UNIVERSIDADE D
COIMBRA

Hugo Ferreira Gonçalves

Global Sensitivity Analysis of the Kinetic Model of a Polymerization Reaction According to the SARA-ATRP Mechanism

Master´s dissertation in the field of Chemical Engineering in the area of Biosystems, supervised by Professor Lino Santos, PhD. and Professor Jorge Coelho, PhD. and submitted to the Department of Chemical Engineering, Faculty of Science and Technology, University of Coimbra

October of 2021

Hugo Ferreira Gonçalves

Global Sensitivity Analysis of the Kinetic Model of a Polymerization Reaction According to the SARA-ATRP Mechanism

Master´s dissertation in the field of Chemical Engineering in the area of Biosystems, supervised by Professor Lino Santos, PhD. and Professor Jorge Coelho, PhD. and submitted to the Department of Chemical Engineering, Faculty of Science and Technology, University of Coimbra

Supervisors:
Professor Lino Santos, PhD.
and
Professor Jorge Coelho, PhD.

Institution:
Department of Chemical Engineering
Faculty of Sciences and Technology of the University of Coimbra

1 2 9 0

UNIVERSIDADE D COIMBRA

October, 2021

"Sensitivity analysis for modellers?
Would you go to an orthopaedist who didn’t use X-ray?"
Fürbinger (1996)

In *memoriam* of my grandmother Beatriz Duarte.
Thank you for everything.

Acknowledgments

The acknowledgments will be written in my language, Portuguese.

Em primeiro lugar, agradeço aos meus orientadores: ao professor Lino Santos, por toda a paciência, interesse e disponibilidade demostrada ao longo do desenvolvimento deste trabalho, sem esquecer as longas discussões que certamente fizeram enriquecer o meu conhecimento na área Engenharia Química; ao Professor Jorge Coelho, pela constante disponibilidade em responder a muitas dúvidas que apareceram ao longo deste trabalho e pelo interesse demonstrado nas aplicações de *Process System Engineering* em ciências dos polímeros.

Ao Doutorando Joel Sansana, pelos conselhos, pela ajuda e pela motivação que contribuíram muito positivamente para o desenvolvimento desta dissertação. Agradeço-lhe também por ter fornecido o seu template de LaTeX.

Ao Engenheiro João Coutinho, pelas constantes trocas ideias e discussões, que foram fundamentais na minha motivação de fazer mais e melhor.

Ao Doutor Carlos Abreu, pela disponibilidade que sempre mostrou para me ajudar.

Ao Doutorando Eugeniu Strelet, por me ter apresentado a Parallel Computing Toolbox, que diminui bastante o tempo computacional dos métodos utilizados neste trabalho.

Ao Professor Marco Reis, pelos conselhos dado para aplicação do método *Monte Carlo Filtering*.

Ao pessoal da C16: João; Joel; Eugeniu; Rúben; Rodrigo; Paulo; Carolina. Certamente ficarão muitas saudades da vossa companhia e principalmente das "mesas redondas" que se tinha diariamente nesta sala.

As pessoas que mais marcaram a minha passagem por esta cidade: André Simões, Beatriz Marona, Claudia Gaspar, Daniela Pereira, Daniel Rodrigues, Diogo Salgueiro, Francisco Brandão, Francisco Cardoso, Gonçalo Araújo, Francisco Cardoso, João Lincho, João Lopes, Mariana Rodrigues, Sofia Leão, Paula Oliveira, Rita Cardoso, Rodrigo Paredes, Rúben Gariso, Vasco Araújo, Vítor Sousa e a todos aqueles que tive o privilégio de conhecer nesta cidade. Obrigado por tudo!

Aos meus familiares: ao meu avô António Ferreira, ao meu avô Luciano Gonçalves, à minha avó Manuela Monteiro, à "Tia" Carmo, à minha afilhada Beatriz Moreira, à "prima" Ana, ao Nuno Gomes, ao meu irmão Diogo Gonçalves, à Inês Guerreiro e à minha madrinha Marta Diogo.

Em último, queria agradecer aos meus pais, Jorge e Elsa Gonçalves, que até nos momentos mais difíceis da vida, se esforçaram para dar-me a melhor educação. Certamente, sem vocês, nada disto seria possível.

Resumo

A polimerização via radical por transferência de átomo com agentes suplementares de ativação e redução, usualmente referida como SARA-ATRP, tem sido alvo de interesse na comunidade científica. A partir desta reação é possível produzir-se polímeros com características específicas, sendo que esta se realiza com uma baixa concentração de catalisador e à temperatura ambiente. Um modelo que descreva bem este sistema reacional pode ser uma ferramenta importante para o *scale-up*, optimização e controlo deste processo. Tendo o modelo desenvolvido, um aspeto importante é saber quais são os parâmetros mais importantes do modelo cinético. Estes resultados vão permitir obter um conhecimento extra sobre o modelo cinético, onde os parâmetros mais relevantes deverão ser o foco de novas experiências laboratoriais para melhorar a capacidade preditiva do modelo. Adicionalmente, isto vai ajudar a decidir o planeamento de experiências laboratoriais e, eventualmente, reduzir o custo e tempo das experiências. O principal objetivo deste trabalho é obter a informação de quais os parâmetros mais importantes do modelo cinético SARA-ATRP, tendo sido para isso utilizados métodos globais de análise de sensibilidade.

Neste trabalho foram aplicados dois métodos globais de análise de sensibilidade. A filtragem de Monte Carlo e um baseado na variância, o método de Sobol. A aplicação destes métodos, requer a geração de vários valores dos parâmetros, dentro do intervalo dos seus possíveis valores. De facto, diferentes intervalos considerados, podem mudar o resultado da análise de sensibilidade. Assim, foi decidido estudar 3 intervalos diferentes em termos da gama de valores que se considera os parâmetros poderem assumir.

A formulação do modelo cinético publicado na literatura foi revisitada neste trabalho. As simulações preliminares, mostraram que o modelo tem uma fraca capacidade em previsão da conversão do monómero e da massa molecular numérica média . Tal motivou a resolução de um problema de estimação de parâmetros, mas apenas estimando os mais importantes, identificados a partir dos resultados de análise de sensibilidade.

Os resultados da análise de sensibilidade obtidos para a conversão do monómero pelos dois métodos são semelhantes. Esta semelhança verifica-se não só no número de parâmetros que são classificados como críticos/importantes como também na classifiação da sua importância. Assim, as constantes cinéticas de propagação, de terminação e a constante aparente de ativação de cadeias dormentes são considerados críticos/importantes pelos dois métodos, nos 3 intervalos

de valores. O parâmetro mais importante é a constante de propagação, seguidamente vem a constante de terminação e a constante aparente de ativação de cadeias dormentes. Já nos resultados da análise de sensibilidade para a polidispersividade, há diferenças significativas que podem ser encontradas tanto nos resultados obtidos pelos dois métodos, como nos resultados obtidos consoante o intervalo de variação dos parâmetros considerado. As constantes cinéticas de adição da primeira unidade de monómero e a constante aparente de redução do desativador são classificadas como importantes/críticos pelos dois métodos nos 3 intervalos, sendo que o *ranking* de importância muda consoante o intervalo de variação considerado. Para a massa molecular média podem ser detetadas algumas diferenças entre o resultados obtidos pelos dois métodos e para os diferentes intervalos de variação dos parâmetros. Os parâmetros que são considerados críticos/importantes são os mesmos que os obtidos para a conversão do monómero com a adicão da constante aparente de redução do desativador.

A partir dos resultados da análise de sensibilidade, decidiu-se estimar novos valores para as constantes de propagação e terminação. Com os valores obtidos pela optimização, o modelo foi novamente simulado e desta vez, mostrou uma melhor capacidade para prever a conversão do monómero e a massa molecular média.

Nesta dissertação foram obtidos resultados importantes, tais como: a determinação das constantes cinéticas mais importantes para cada um dos *outputs* do modelo estudados; e a melhoria da capacidade de previsão do modelo, após um problema de estimação de parâmetros. Alguns dos trabalhos futuros a realizar, podem passar em utilizar os resultados da análise de sensibilidade para tentar melhorar a capacidade de previsão do modelo. Contudo, é necessário recolher mais dados experimentais com intuito de diminuir tanto o erro experimental, como erros do modelo.

Palavras-chave: *Análise de sensibilidade global; Método de filtragem de Monte Carlo; Método de Sobol; Modelação de polimerizações; SARA-ATRP.*

Abstract

Supplementary activators and reducing agents atom transfer radical polymerization (SARA-ATRP) is a ATRP variation that has gained attention in the scientific community. Through this reaction it is possible to produce tailor-made polymers using low amounts of catalyst at room temperature. A model which can accurately describe these reactions behaviour it is a powerful tool for scale up, optimization and control of such process. With the kinetic model developed, an important step is to obtain information on which are the most important model parameters.The results will allow to obtain extra knowledge about the kinetic model regarding which are the most relevant parameters that should be the focus of further laboratory experiments to improve the model predictive ability. Moreover, this will help to decide the laboratory experiments planning and eventually to reduce the cost and time required for the experiments. The main goal of this thesis is to obtain information on which are the most important parameters of a SARA-ATRP kinetic model. In order to achieve this goal, global sensitivity analysis techniques are used.

