<td>7.26</td>
<td>3.51</td>
</tr>
<tr>
<td>33</td>
<td>573</td>
<td>1.0</td>
<td>2.55</td>
<td>1.92</td>
<td>8.97</td>
<td>7.50</td>
<td>1.99</td>
</tr>
<tr>
<td>34</td>
<td>573</td>
<td>1.0</td>
<td>2.55</td>
<td>2.55</td>
<td>8.27</td>
<td>6.35</td>
<td>1.73</td>
</tr>
<tr>
<td>35</td>
<td>573</td>
<td>2.0</td>
<td>2.55</td>
<td>1.28</td>
<td>9.34</td>
<td>6.89</td>
<td>3.04</td>
</tr>
<tr>
<td>36</td>
<td>573</td>
<td>2.0</td>
<td>2.55</td>
<td>1.92</td>
<td>9.59</td>
<td>6.10</td>
<td>1.69</td>
</tr>
</tbody>
</table>
E.3 Additional Uncertainty Quantification Results for Situation where $F_{nC6,m}^0$ and $F_{H2,m}^0$ are Uncertain

Figure E-1 shows the results of uncertainty quantification for six of the parameters (the results for the other two parameters are shown in Figures 5-2 and 5-3 in the main text) based on 200 bootstrap parameter estimation runs. The corresponding model prediction uncertainties for the mean prediction and a single-future prediction of $F_{3MP}$ are shown in Figures E-2 and E-3 for two different experimental conditions. Figure E-4 shows the uncertainty boxplots for yield predictions corresponding to a single-future measurement that could be obtained using two different target experimental conditions. Figures E-5 and E-6 show the uncertainty boxplots for predictions of 2MP and 3MP selectivity associated with the two different experimental conditions.
Figure E. 1 Boxplots for uncertainties of the parameters based on 200 bootstrap parameter estimates when both uncertain inputs $F_H^0$ and $F_C^0$ are uncertain: a) kinetic rate of the second reaction at the reference temperature, b) kinetic rate of the third reaction at the reference temperature, c) equilibrium constant at the reference temperature, d) activation energy for the first reaction, e) activation energy for the second reaction and f) activation energy for the third reaction. End of the whiskers correspond to an approximate 95% confidence intervals.
Figure E. 2 Prediction uncertainty boxplots for a) mean prediction of $F_{3MP}$ and b) a single-future measurement of $F_{3MP}$ constructed based on 200 MC simulations. Predictions correspond to the experimental condition at the 15th run in Table E.1 ($T = 513\, \text{K}, P = 1\, \text{MPa}, F_{H_2 m}^0 = 1.3 \times 10^{-3}\, \text{mol s}^{-1}$, and $F_{nC_6 m}^0 = 2.56 \times 10^{-5}\, \text{mol s}^{-1}$). The whisker ends correspond to approximate 95% confidence intervals.

Figure E. 3 Prediction uncertainty boxplots for a) mean prediction of $F_{3MP}$ and b) a single-future measurement of $F_{3MP}$ constructed based on 200 MC simulations corresponding to a new experimental condition ($T = 580\, \text{K}, P = 1\, \text{MPa}, F_{H_2 m}^0 = 1.3 \times 10^{-3}\, \text{mol s}^{-1}$, and $F_{nC_6 m}^0 = 2.56 \times 10^{-5}\, \text{mol s}^{-1}$). The whisker ends correspond to approximate 95% confidence intervals.
**Figure E. 4** Boxplots for yield predictions constructed based on 200 MC simulations: a) for a single-future yield predictions at the target experimental condition with $T = 513 \text{ K}, P = 1 \text{ MPa}, F_{H_2}^{0} m = 1.3 \times 10^{-3} \frac{\text{mol}}{s}$ and $F_{\text{H}_6}^{0} \text{m} = 2.56 \times 10^{-5} \frac{\text{mol}}{s}$; b) for single-future yield predictions at a new experimental condition with higher temperature where $T = 580 \text{ K}, P = 1.0 \text{ MPa}, F_{H_2}^{0} m = 1.3 \times 10^{-3} \frac{\text{mol}}{s}$ and $F_{\text{H}_6}^{0} \text{m} = 2.56 \times 10^{-5} \frac{\text{mol}}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).
Figure E. 5 Boxplots for $S_{2MP}$ predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with $T = 513 K, P = 1 MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$; c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where $T = 580 K, P = 1.0 MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).
Figure E. 6 Boxplots for $S_{3MP}$ predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with $T = 513 K, P = 1 MPa, F_{H_2} m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6} m = 2.56 \times 10^{-5} \frac{mol}{s}$; c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where $T = 580 K, P = 1.0 MPa, F_{H_2} m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6} m = 2.56 \times 10^{-5} \frac{mol}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).