In this work, two different global methods were applied: a regionalized based method the Monte Carlo filtering; and a variance based the Sobol method. To perform these methods, it is necessary to sample various inputs values inside their possible range of variation (i.e., input space). In fact, different input spaces can affect the outcome of the sensitivity analysis. In this thesis 3 different inputs spaces were studied.

The kinetic model of this polymerization system published in the literature was reviewed. The preliminary numerical simulations showed a poor ability of the model at predicting the monomer conversion and the number average molecular weight . To improve the model, it was decided to solve a parameter estimation problem, but only estimating the important parameters. In this way, the results of the sensitivity analysis could be used.

The results obtained by the two methods for the monomer conversion are very similar, in both the parameters that are considered critical/important as well in their importance ranking. Nevertheless, small differences can be detected. The kinetic rates constants of propagation, of termination and the apparent rate constant of dormant chains activation are all considered critical/important, by the two methods, in all the three input spaces studied. The ranking of importance is the rate constant of propagation followed by rate constant of termination and the

apparent rate constant of dormant chains activation. In the polydispersity index output, significant differences can be identified, both in the result obtained by the two methods as well in the result obtained between the input spaces intervals studied. The kinetic rate constants of the first monomer addition and the apparent rate of deactivator reduction are classified as critical/important by the two methods in the three input space intervals, being that the importance ranking changes between them. Finally, for the number average molecular weight some differences can be found between the methods and the intervals studied. For this output, the parameters that are consider critical/important, by the two methods in the three intervals studied, are the same as the ones obtained for the monomer conversion with the addition of the apparent rate constant of deactivator reduction.

Trough the results of the sensitivity analysis it was decided to only estimate the parameters the kinetic rate of propagation and termination. The model was simulated with the new parameter values obtained in the optimization. This time, the model showed a better capability to predict the monomer conversion and the number average molecular weight.

In this thesis, important results were obtained that can be used in future works, such as: the determination of the important kinetic rate constants for the 3 model outputs; and the improvement of the model predictive ability after a parameter estimation problem. In future work the results of the sensitivity analysis can be use to improve the predictive ability of the model. Nevertheless, is necessary to collect more data in order to reduce possible experimental and model errors.

Keywords: *Global sensitivity analysis; Monte Carlo filtering method; Sobol method; Polymerization modeling; SARA-ATRP*

Abbreviations

***e*ATRP** electrochemically meadiated ATRP.

***photo*ATRP** photochemically meadiated ATRP.

ARGET-ATRP activators regenerated by electron transfer ATRP.

ATRP atom transfer radical polymerization.

CDFs cumulative distribution functions.

CRP controlled radical polymerization.

EE elementary effects.

EnvSci environmental science.

FAST Fourier amplitude sensitivity test.

FRP free radical polymerization.

GSA global sensitivity analysis.

ICAR-ATRP initiators for continuous activator regeneration ATRP.

K-S Kolmogorov-Smirnov.

LRP living radical polymerization.

MCF Monte Carlo filtering.

NMP nitroxide mediated polymerization.

NRMSE normalized root mean square error.

OAT one-at-time.

ODEs ordinary differential equations.

OSET outer-sphere electron transfer.

PDI polydispersity index.

PLA poly lactid acid.

PRE persistent radical effect.

PSO particle swarm optimization.

QbD quality by design paradigm.

RAFT reversible addition-fragmentation chain transfer.

RDRP reversible deactivation radical polymerisation.

RMSE root mean square error.

RSA regionalized sensitivity analysis.

SA sensitivity analysis.

SARA-ATRP supplementary activators and reducing agents ATRP.

Symbols

B behavioural output.

DP_n degree of polymerization.

$E(Y)$ output expected value.

$F_n(X_i|B)$ cumulative distribution function of the ith input values that obtained a behavioural output.

$F_n(X_i|\overline{B})$ cumulative distribution function of the ith input values that obtained a non-behavioural output.

J objective function.

$J_{i,\,max}$ maximum value of the ith objective function.

K number of inputs under studied.

K_{ATRP} ATRP equilibrium constant.

K_{dis} dissociation of inorganic salt equilibrium constant.

K_{s} dissolution of inorganic salt equilibrium constant.

L lower bound of the parameter search interval for the optimization problem.

M_M monomer molar mass [g/mol].

M_n number average molecular weight [g/mol].

M_n^{th} theoretical number average molecular weight [g/mol].

M_w weight average molecular weight [g/mol].

N_{exp} number of experimental points.

R_{p} rate of propagation [M s^{-1}].

S_{T_i} total sensitivity index for the ith input.

S_i first order sensitivity index for the ith input.

S_i^L local sensitivity index for the ith input.

U upper bound of the parameter search interval for the optimization problem.

$V(Y)$ output unconditional variance.

X model input.

Y model output.

Δ increment used in the elementary effects method.

δ_k kth order moment for dead chains.

λ_k kth order moment for propagating chains.

μ_i mean for the ith input.

μ_k kth order moment for dormant chains.

$\overline{B}$ non-behavioural output.

$\overline{y_{exp}}$ experimental values average.

σ_i standard deviation for the ith input.

$d_{n,\overline{n}}$ two sample Kolmogorov-Smirnov test statistics.

$f_n(X_i|B)$ probability density function of the ith input values that obtained a behavioural output.

$f_n(X_i|\overline{B})$ probability density function of the ith input values that obtained a non-behavioural output.

k optimization decision variables.

$k_{\mathrm{a0p}}^{\mathrm{app}}$ apparent rate of dormant chains activation by $\mathrm{SO_2}^{\bullet}$[s^{-1}].

$k_{\mathrm{a0}}^{\mathrm{app}}$ apparent rate of initiator activation by $\mathrm{SO_2}^{\bullet}$[s^{-1}].

k_{a1p} rate constant of dormant chains activation [$\mathrm{M}^{-1}\mathrm{s}^{-1}$].

k_{a1} rate constant of initiator activation [$\mathrm{M}^{-1}\mathrm{s}^{-1}$].

k_{add} rate constant of the first monomer addition [$\mathrm{M}^{-1}\mathrm{s}^{-1}$].

k_{ass1} rate constant of association of the $\mathrm{Cu^{I}/L}$ halidophilicity equilibrium [$\mathrm{M}^{-1}\mathrm{s}^{-1}$].

k_{ass2} rate constant of association of the $\mathrm{Cu^{II}X/L}$ halidophilicity equilibrium [$\mathrm{M}^{-1}\mathrm{s}^{-1}$].

k_{d1p} rate constant of propagating chains deactivation [$\mathrm{M}^{-1}\mathrm{s}^{-1}$].

k_{d1} rate constant of initiator deactivation [$\mathrm{M^{-1}s^{-1}}$].

k_{diss1} rate constant of dissociation of the $\mathrm{Cu^{I}/L}$ halidophilicity equilibrium [$\mathrm{s^{-1}}$].

k_{diss2} rate constant of dissociation of the $\mathrm{Cu^{II}X/L}$ halidophilicity equilibrium [$\mathrm{s^{-1}}$].

k_{d} rate constant of initiator dissociation [$\mathrm{M^{-1}s^{-1}}$].

k_{p} rate constant of propagation [$\mathrm{M^{-1}s^{-1}}$].

k_{t0} rate constant of termination between two radicals [$\mathrm{M^{-1}s^{-1}}$].

k_{tR} rate constant of termination between a radical and a propagating chain [$\mathrm{M^{-1}s^{-1}}$].

k_{tc} rate constant of termination by combination [$\mathrm{M^{-1}s^{-1}}$].

k_{td} rate constant of termination by disproportionation [$\mathrm{M^{-1}s^{-1}}$].

k_{t} rate constant of termination [$\mathrm{M^{-1}s^{-1}}$].

$p - \mathrm{value}$ proof value of the statistical test.

r number of elementary effects.

w_i weight for the ith objective function.

x monomer conversion.

$y_{exp,\,i}$ ith experimental value.

$y_{prev,\,i}$ ith model predicted value.

Contents

Acknowledgments **v**

Resumo **vii**

Abstract **ix**

Abbreviations **xi**

Symbols **xiii**

List of Tables **xxi**

List of Figures **xxiii**

1 Objectives and dissertation outline **1**

1.1 Thesis goal 1

1.2 Outline 1

2 Polymerization reaction mechanism and modeling **3**

2.1 Free radical polymerization 3

2.2 Reversible deactivation radical polymerization 5

2.3 Atom transfer radical polymerization 6

2.4 SARA-ATRP 8

2.5 Mathematical modeling 9

2.5.1 Method of moments 10

3 Sensitivity analysis **13**

3.1 Definition . . . 13

3.2 Local sensitivity analysis methods . . . 14

3.3 Global sensitivity analysis methods . . . 15

3.3.1 Elementary effects method . . . 15

3.3.2 Variance based methods . . . 16

3.3.3 Regionalized sensitivity analysis . . . 17

3.4 Methods comparison . . . 18

3.5 Applications of sensitivity analysis . . . 18

3.6 Detailed methods . . . 20

3.6.1 Monte Carlo filtering . . . 21

3.6.2 Sobol method . . . 23

3.7 Settings for GSA . . . 28

3.7.1 Model output . . . 28

3.7.2 Input space . . . 29

3.7.3 Sampling approach . . . 29

4 SARA-ATRP kinetic model formulation and numerical simulation **31**

4.1 Reaction scheme . . . 31

4.2 Model assumptions . . . 33

4.3 Dynamic model formulation . . . 34

4.3.1 The moments differential equations . . . 34

4.3.2 Non-polymeric species molar balances . . . 34

4.4 Numerical simulation . . . 36

5 Global sensitivity analysis results **41**

5.1 Simplifications . . . 41

5.1.1 Parameters studied . . . 41

5.1.2 Notation 41
5.2 Monte Carlo filtering 42
5.2.1 Choosing the number of samples (N) 42
5.2.2 Monomer conversion output results 45
5.2.3 PDI output results 49
5.2.4 M_n output results 56
5.3 Sobol method 61
5.3.1 Methods comparison 61
5.3.2 Choosing the number of samples (N) 63
5.3.3 Monomer conversion output results 65
5.3.4 PDI output results 67
5.3.5 M_n output results 70
5.4 Methods comparison 72

6 Parameter estimation problem **73**
6.1 Problem formulation 73
6.2 Optimization results 74
6.3 Numerical simulation 75

7 Conclusions and Future work **83**
7.1 Conclusions 83
7.2 Future work 85

8 Bibliography **87**

Appendices **95**
A Sobol method demonstrations 97
A.1 Sobol demonstrations 97
A.2 Equation 3.14 deduction 98

B Moments differential equations development 98

C Experimental points 102

D Approximated graphical representation of the empirical correlation matrix . . . 103

E Implementation test 108

F Algorithm for the weights determination 109

List of Tables

2.1 Meaning of the rate constants in (2.1). 4

2.2 Individual momentum of each species [Adapted from Lyra et al. (2019)]. 11

3.1 Comparison between the methods used for SA (Saltelli et al., 2008; Qian and Mahdi, 2020; Yao, 2015). 18

3.2 Parameter classification in function of the p value. 22

3.3 Statistical estimators. 27

3.4 Parameters input spaces for the sensitivity analysis. 29

4.1 Meaning of the rate constants in 4.1. 32

4.2 Kinetic parameters reported in Krys et al. (2017). 36

4.3 Initial conditions. 37

4.4 Error between model prediction and experimental data. 38

5.1 Number of outputs off-target and in-target with 1000 model evaluations. 43

5.2 1st interval monomer conversion MCF result. 45

5.3 2nd interval monomer conversion MCF result. 45

5.4 3rd interval monomer conversion MCF result. 45

5.5 1st interval PDI MCF result. 50

5.6 2nd interval PDI MCF result. 50

5.7 3rd interval PDI MCF result. 50

5.8 1st interval M_n MCF result. 57

5.9 2nd interval M_n MCF result. 57

5.10 3rd interval M_n MCF result. 57

5.11 First order sensitivity indices true values of the parameters in Equation 5.1. . . 62

5.12 CPU time to calculate the first and total indices in function of N. 65

6.1 $J_{i,max}$ with 10000 objective function evaluations. 75

6.2 w_i values. 76

6.3 Optimal parameter values. 76

6.4 NRMSE for the new simulations. 77

6.5 Initial conditions for the 3 experiments. 78

6.6 NRMSE obtain for the 3 experiments. 79

B.1 Simplification results of Equation 15 for the 0th, 1st and 2nd moments. 100

B.2 Simplification of (20) for the 0th, 1st and 2nd moments. 102

C.3 Experimental points for the model validation in Chapter 4. 102

C.4 Experimental points of the 2nd experience. 103

C.5 Experimental points of the 3rd experience. 103

E.6 Global sensitivity analysis results of Equation 22. 109

List of Figures

2.1 Monomer addition to the growing polymer in a radical polymerization [Adapted from Young and Lovell (2011)] 3

3.1 Graphical representation of a 3 dimensional input space [Retrieve from Saltelli et al. (2019)]. 15

3.2 Example of a CDFs plot. 23

3.3 Input space generated by two different number generators for the same number of points. 30

4.1 Profiles obtained along time. 38

4.2 Profiles obtained along the monomer conversion. 39

5.1 Two sample Kolmogorov-Smirnov test statistics in function of $log_{10}(N)$ for the monomer conversion output. 43

5.2 Two sample Kolmogorov-Smirnov test statistics in function of $log_{10}(N)$ for the PDI output. 44

5.3 Two sample Kolmogorov-Smirnov test statistics in function of $log_{10}(N)$ for the M_n output. 44

5.4 CDFs of k_{p} for the monomer conversion. 47

5.5 CDFs of k_{t} for the monomer conversion. 47

5.6 CDFs of $k_{\mathrm{a0p}}^{\mathrm{app}}$ for the monomer conversion. 47

5.7 CDFs of $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ for the monomer conversion. 47

5.8 CDFs of k_{d1p} for the monomer conversion. 48

5.9 CDFs of k_{add} for the monomer conversion. 48

5.10 Graphical representation of the empirical correlation matrix for the monomer conversion. 49

5.11 CDFs of $k_{\mathrm{a0}}^{\mathrm{app}}$ for the PDI. 52

5.12 CDFs of $k_{\mathrm{a0p}}^{\mathrm{app}}$ for the PDI. 53

5.13 CDFs of $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ for the PDI. 53

5.14 CDFs of k_{p} for the PDI. 53

5.15 CDFs of k_{t} for the PDI. 53

5.16 CDFs of k_{a1p} for the PDI. 54

5.17 CDFs of k_{d1p} for the PDI. 54

5.18 CDFs of k_{add} for the PDI. 54

5.19 CDFs of k_{t0} for the PDI. 54

5.20 Graphical representation of the empirical correlation matrix for the PDI. 56

5.21 CDFs of $k_{\mathrm{a0}}^{\mathrm{app}}$ for the M_n. 58

5.22 CDFs of $k_{\mathrm{a0p}}^{\mathrm{app}}$ for the M_n. 58

5.23 CDFs of $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ for the M_n. 59

5.24 CDF of k_{p} for the M_n. 59

5.25 CDFs of k_{t} for the M_n. 59

5.26 CDFs of k_{add} for the M_n. 59

5.27 Graphical representation of the empirical correlation matrix for the M_n. 61

5.28 Results for the parameter X_1. 62

5.29 Results for the parameter X_2. 62

5.30 Results for the parameter X_3. 63

5.31 First and total order indices in function of $log_{10}(N)$ for the monomer conversion. 64

5.32 First and total order indices in function of $log_{10}(N)$ for the PDI. 64

5.33 First and total order indices in function of $log_{10}(N)$ for the M_n. 64

5.34 Sensitivity indices for the monomer conversion. 66

5.35 Sensitivity indices for the PDI output. 67

5.36 Sensitivity indices for the PDI output without the first experimental point. . . . 69

5.37 Sensitivity indices for the M_n output. 70

6.1 Time profiles with the new parameter values. 76

6.2 Monomer conversion profiles with the new parameter values. 77

6.3 Monomer conversion vs time profile. 78

6.4 PDI vs time profile. 78

6.5 M_n vs time profile. 79

6.6 PDI vs monomer conversion profile. 80

6.7 M_n vs monomer conversion profile. 80

D.1 1st interval for the monomer conversion. 104

D.2 2nd interval for the monomer conversion. 104

D.3 3nd interval for the monomer conversion. 105

D.4 1st interval for the PDI. 105

D.5 2nd interval for the PDI. 106

D.6 3rd interval for the PDI. 106

D.7 1st interval for the M_n. 107

D.8 2nd interval for the M_n. 107

D.9 3rd interval for the M_n. 108

F.10 Steps to calculte w_i. 109

Chapter 1

Objectives and dissertation outline

1.1 Thesis goal

The main goal of this thesis is to investigate on which are the most important parameters of a supplementary activators and reducing agents atom transfer radical polymerization (SARA-ATRP) kinetic model reported in Krys et al. (2017). This reaction has been gaining attention in the scientific community, due to the ability of producing tailor-made polymers with a low polydispersity index (PDI) using low amounts of catalyst at room temperature. In order to achieve the goal of this thesis, global sensitivity analysis (GSA) methods are used.

This GSA will allow the answer of some important questions, such as (Saltelli et al., 2004):"which of the uncertain input factors is more important in determining the uncertainty in the output of interest?"; or "if we could eliminate the uncertainty in one of the input factors, which factor should we choose to reduce the most the variance of the output?". These answers will not only be important to understand the reaction mechanism but also to give information to the experimentalists about which parameters should be the focus of further estimates, in order to improve the model predictive ability. This idea, if correctly applied, can reduce the cost and time of experimental work and promote an efficient experimental planning.

1.2 Dissertation outline

This thesis comprises 6 chapters. In Chapter 2, it is presented an introduction to the reaction system under study, with a brief review radical polymerization including the development of the SARA-ATRP system. This chapter, also provides a brief introduction to the polymerization modeling methods, focusing on the one used in this thesis, the method of moments.

In Chapter 3 the topic of sensitivity analysis (SA) is addressed. First, the meaning of SA is defined and some methods commonly used for performing it are presented. Then, a review of a few interesting applications of SA in different scientific areas is presented to emphasize that

this is an important tool, that can be used in a wide range of applications. The methods used in this thesis are explained in more detail and some important settings for the SA are defined, such as: the model output; the input space; and the sampling approach.

In Chapter 4 firstly, the kinetic model used in this thesis, developed in Krys et al. (2017), is described. Then, the modeling assumptions are defined and the model mathematical expressions are developed. Using the kinetic parameters reported in Krys et al. (2017) the first model simulations are performed. The kinetic model simulations results are compared with the experimental data. Following this simulation, it was concluded that it is necessary to estimate new kinetic parameter values, in order to improve the model predictive ability.