E.4 Uncertainty Quantification Results for Situation where $F_{nC_6,m}^0$ is the only Uncertain Input Measurement

The results of parameter uncertainty quantification using bootstrap are shown in Figures E-7 and E-8 for the situation where there is only one uncertain input (i.e., $F_{nC_6,m}^0$). The corresponding prediction uncertainty quantification results are shown in Figures E-9 to E-13. These figures contain two kinds of prediction uncertainty results (i.e., uncertainties in the mean prediction and uncertainties in a single future prediction). Figures E-9 and E-10 show the prediction uncertainty boxplots for outflow rate predictions corresponding to two experimental conditions. Figure E-11 shows the uncertainties in predictions of the yield for two experimental conditions. Figures E-12 and E-13 show prediction uncertainty boxplots for 2MP and 3MP selectivity, respectively, corresponding to two different target experimental conditions.
Figure E. 7 Boxplots for uncertainties of the parameters based on 200 bootstrap parameter estimates when independent variable $F_{\text{nc}}^{0}$ is uncertain: a) kinetic rate of the first reaction at the reference temperature, b) kinetic rate of the second reaction at the reference temperature, c) kinetic rate of the third reaction at the reference temperature and d) equilibrium constant at the reference temperature. End of the whiskers correspond to an approximate 95% confidence intervals.
Figure E. 8 Boxplots for uncertainties of the parameters based on 200 bootstrap parameter estimates when independent variable $F_{C_6}^0$ is uncertain: a) activation energy of the first reaction, b) activation energy of the second reaction, c) activation energy of the third reaction and d) enthalpy of physisorption. End of the whiskers correspond to an approximate 95% confidence intervals.
Figure E. 9 Prediction uncertainty boxplots for a) mean prediction of $F_{2MP}$, b) a single-future measurement of $F_{2MP}$, c) mean prediction of $F_{3MP}$, d) a single-future measurement of $F_{3MP}$, e) mean prediction of $F_{C3}$ and f) a single-future measurement of $F_{C3}$ constructed based on 200 MC simulations. Predictions correspond to the experimental condition at the 15th run in Table E.1 ($T = 513 \, K, P = 1 \, MPa$, and $F_{H2 \, m}^0 = 1.3 \times 10^{-3} \, \text{mol s}^{-1}$ and $F_{nC6 \, m}^0 = 2.56 \times 10^{-5} \, \text{mol s}^{-1}$) when independent variable $F_{C6}$ is uncertain. The whisker ends correspond to approximate 95% confidence intervals.
Figure E. 10 Prediction uncertainty boxplots for a) mean prediction of $F_{2MP}$, b) a single-future measurement of $F_{2MP}$, c) mean prediction of $F_{3MP}$, d) a single-future measurement of $F_{3MP}$, e) mean prediction of $F_{C3}$ and f) a single-future measurement of $F_{C3}$ constructed based on 200 MC simulations. Predictions correspond to the experimental at a new condition ($T = 580 K, P = 1 MPa$, and $F_{H2, m}^0 = 1.3 \times 10^{-3} \frac{mol}{s}$, and $F_{nC6, m}^0 = 2.56 \times 10^{-5} \frac{mol}{s}$) where independent variable $F_{C6}^0$ is uncertain. The whisker ends correspond to approximate 95% confidence intervals.
Figure E. 11 Boxplots for yield predictions constructed based on 200 MC simulations: a) for mean yield predictions and b) for a single-new prediction at the target experimental condition with $T = 513 \, K, P = 1 \, MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$; c) for mean yield predictions and d) a single-new prediction at a new experimental condition with higher temperature where $T = 580 \, K, P = 1.0 \, MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).
Figure E. 12 Boxplots for $S_{2MP}$ predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with $T = 513 K, P = 1 MPa, F_{H_2}^{0,m} = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$; c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where $T = 580 K, P = 1.0 MPa, F_{H_2}^{0,m} = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).
Figure E. 13 Boxplots for $S_{3MP}$ predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with $T = 513 \, K$, $P = 1 \, MPa$, $F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$; c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where $T = 580 \, K$, $P = 1.0 \, MPa$, $F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).
E.5 Uncertainty Quantification Results for the Situation where $F_{H_2 m}^0$ is the only Uncertain Input

Figures E-14 to E-20 show uncertainty boxplots for parameter estimates and model predictions when $F_{H_2 m}^0$ is the only uncertain input. Figures E-14 and E-15 show the uncertainty boxplots for estimated parameters using 200 bootstrap runs. Figures E-16 and E-17 show the uncertainties in predictions of outflow rates (i.e., $F_{2MP}, F_{3MP}$ and $F_{C_3}$) for two different experimental conditions. Figure E-18 shows the uncertainties in yield predictions for two experimental conditions. Figures E-19 and E-20 show the prediction uncertainty quantification for $S_{2MP}$ and $S_{3MP}$, respectively. Note that prediction uncertainties were obtained for both the mean prediction and an anticipated single future experiment.