In Chapter 5, the results of the two global SA methods performed in this thesis are presented and interpreted, the Monte Carlo filtering (MCF) and the Sobol Method. Finally, a brief comparison between the two methods is discussed.

In Chapter 6 the kinetic estimation problem is formulated where the decision variables of the optimization problem are specified following the results of the global sensitivity analysis. The kinetic model simulations results with the new kinetic parameter values are compared with other experimental data. This is done in order to prevent the model overfitting to the data used in the parameter estimation problem. Finally, the conclusions and possible future works are presented in Chapter 7.

Chapter 2

Polymerization reaction mechanism and modeling

2.1 Free radical polymerization

Free radical polymerization (FRP) is a chain growth polymerization mechanism used in the production of around 100 million tons of polymers per year. It is the most widely used polymerization technique for the production of polymers for various applications, such as: fibers, plastics and rubbers (Matyjaszewski, 2012; Coessens et al., 2001).

This polymerization mechanism is characterized by a growing polymer chains with an unpaired electron in one of the carbon atoms (radical species) known as active center. This mechanism is illustrated in Figure 2.1, where one of the electrons in the π-bond joins the unpaired one, while the other moves to one of the carbons in the C=C bond, forming a new active center. This process will occur successively, by the constant addition of monomer units, allowing the polymer to growth until the termination of the active site (Young and Lovell, 2011; Odian, 2004).

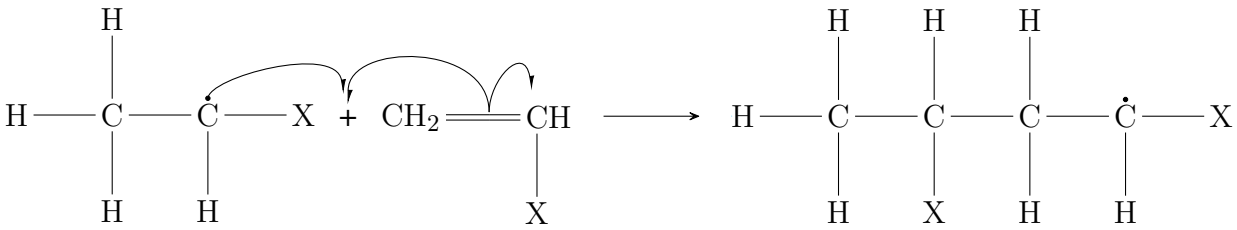


Figure 2.1: Monomer addition to the growing polymer in a radical polymerization [Adapted from Young and Lovell (2011)] .

The FRP reaction can be divided in 3 main stages: initiation; propagation; and termination. Initiation is composed by two steps. The first one is the formation of radicals (R) from the initiator (I) (Equation 2.1a) that can occur in two main ways: by homolytic scission or homoloysis.

In the second step, the first monomer (M) unit is added to the free radical, producing the first propagating chain ($\mathrm{P_1}$) (Equation 2.1b). The next stage is the propagation (Equation 2.1c), which is the stage where the polymer grows trough the successive addition of monomer units, by the mechanism described, forming propagating chains with n monomer units added (P_n). The final stage, termination, can occur in two ways: by combination, when two chains combine and form a single polymer (D_{n+m}) (Equation 2.1d); or by disproportionation, when the product of termination is two individual chains (D_n or D_m) (Equation 2.1e) (Young and Lovell, 2011; Odian, 2004). The meaning of the rate constants in (2.1) are presented in Table 2.1.

$$\mathrm{I} \xrightarrow{\mathrm{k_d}} 2\,\mathrm{R}^{\bullet} \tag{2.1a}$$

$$\mathrm{R}^{\bullet} + \mathrm{M} \xrightarrow{\mathrm{k_{add}}} \mathrm{P_1}^{\bullet} \tag{2.1b}$$

$$\mathrm{P}_n{}^{\bullet} + \mathrm{M} \xrightarrow{\mathrm{k_p}} \mathrm{P}_{n+1}{}^{\bullet} \tag{2.1c}$$

$$\mathrm{P}_n{}^{\bullet} + \mathrm{P}_m{}^{\bullet} \xrightarrow{\mathrm{k_{tc}}} \mathrm{D}_{n+m} \tag{2.1d}$$

$$\mathrm{P}_n{}^{\bullet} + \mathrm{P}_m{}^{\bullet} \xrightarrow{\mathrm{k_{td}}} \mathrm{D}_n + \mathrm{D}_m \tag{2.1e}$$

Table 2.1: Meaning of the rate constants in (2.1).

Nomenclature	description
k_{d}	rate constant of initiator dissociation
k_{add}	rate constant of the first monomer addition
k_{p}	rate constant of propagation
k_{tc}	rate constant of termination by combination
k_{td}	rate constant of termination by disproportionation

The kinetics of the FRP is characterized by a very slow initiation and a very fast propagation and termination of propagating chains, being almost impossible to control the polymer structure. Typically, the final product has high values of polydispersity index (PDI), close to two, due to the very broad molecular weight distribution. Also, using the FRP mechanism is not suitable to synthesise block co-polymers, due to the short lifetime of the propagating chains, of about 1 second for every 1000 monomer additions (Abreu, 2018; Matyjaszewski and Davis, 2003). These limitations, will affect the characteristics of the polymers produced (Coessens et al., 2001). In order to address these limitations, reversible deactivation radical polymerisation (RDRP) (formerly known as living radical polymerization (LRP) and/or controlled radical polymerization (CRP)) have been developed. These mechanisms allow the production of polymers with well defined structures, opening a new door to the world of tailor made polymers (Matyjaszewski, 2012).

2.2 Reversible deactivation radical polymerization

Reversible deactivation radical polymerization (RDRP) techniques allow: the synthesis of well-defined polymers with a broad molecular weight distribution (Abreu, 2018; Matyjaszewski and Davis, 2003; Matyjaszewski, 2012); production of polymer chains with different typologies (like stars, brushes, etc) (Matyjaszewski, 2012) ; and the possibility of making block or other types of co-polymers (Matyjaszewski, 1998, 2012). The success of this type of polymerization is manly due to a dynamic equilibrium between the propagating chains and a dormant specie, which is a funcionalized chain that can not propagate or terminate (Abreu, 2018). This equilibrium allows the life span of a propagating chain to increase from 1 second, in FRP, to more than 1 hour (Mishra and Kumar, 2012).

The methods used in RDRP can be classified according to two main characteristics: dormant and persistent species structures; and chemical mechanism of persistent species exchange. The three most efficient and successful methods, are the nitroxide mediated polymerization (NMP), the reversible addition-fragmentation chain transfer (RAFT) and the metal catalysed atom transfer radical polymerization (ATRP) methods (Matyjaszewski, 2000; Mishra and Kumar, 2012). The most commonly used polymerization method is the ATRP.

In the ATRP the dynamic equilibrium relies on the persistent radical effect (PRE) which was described by Hanns Fischer in his three articles in Fischer (1997, 1999, 2001). In order to explain this equilibrium effect, the reactions (2.2a) and (2.2b) are considered, where, following the classification made in Fischer (1997), X is the persistent species, $\mathrm{P}_n\mathrm{X}$ denotes the dormant chain and k_t is the rate constant of termination (termination by combination + termination by disproportionation).

$$\mathrm{P}_n\mathrm{X} \underset{\mathrm{k_{deact}}}{\overset{\mathrm{k_{act}}}{\rightleftharpoons}} \mathrm{P}_n{}^{\bullet} + \mathrm{X} \tag{2.2a}$$

$$\mathrm{P}_n{}^{\bullet} + \mathrm{P}_n{}^{\bullet} \xrightarrow{\mathrm{k_t}} \mathrm{D}_{n+m} \tag{2.2b}$$

At the beginning of the reaction, if there is zero concentration of radicals (i.e, P_n) , the concentration of radicals and of persistent species increase at the same rate in equal amounts (Equation 2.2a). The radical can irreversibly self-terminate (Equation2.2b), or alternatively react with the persistent species back to the dormant state. However, the transient species can only react by cross reaction with the radicals. As time proceeds, due to the self-termination of radicals, there will be a continuous accumulation of the persistent species, and the self-termination will be less likely to occur. Consequently, this makes the cross reaction the dominant product formation path-way, reaching a state of quasi-equilibrium (Fischer, 1997, 1999, 2001).

When the radical is formed, it can add monomer units to its chain, similarly to FRP (Equation 2.1c). However, in this reaction the rate of polymerization and other characteristics will be strongly dependent on the equilibrium dynamics. In general, the polymer chain stays dor-

mant for about 1 minute after being transient radical for only 1 millisecond (Abreu, 2018; Matyjaszewski and Davis, 2003). This will reduce the chain termination, other side reactions, uncontrolled propagation of monomer units, and, it is expected, a linear evolution of the theoretical number average molecular weight (M_n^{th}) along the monomer conversion, as represented in Equation 2.3. Here, $[\mathrm{M}]_0$ and $[\mathrm{I}]_0$ represent the initial concentration of monomer and initiator, and x and M_M the monomer conversion and molar mass, respectively (Matyjaszewski and Davis, 2003). In addition, to guarantee a controlled polymerization, in terms of polymer properties (number average molecular weight (M_n) and PDI), it is necessary to ensure that all the initiator is consumed at the early stage of the reaction (Matyjaszewski and Davis, 2003).

$$Mn^{th} = \frac{[\mathrm{M}]_0}{[\mathrm{I}]_0} M_M\, x \tag{2.3}$$

The following section will be devoted to the ATRP method, because it is the one under study in this work.

2.3 Atom transfer radical polymerization

ATRP was first introduced by two different investigation groups, whose works are reported in Kato et al. (1995) and Wang and Matyjaszewski (1995), respectively. Since this discovery, this method has become one of the most used techniques for the productions of polymers with narrow PDI, pre-determined M_n and with chain-end functionality (Guliashvili et al., 2012; Santos, 2020). The success of this polymerization can be attributed to the great versatility, the compatibility with a variety of monomers, and the ability to be carried out under mild reaction conditions (Santos, 2020).

ATRP is based on the transition-metal complex mediated equilibrium between a dormant species (that can be an alkyl halide initiator (RX) or an macromolecular specie ($\mathrm{P}_n\mathrm{X}$)) and a propagating radical (R) or chain (P_n) (Guliashvili et al., 2012; Matyjaszewski, 2012). This equilibrium is represented in (2.4).

$$\mathrm{P}_n\mathrm{X} + \mathrm{Mt}^m/\mathrm{L} \underset{\mathrm{k_{d1p}}}{\overset{\mathrm{k_{a1p}}}{\rightleftharpoons}} \mathrm{P}_n{}^{\bullet} + \mathrm{X{-}Mt}^{m+1}/\mathrm{L} \tag{2.4}$$

In Equation 2.4 the dormant species is activated by a metal transition complex in a lower oxidation state (Mt^m/L, where Mt is the transition metal specie, m is the oxidation state and L is the ligand), also called activator. The metal transition complex is oxidized, losing one electron, and binds to the halogen atom from the dormant chain, forming the propagating chain and $\mathrm{X{-}Mt}^{m+1}/\mathrm{L}$ (deactivator), which is in a higher oxidation state. This reaction occurs with a rate k_{a1p}. The growing chains quickly react in a reversible reaction with $\mathrm{X{-}Mt}^{m+1}/\mathrm{L}$, with a rate of k_{d1p}. This path is favoured over activation due to the persistent radical effect explained earlier, where the halogen atom functions as the persistent species (Matyjaszewski, 2012). The

addition of monomer and the small, but existing, termination of polymeric chains occurs when the polymer is active in the same way as in FRP (see (2.1)) (Matyjaszewski and Davis, 2003).

The catalyst used in ATRP, Mt^m in Equation 2.4, is one of the most important components in this technique. It also contributes the most for the value of K_{ATRP}, which is the equilibrium constant ($K_{\mathrm{ATRP}} = \frac{k_{\mathrm{a1p}}}{k_{\mathrm{d1p}}}$). The catalyst will also determine the exchange dynamics between the dormant and propagating chains. The most used metal is by far copper (Cu), due to is versatility and low cost. Nevertheless, other metals such as manganese, ruthenium, iron, cobalt, etc., can be utilized as well. The persistent species, which is a halide group ($-\mathrm{X}$), should migrate rapidly and selectively between the growing polymer chain and the metal transition complex to ensure a controlled polymerization. Usually Bromide (Br) or chlorine (Cl) are used (Matyjaszewski and Davis, 2003).

The rate of polymerization (R_{p}) for an ATRP system can be defined, considering some approximations, by Equation 2.5, where the metal transition complex is copper (i.e., $\mathrm{Cu^{II}X/L}$ is the deactivator and $\mathrm{Cu^{I}/L}$ is the activator) (Matyjaszewski, 2012). Here, R_{p} depends on the propagation and equilibrium constants and the concentration of activator, deactivator, monomer and dormant species. Nevertheless, the structure of the ligand, the monomer, the pressure, the solvent, the temperature and the activity of the catalyst used, affect the K_{ATRP}. Consequently, all these factors will influence R_{p} (Matyjaszewski, 2012; Matyjaszewski and Davis, 2003; Santos, 2020).

$$R_{\mathrm{p}} = k_{\mathrm{p}}[\mathrm{M}][\mathrm{P}_n^{\bullet}] = k_{\mathrm{p}} K_{\mathrm{ATRP}} \frac{[\mathrm{P}_n\mathrm{X}][\mathrm{Cu^{I}/L}][\mathrm{M}]}{[\mathrm{Cu^{II}X/L}]} \tag{2.5}$$

The PDI equation can also be derived for an ATRP system with some approximations. It is given in Equation 2.6, where DP_n is the target degree of polymerization (Matyjaszewski, 2012). From this equation, it follows that the PDI depends on the propagation and deactivation constants, the degree of polymerization, the monomer conversion and the concentration of deactivator and dormant chains. This equation indicates that a higher concentration of deactivator or a higher DP_n or a lower concentration of dormant chains, will produce polymers with a smaller PDI value (Matyjaszewski, 2012). However, when an higher DP_n is targeted, termination and other side reactions increase their importance, and may negatively affect the PDI (Matyjaszewski and Davis, 2003).

$$PDI = 1 + \frac{1}{DP_n} + \left(\frac{k_{\mathrm{p}}[\mathrm{P}_n\mathrm{X}]}{k_{\mathrm{d1p}}[\mathrm{X{-}Cu^{II}/L}]}\right)\left(\frac{2}{x} - 1\right) \tag{2.6}$$

One of the majors drawbacks in the original ATRP polymerization is the high concentration of catalyst used, about 1000-10 000 ppm of cooper, which is not environmentally friendly and cost effective (Krys and Matyjaszewski, 2017). On top of that, this catalyst co-precipitates in the final polymers and can give them a coloured tone and turn them toxic. Therefore, it is very important to purify the final product for safety and aesthetic reasons (Shen et al., 2004). These drawbacks have created a need for the development of similar reactions, that use a much smaller concentration of catalyst (Abreu et al., 2017).

Based on these needs, new ATRP methods were developed in which a smaller amount of catalyst is used, but still managing to keep the polymerization under control. Most of these techniques are based on the continuous regeneration of $\mathrm{Cu^{I}/L}$. This is usually done by adding another redox cycle, which supplements the ATRP equilibrium (Krys and Matyjaszewski, 2017). The most commonly known methods are: activators regenerated by electron transfer ATRP (ARGET-ATRP); initiators for continuous activator regeneration ATRP (ICAR-ATRP); supplementary activators and reducing agents ATRP (SARA-ATRP); electrochemically meadiated ATRP (*e*ATRP); and photochemically meadiated ATRP (*photo*ATRP) (Abreu, 2018; Krys and Matyjaszewski, 2017; Lyra et al., 2019). In the forthcoming developments, the SARA-ATRP will be described in more detail.

2.4 SARA-ATRP

The SARA-ATRP method uses an agent (SARA agent) that is usually zero-valent cooper ($\mathrm{Cu^{0}}$), which is easily removable from the reaction medium after the polymerization. This component can regenerate $\mathrm{Cu^{I}/L}$ by a comporpotionation reaction with $\mathrm{X{-}Cu^{II}/L}$. It also serves as an supplemental activator of alkyl-halide initiators and dormant chains (Krys et al., 2017; Krys and Matyjaszewski, 2017).

The mechanism of the reaction with $\mathrm{Cu^{0}}$ has been deeply discussed in the scientific community. Percec et al. (2006) has suggest that $\mathrm{Cu^{0}}$ is totally responsibly for the activation of alkyl-halide initiators and dormant chains by an outer-sphere electron transfer (OSET). Being that $\mathrm{Cu^{I}/L}$ is only formed by $\mathrm{Cu^{0}}$, but rapidly disporpotionates back to $\mathrm{Cu^{0}}$ and $\mathrm{X{-}Cu^{II}/L}$ (Krys and Matyjaszewski, 2017). This point of view, of a reaction only mediated by $\mathrm{X{-}Cu^{II}/L}$ and $\mathrm{Cu^{0}}$, stands in contrast with the concept explained above, of a controlled polymerization due to an equilibrium that relies on the PRE (Abreu, 2018). The SARA-ATRP and SET-LRP (the name given to the mechanism suggested in Percec et al. (2006)) were the subject of many works, with the goal of reaching a definitive conclusion on what would be the correct mechanism (Abreu, 2018). To put an end to this discussion, an article with the title "SARA-ATRP or SET-LRP. End of controversy?" was published by the Matyjaszeski´s investigation group, that trough simulations, using PREDICI, and experimental results, were able to prove and support the SARA-ATRP mechanism. This means, that in fact, the ATRP equilibrium is the main reaction, and $\mathrm{Cu^{0}}$ only acts as a supplemental activator of alkyl-halide initiators and dormant chains (Abreu, 2018; Konkolewicz et al., 2014).

There has been a search for other alternatives of SARA agents, in order to ensure a more safe, inexpensive, environmental friendly and less toxic SARA-ATRP reaction (Abreu et al., 2012). Some alternatives have already been used to replace $\mathrm{Cu^{0}}$ such as $\mathrm{Fe^{0}}$ or $\mathrm{Ag^{0}}$, but the use of inorganic sulfites has caught the interest of the scientific community (Boyer et al., 2016; Krys and Matyjaszewski, 2017). Inorganic Salts, such as $\mathrm{Na_2S_2O_4}$, have been successfully employed

as SARA agents. These species are capable of supplemental activate the alkyl-halide initiators and dormant chains and to regenerate the activator, similarly to Cu^0, and still ensure a controlled polymerization (Abreu et al., 2012). These agents are approved by the U.S Food and Drug administration and are bio compatible, being possible to prepare polymers for bio-applications without purification of the final product (Boyer et al., 2016).