Figure E. 14 Boxplots for uncertainties of the parameters based on 200 bootstrap parameter estimates when independent variable $F_{H_2,m}^0$ is uncertain: a) kinetic rate of the first reaction at the reference temperature, b) kinetic rate of the second reaction at the reference temperature, c) kinetic rate of the third reaction at the reference temperature and d) equilibrium constant at the reference temperature. End of the whiskers correspond to an approximate 95% confidence intervals.
Figure E. 15 Boxplots for uncertainties of the parameters based on 200 bootstrap parameter estimates when independent variable $F_{H_2 m}^0$ is uncertain: a) activation energy of the first reaction, b) activation energy of the second reaction, c) activation energy of the third reaction and d) enthalpy of physisorption. End of the whiskers correspond to an approximate 95% confidence intervals.
Prediction uncertainty boxplots for a) mean prediction of $F_{2MP}$, b) a single-future measurement of $F_{2MP}$, c) mean prediction of $F_{3MP}$, d) a single-future measurement of $F_{3MP}$, e) mean prediction of $F_{C3}$ and f) a single-future measurement of $F_{C3}$ constructed based on 200 MC simulations. Predictions correspond to the experimental condition at the 15th run in Table E.1 ($T = 513 \, K$, $P = 1 \, MPa$, and $I_{H_2} = 1.3 \times 10^{-3} \, mol \, s^{-1}$, and $I_{N_2} = 2.56 \times 10^{-5} \, mol \, s^{-1}$) when independent variable $F_{H_2}^0$ is uncertain. The whisker ends correspond to approximate 95% confidence intervals.
Figure E. 17 Prediction uncertainty boxplots for a) mean prediction of $F_{2MP}$, b) a single-future measurement of $F_{2MP}$, c) mean prediction of $F_{3MP}$, d) a single-future measurement of $F_{3MP}$, e) mean prediction of $F_{C3}$ and f) a single-future measurement of $F_{C3}$ constructed based on 200 MC simulations. Predictions correspond to the experimental at a new condition ($T = 580 \, K$, $P = 1 \, MPa$, and $F_{H2_m} = 1.3 \times 10^{-3} \frac{mol}{s}$, and $F_{nCO_m} = 2.56 \times 10^{-5} \frac{mol}{s}$) where independent variable $F_{H2}^0$ is uncertain. The whisker ends correspond to approximate 95% confidence intervals.
Figure E. 18 Boxplots for yield predictions constructed based on 200 MC simulations: a) for mean yield predictions and b) for a single-new prediction at the target experimental condition with \( T = 513 \, K, P = 1 \, MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s} \) and \( F_{nC_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s} \); c) for mean yield predictions and d) a single-new prediction at a new experimental condition with higher temperature where \( T = 580 \, K, P = 1.0 \, MPa, F_{H_2}^0 m = 1.3 \times 10^{-3} \frac{mol}{s} \) and \( F_{nC_6}^0 m = 2.56 \times 10^{-5} \frac{mol}{s} \). The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).
Figure E. 19 Boxplots for $S_{2MP}$ predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with $T = 513 \, K$, $P = 1 \, MPa$, $F_{H_2 m}^0 = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6 m}^0 = 2.56 \times 10^{-5} \frac{mol}{s}$; c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where $T = 580 \, K$, $P = 1.0 \, MPa$, $F_{H_2 m}^0 = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6 m}^0 = 2.56 \times 10^{-5} \frac{mol}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).
Figure E. 20 Boxplots for $S_{3MP}$ predictions constructed based on 200 MC simulations: a) for mean predictions and b) a single-new prediction at the target experimental condition with $T = 513 K, P = 1 MPa, F_{H_2,m}^0 = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6}^0 = 2.56 \times 10^{-5} \frac{mol}{s}$; c) for mean predictions and d) for a single-new prediction at a new experimental condition with higher temperature where $T = 580 K, P = 1.0 MPa, F_{H_2,m}^0 = 1.3 \times 10^{-3} \frac{mol}{s}$ and $F_{nC_6}^0 = 2.56 \times 10^{-5} \frac{mol}{s}$. The ends of the whiskers correspond to the 2.5 and 97.5 percentiles (i.e., approximate boundaries of 95% confidence intervals).
E.6 Reference for Appendix E