The mechanism of a SARA-ATRP using $\mathrm{Na_2S_2O_4}$, was proposed in Krys et al. (2017), where the authors report the estimation of the kinetic parameters and simulate the model proposed using the PREDICI software. The results show a good agreement between the model prediction and the experimental data, validating the mechanism proposed in that work. In comparison to a typical ATRP mechanism, more reactions were added, corresponding to the dissolution and dissociation of the inorganic sulfite and the supplemental activation of dormant chains and alkyl-halide initiators and the regeneration of $\mathrm{Cu^I/L}$. This will be addressed later in the model development in Chapter 4.

2.5 Mathematical modeling

Mathematical models bring many advantages to polymer science research, such as: the study and understanding of the polymerization reaction kinetic mechanism; the development of new processes for production of polymers with controlled characteristics; their implementation in advanced control applications to ensure a safe reaction and the final product desired quality; and to scale up polymerization processes (Asteasuain, 2018; Mastan and Zhu, 2015). One important point to highlight, is that modeling can not substitute the need for experiments, but if the two procedures are used, the time and cost of laboratory work can be reduce (Asteasuain, 2018).

In polymer science, the main modeling methods used can be divided into two groups; deterministic or kinetic-based; and statistic or stochastic (Asteasuain, 2018; Mastan and Zhu, 2015). The first ones are based on the derivation of population balances (i.e., reactants mass balances), in the form of a system of differential-algebraic equations. The solution of this set of equations gives the variation of specific proprieties along time. On the other hand, the statistic or stochastic modeling methods do not need the derivation of the population balances, because they are based on probability theories, being this modeling techniques more suitable for complex systems. One big disadvantage of these methods are their incapability of saving the reaction "history"(i.e., evolution along time). Here, the Kinetic Monte Carlo technique, which is the combination of deterministic and stochastic approaches, developed by Gillespie (1976, 1977) addresses this limitations. Nevertheless, these methods will need a long computational time to give accurate results and the objective of improving their computational efficiency is an area of active investigation (Asteasuain, 2018; Mastan and Zhu, 2015; Mastan et al., 2015). The modeling method used in this work is deterministic.

Using a deterministic method, it is necessary to write and obtain the result of a set of differential-algebraic equations. In a polymerization reaction there are different types of chains (e.g., dormant and propagating chains) and chains with different sizes (i.e., number of monomer units added), being necessary to derive the mass balance for each one of this chains (Mastan and Zhu, 2015). For example, if the goal of the polymerization is to obtain polymers with the maximum of 10000 monomer units added, there will be necessary to have 10000 differential equations and, if there are two types of polymers the number of equations raises to 20000. This makes the implementation and computational cost very high and not feasible. In order to address this limitation, the method of moments is adopted (Mastan and Zhu, 2015).

2.5.1 Method of moments

The method of moments allows the formulation of a model with a small number of differential-algebraic equations, decreasing the implementation effort and the computational cost of the simulation. One disadvantage of this method, is that only allows the calculation of average proprieties, such as M_n, and it is not possible to obtain a full distribution of molecular weights (Mastan and Zhu, 2015).

A statistical moment is defined by Equation 2.7, where k is the moment order and $[P_n]$ represents the concentration of propagating chains with n monomer repeating units (as an example). The 0th moment represents the total concentration of the polymer ($\sum_{n=1}^{\infty} n^k[\mathrm{P}_n{}^\bullet]$) and the 1st moment is the total concentration of all monomer units in all P_n chains ($\sum_{n=1}^{\infty} n^1[P_n]$) (Hungenberg and Wulkow, 2018; Vieira et al., 2015).

$$\lambda_k = [\mathrm{P}_1{}^\bullet] + \sum_{n=2}^{\infty} n^k[\mathrm{P}_n{}^\bullet] \tag{2.7}$$

Through the statistical moments it is possible to calculate the M_n and the weight average molecular weight (M_w) by Equation 2.8 and 2.9 (Hungenberg and Wulkow, 2018).

$$M_n = \frac{[\mathrm{P}_1{}^\bullet] + \sum_{n=2}^{\infty} n^1[\mathrm{P}_n{}^\bullet]}{[\mathrm{P}_1{}^\bullet] + \sum_{n=2}^{\infty} n^0[\mathrm{P}_n{}^\bullet]} M_M = \frac{\lambda_1}{\lambda_0} \tag{2.8}$$

$$M_w = \frac{[\mathrm{P}_1{}^\bullet] + \sum_{n=2}^{\infty} n^2[\mathrm{P}_n{}^\bullet]}{[\mathrm{P}_1] + \sum_{n=2}^{\infty} n^1[\mathrm{P}_n{}^\bullet]} M_M = \frac{\lambda_2}{\lambda_1} \tag{2.9}$$

By definition the PDI can be calculated through Equation 2.10.

$$PDI = \frac{\lambda_0 \lambda_2}{(\lambda_1)^2} \tag{2.10}$$

Also, it is very useful to obtain differential equations of the 0th, 1st and 2nd order moments to compute the values of M_n, M_w and PDI along the reaction time. This can be done by computing the time derivative of λ_k^p, which leads to Equation 2.11 (Hungenberg and Wulkow, 2018; Vieira

et al., 2015; Al-Harthi et al., 2006).

$$\frac{\mathrm{d}\lambda_k}{dt} = \frac{\mathrm{d}[\mathrm{P}_1^\bullet]}{dt} + \frac{\mathrm{d}}{dt}\left(\sum_{n=2}^{\infty} n^k [\mathrm{P}_n^\bullet]\right) \tag{2.11}$$

In the SARA-ATRP system under study, there are three types of polymers (i.e., monomer containing species), the dormant ($\mathrm{P}_n\mathrm{X}$), the propagating ($\mathrm{P}_n^\bullet$) and the dead (D_n). Thus, the total moment is the sum of the individual momentum of these three types of polymer chains. The notation in Table 2.2 will be used along this thesis to represent the momentum of each individual species (Lyra et al., 2019).

Table 2.2: Individual momentum of each species [Adapted from Lyra et al. (2019)].

Specie	momentum of kth order
μ_k	$\mu_k = [\mathrm{P}_1\mathrm{X}] + \sum_{n=2}^{\infty} n^k [\mathrm{P}_n\mathrm{X}]$
λ_k	$\lambda_k = [\mathrm{P}_1^\bullet] + \sum_{n=2}^{\infty} n^k [\mathrm{P}_n^\bullet]$
δ_k	$\delta_k = [\mathrm{D}_1] + \sum_{n=2}^{\infty} n^k [\mathrm{D}_n]$

Following the notation given in Table 2.2 the M_n, M_w and PDI can be calculated using Equations 2.12, 2.13 and 2.14, respectively.

$$M_n = \frac{\mu_1 + \lambda_1 + \delta_1}{\mu_0 + \lambda_0 + \delta_0} M_M \tag{2.12}$$

$$M_w = \frac{\mu_2 + \lambda_2 + \delta_2}{\mu_1 + \lambda_1 + \delta_1} M_M \tag{2.13}$$

$$PDI = \frac{(\mu_2 + \lambda_2 + \delta_2)(\mu_0 + \lambda_0 + \delta_0)}{(\mu_1 + \lambda_1 + \delta_1)^2} \tag{2.14}$$

It follows that it is necessary to obtain nine differential equations (0th, 1st and 2nd order momeents for the 3 species), to calculate the M_n, M_w and PDI along the reaction time. Also, it is easy to derive the equations for the other chemical species, such as the monomer, the activator or deactivator, etc., using the zero order moment. For example, the monomer mass balance is given by Equation 2.15 (considering that the only reaction is monomer addition to the propagating chain)

$$\frac{d\mathrm{M}}{dt} = -k_\mathrm{p}[\mathrm{M}] \sum_{n=1}^{\infty} [\mathrm{P}_n^\bullet] = -k_\mathrm{p}[\mathrm{M}]\lambda_0 \tag{2.15}$$

The method of the moments has been widely used in studies of the ATRP polymerization, not only to predict the monomer conversion and the PDI, but also to study the reaction kinetics and the sensitivity to the kinetic parameters and initial conditions. Some examples of these studies are briefly described bellow:

- In Lyra et al. (2019) the method of moments was used to predict the monomer conversion, the M_n and the PDI of an ARGET-ATRP system. Following the estimation of some kinetic

parameters, the model was able to make a good prediction of the experimental data. Then, it was used to study the kinetic model sensitivity to the initial conditions and the kinetic parameters, obtaining important conclusions of their effect on the conversion, M_n and PDI;

- In Vieira and Lona (2016) a model based on the method of moments was used to establish the limit temperature for the ATRP of styrene. However, the model prediction has not been tested with experimental data, an approximate limit value was specified through several simulations with different temperature values;

- In Vieira et al. (2015) the model was used to optimize the condition of a ATRP system. In order to obtain the values of temperature, catalyst and initiator concentration that can produce polymers with the lowest PDI but ensure the highest possible monomer consumption;

- In Al-Harthi et al. (2006) an ATRP system model using bifunctional initiators was formulated. The use of these initiators allow to increase simultaneously the rate of polymerization and the polymer final molecular weight, which is not possible with monofunctional initiators. The developed model was compared to a model with monofunctional initiators, confirming this theory;

- In Zhou and Luo (2014) a model was formulated to simulate a SARA-ATRP polymerization reaction, but using Cu^0 as the SARA agent. In this work, the model includes diffusion limitation, and studies the effect of the kinetic rate constants corresponding to the supplemental activation of propagating radicals/chains and regeneration of $\mathrm{Cu}^{\mathrm{I}}/\mathrm{L}$ by Cu^0 in the monomer conversion, M_n and PDI.

These previous works not only prove that the momentum method is a powerful tool for modeling polymerization reactions but the importance of modeling itself, and its utility to polymer science. Other works regarding the use of the method of moments can be found in the literature such as in Massicotte (2015) where ARGET-ATRP polymerization is modeled and simulated.

Chapter 3

Sensitivity analysis

In this chapter a brief introduction to the sensitivity analysis (SA) topic will be made, involving: the definition of SA; a brief presentation of commonly used methods for SA, being the SA methods addressed in this work analyzed in a more detailed way; discussion of some SA applications in the literature; and the settings used in the SA performed in this thesis.

3.1 Definition

Sensitivity analysis was defined by Saltelli (2002b) as the study of how the uncertainty in the output of a model can be distributed over different sources of uncertainty in the model inputs (e.g., model parameters). A sensitivity analysis is performed when one or more of the following objectives is intended to be achieved (Qian and Mahdi, 2020):

- Ranking the model inputs in function of there importance to the variability of the output. Those achieving the higher score are the focus of experimental and numerical estimation, because they are the ones that, if correctly determined, will lead to the largest decrease in output uncertainty;

- Identify which model inputs have little or no effect on the output variability. this allows the modeller to reduce the complexity of the model by not taking into account this inputs or setting them in fixed values;

- Mapping the effects of the inputs on the output, with the aim of finding the parameter values where the model is stable or where it achieves the optimum conditions.

Along this chapter the model output will be defined by the letter Y and the vector of inputs by X, being $Y = f(X_1, X_2, ..., X_i, ..., X_K)$. There are two main types of sensitivity methods: The local and global sensitivity methods.

3.2 Local sensitivity analysis methods

Local or one-at-time (OAT) sensitivity methods study the variation of a single input while the others are kept constant (Cariboni et al., 2007). This analysis can be done by changing each input manually by a margin of their nominal value. A possible sensitivity measure (S_i^L) can be calculated by Equation 3.1a, where Y_i^{max} represents the value of the model output generated by the maximum input value (X_i^{max}), being that Y_i^{min} and X_i^{min} have a similar definition (Saltelli et al., 2019). Another alternative to measure sensitivity, can be the calculation of a sensitive index by computing the partial derivatives of the output with respect to the input study, as represented in Equation 3.1b. Here, γ_i is a constant whose value change depending on the method used (for more information the reader is refereed to Qian and Mahdi (2020)) and X_i^* is the point where the derivative is evaluated (Saltelli et al., 2019). ´

$$S_i^L = \frac{Y_i^{max} - Y_i^{min}}{X_i^{max} - X_i^{min}} \tag{3.1a}$$

$$S_i^L = \gamma_i \frac{\partial Y}{\partial X_i}\Big|_{X=X_i^*} \tag{3.1b}$$

Local methods are relatively simple and, when applied, usually have a low computational cost. These approaches should not be used when the model features strongn nonlinearities. The reason for this, is that these methods do not cover all the input space (i.e., all possible combinations of input values). This is aggravated when the number of inputs studied increases (Qian and Mahdi, 2020; Saltelli et al., 2019). To illustrate this problem, an example is retrieved from Saltelli et al. (2019). For example, in a sensitivity analysis for the study of 3 inputs, being their range of variation of 1 unit, the input space can be represented by a three-dimensional cube with a side length of one (Figure 3.1). If the inputs are moved one at a time from the center with a distance of $\frac{1}{2}$, these points will be on the faces of the cube not exploring their corners. In reality, as explained in Saltelli et al. (2019), these points are on the surface of a sphere that is inside and tangent to the cube, represented by a red color in Figure 3.1. The ratio between the sphere and cube volume in this case is about 0.52, meaning that a lot of space is not explored. As one can imagine, due to the so-called "curse of dimensionality" when the number of inputs increases the volume ratio will decrease, expanding the unexplored space. For example, if there are 10 inputs the ratio between the "hypersphere" and the "hypercube" will be 0.0025, proving that local approaches should not be used, unless the model is linear. Another factor that needs to be accounted is that interactions between model inputs are only detected by moving all simultaneous, which is not done in local approaches (Saltelli and Annoni, 2010; Saltelli et al., 2019). In these cases, when the model is nonlinear, global methods should be used.

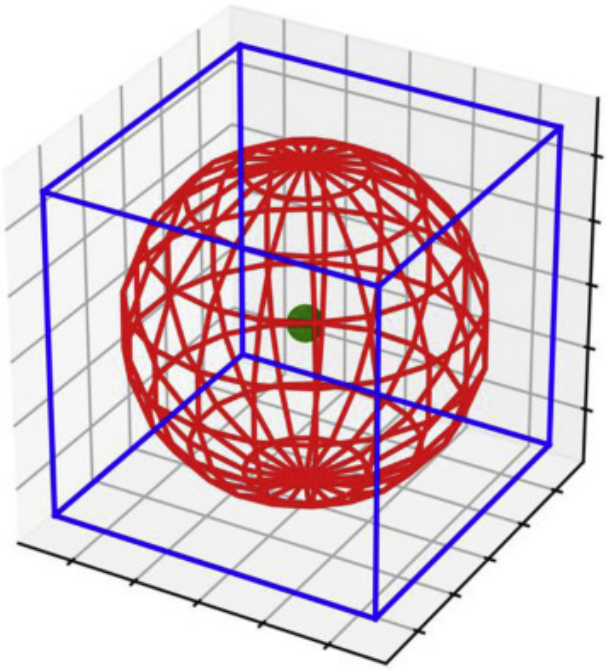


Figure 3.1: Graphical representation of a 3 dimensional input space [Retrieve from Saltelli et al. (2019)].

3.3 Global sensitivity analysis methods

In the global sensitivity analysis (GSA) methods the sensitivity is evaluated at multiple points of the input space. Then these sensitivity measures are aggregated, for example, by calculating the average value (Wang and Solomatine, 2019; Qian and Mahdi, 2020). There are several methods that can be used to perform a GSA, and the choice of which method to use, depends on the modeller objective, number of inputs, the output study and the audience to which the results will be presented (Saltelli et al., 2008). In this work, a brief introduction of the main global methods used will be done. Also, the theoretical and numerical procedure of the GSA methods applied in this thesis will be described in more detail. Although, there are many other methods that can be used for a GSA study. It would not be possible to carry out this work in due time, describing all. This thesis focus on three commonly applied methods reported in the works of Saltelli et al. (2008) and Wang and Solomatine (2019). Also, an argument in favour of these, is that they are a good starting point to initiate a GSA study, because of their simplicity and adequacy to handle non-linear models. The reader is refereed to the review of Qian and Mahdi (2020) for more details on the others methodologies.

3.3.1 Elementary effects method

The elementary effects (EE) or Morris method can be seen as a global version of the local approach described above, being one of the most used methods to perform GSA. The elementary effect of X_i can be defined by Equation 3.2, where Δ is the step-size or increment, of the

parameter variation (Yao, 2015; Qian and Mahdi, 2020).

$$EE_i = \frac{Y(X_1, X_2, ..., X_{i-1}, X_i + \Delta, ..., X_k) - Y(X_1, X_2, ..., X_{i-1}, X_i, ..., X_n)}{\Delta} \tag{3.2}$$

Basically the EE method approximates the sensitivity derivatives, in multiple points, by using finite differences with the objective of obtaining an distribution of elementary effects for each input. Then, the mean (μ_i) and the standard deviation (σ_i) of this distribution are calculated. μ_i measures the overall influence of the input, and σ_i gives information about if the input factor is non-linear or/and is taking part in interactions (Saltelli et al., 2019). The process to generate these distributions is simple. First, the inputs are rescaled to an uniform distribution between $[0, 1]$. Initially, the inputs are set randomly inside this interval, then one factor is incremented or decremented (by a Δ) and Equation 3.2 is applied. From this new value of EE, a different input is changed. This is done until it is calculated one elementary effect for each input. Then, the process will be repeated by setting the inputs in different starting values, obtaining r elementary effects per input, and computing the corresponding distribution (Qian and Mahdi, 2020). This method has a medium computational cost, being mainly used for factor screening (i.e., deciding which are the most important model parameters) when there is a high number of inputs, between 20-100 (Saltelli et al., 2008; Qian and Mahdi, 2020).

3.3.2 Variance based methods

Variance based sensitivity analysis (VSA) methods are the most popular technique used to perform a GSA (Wang and Solomatine, 2019). These methods calculate the proportion of the model variance caused by each input factor. This includes their single effect and interactions with 2 or more inputs (Wang and Solomatine, 2019; Qian and Mahdi, 2020). These methods are based on the calculation of sensitivity indices, being the value of the single effect (given by the value of the first order sensitivity index (S_i)) for the input X_i calculated by Equation 3.3. The meaning of the elements in this equation will be further discussed in the next section.

$$S_i = \frac{V_{X_i}(E_{X_i}(Y|X_i))}{V(Y)} \tag{3.3}$$

The calculation of the sensitivity indices will need the computation of multidimensional integrals, by brute force methods, which is not feasible due to long computational time needed, especially for complex models (Saltelli et al., 2008; Wang and Solomatine, 2019). In order to address these limitations, methods like Sobol and the Fourier amplitude sensitivity test (FAST) were developed. The Sobol method is an efficient sample based approach, where the model variance is decomposed into contributions from each input and their interactions, being the most popular variance based method (Wang and Solomatine, 2019; Qian and Mahdi, 2020). This type of approach is usually done for ranking and factor screening, and despite the new alternative methods, such as Sobol, the computational cost is still high and their used is only

recommended when the input factors are less that 20 (Saltelli et al., 2008). In order to reduce the computational effort required by these methodologies, some authors use approaches based on metamodeling. This topic is out of the scope of this thesis. For further details on this metamodeling approaches the reader is refereed to Saltelli et al. (2008) or Wang and Solomatine (2019).

3.3.3 Regionalized sensitivity analysis

Regionalized sensitivity analysis (RSA) is an approach where it is intended to find the region of parameter values that lead to a certain behaviour of the model, and to identify which parameters are responsible for not obtaining that particular behaviour (Wu et al., 2017; Wang and Solomatine, 2019). An important step in this method, is to define what is the "good" behaviour and the "bad" behaviour of the model. For example, in a polymerization reaction the behavioural output could be a PDI smaller that 1.10, and all the values superior to 1.10 would be considered non-behavioural. Then, the model is evaluated with different input values and the simulation results are split into two groups: those who are in the good behaviour zone; and those who are not (Wang and Solomatine, 2019). Next, two Cumulative Distribution Functions (CDFs) are generated for all the inputs. One with the input values that generate a behavioural output and other with the ones that generate a non-behavioural output. This CDFs can be used to determined the importance of each input and to locate the "good" parameter values (this will be explained in more detail in the next section) (Wang and Solomatine, 2019). An usual method of RSA is the Monte Carlo filtering (MCF). This method is described, in detail, in two books of Andrea Saltelli (Saltelli et al. (2004, 2008)). One disadvantage of these methods, is that many interactions between the inputs are not detected. RSA is usually used for factor mapping (i.e., see where the "good" values of the inputs are), and has a medium computational cost (compared to EE and Sobol methods). Its application is usually recommended for less than 20 inputs (Saltelli et al., 2008).

3.4 Methods comparison

In Table 3.1 is represented a brief comparison between the sensitivity analysis methods presented in this thesis.

Table 3.1: Comparison between the methods used for SA (Saltelli et al., 2008; Qian and Mahdi, 2020; Yao, 2015).

Methods	Local methods	EE methods	VSA methods	RSA methods
Number of inputs	No information	20-100	<20	<20
Computational cost	Low	Medium	High	High
Coping with non-linearity	-	+	+	+
Detection of interactions	-	+	+	+/-
Difficulty of the analysis	Low*	Medium*	High	Medium

* Expected difficulty. Only the VSA and RSA methods were implemented in this thesis

3.5 Applications of sensitivity analysis

Applications of sensitivity analysis can be found in the most varied fields, not only in exact sciences, but also in economics and social sciences (Saltelli et al., 2019). In the literature, it is possible to find very interesting works that deal with current problems. A work that stood out in this research, due to the present world situation, was the GSA in Zhang et al. (2021) of a COVID-19 model using the Sobol method. The authors main goal was to identify the inputs that are the most responsible for the 4 outputs uncertainties (Susceptible Population, Exposed individuals, Infected population and Dead bodies), with the aim that these results can help governments, or other entities, to explore various policy options to control the COVID-19 pandemic. This is in line with a recent technical report by the European commission, entitled "Uncertainty and Sensitivity Analysis for policy decision making", which tries to promote the use of SA and Uncertainty analysis in models that assist policy decisions (Azzini et al., 2020).

Another field of application of SA, which is an important topic nowadays, is in environmental science (EnvSci), being currently the subject that release the most amount of SA related papers (Saltelli et al., 2019). One example of the application in EnvSci is reported by Campolongo and Braddock (1999), where a GSA is performed, using a Morris method based approach, in a model that predicts the climate change as a result of the emission of greenhouse gases by human activity. At the end of the analysis, the authors conclude, what are the variables whose measurement should be improved in order to improve the model prediction capacity. Another example in this field is an application that brings together EnvSci and economy in Bolker et al. (2021). Here a GSA is performed on a integrated climate-economic model, that is basically an economic model that takes into consideration the climate dynamics, which can severally impact

the economy. In this type of models SA is very important, because small changes in input values can cause large variations in the output, so their relative importance must be investigated. These three examples clearly show the importance and how valuable can be SA to assist in current world problems.

The use of SA methodologies is also present in Engineering related articles, such as in the work of Wu et al. (2017). In this Civil Engineering paper a SA is performed in a model that predicts pavement distress, being a important tool for pavement design. For the SA the author uses the MCF method, because the designers are interested to explore the zone/"region" of the model output around the design limit and thus finding out the parameters that are most responsible for exceeding that design criteria. In the end, the author obtains information on the inputs ranking, which matches the conclusion of other authors using different SA methods. With the a additional information about whether the effect of the inputs in satisfying the design criteria is positive or negative across the entire range of variation (i.e., all possible values of the input space). This one of the advantages, stated by the authors, for the use of RSA methods in this specific case. Other works can be found in completely different areas of engineering, such as in Biomedical Engineering. One such example is the article Zi et al. (2008), that reports the development of a software to perform GSA in biological models. In fact a SA is so important in this area that a recent review of SA methods untitled "Sensitivity analysis methods in the biomedical sciences" was published in 2020 (Qian and Mahdi, 2020).

In **Chemical Engineering** it is also possible to find some works that preform SA in their models. In fact, in the book of Saltelli et al. (2008), a chemical reactor is used as a case study to exemplify the use of GSA techniques. In these example, the thermal runaway of a batch reactor is investigated, with the aim of finding the model inputs that will offer the better chance of reducing the temperature variance. This GSA study was performed using several techniques. This can also be looked as a example of how SA can help to ensure chemical process safety, by investigate which inputs are the most responsible for creating dangerous conditions. Another example, is a SA performed on a Spray Drying unit by Bhonsale et al. (2019). The main objective of this article is to demonstrate that GSA can be a model based approach to apply the principles of quality by design paradigm (QbD) in the Pharmaceutical industry, where the Spray dry unit is used.

Addressing specifically the area of **polymer sciences** reaction modeling, it is only possible to find a limit amount of papers, modeling-related, that uses GSA techniques. One of the works is described in Salas et al. (2019). Here, is intended to make an online kinetic parameter estimation of a polymerization to synthesize Polyolefins (i.e., polypropylene and polyethylene) in a semi-batch reactor, by measuring the inlet of ethylene flow rate. This model has large number of kinetic parameters, as it is common in polymerization reactions. Thus, the authors decided to make a GSA to identify only the most important parameters to be estimated, keeping the others in their nominal value. Using the Sobol method only two parameters were considered important

to the model output. As a consequence, this approach decreased the computational cost of the kinetic parameter estimation problem. Another work that uses GSA is published in Zubov and Sin (2018) of a poly lactid acid (PLA) production model. The main goal is to develop a model that relates the reactions conditions with the rheological proprieties in the final product. In order to obtain a modeling framework that allows the development of PLA based materials with specific rheological characteristics. A key part in this work is a GSA with the goal of obtaining insights into the contribution of the model inputs to the predicted polymer quality.

Nevertheless, most of the papers in polymerization modeling use local methods. An interesting example, due to the topic of these work, is a SA performed in a RDRP, by computing the sensitivity derivatives in order to asses which parameters have the greatest importance to the monomer conversion, M_n, M_w and the PDI (Flores-Tlacuahuac et al., 2003). From this analysis the authors conclude that for the monomer conversion and M_n the most important kinetic parameters are the propagation, initiation and termination constant, being moderately influenced by the equilibrium parameters (activation and deactivation rate constants). These last two parameters are the most important for the M_w and PDI. Nevertheless, other works can be found that also used derivatives, such as in Thomas and Kiparissides (1984), where a SA is perform in a free-radical polymerization model. In Lyra et al. (2019) a SA is applied by using the one-at-time method.

These works present important conclusions, but as explained above local methods only explore a very confined space of all possible input combinations. In addition, the polymerization kinetic models are normally constituted by several parameters and non-linear. Thus, global methods must be used to make a correct SA. This preference for local methods is not only a problem of polymer science modeling but can also be visualized in other areas. For example, in a recent study presented in Saltelli et al. (2019), were analysed 28 articles that perform SA in chemical engineering related models. Only 12 of this works used GSA methods being that 16 models were non-linear (for the rest of the models analysed it was not possible to determine whether they are linear or non-linear). This lack of use of global techniques and possible bad practices when a SA is realized, can be due to numerous factors, such as (Saltelli et al., 2019): sensitivity analysis is not considered an unified discipline in the various branches of modeling; lack of good statistical background by the modelers; the meaning of sensitivity analysis is not understood; afraid of finding possible errors in the model; lack of comparative studies in the various areas of modeling.

3.6 Detailed methods

In this work is intended to make a GSA to discover the most important kinetic parameters in SARA-ATRP system by exploring all the parameter space. This will be done in order to gain extra knowledge on the reaction mechanism, with the intention of this results being further used

in possible model improvements. Proving that GSA can be an useful and an easy to apply tool in polymerization modeling. For this matter it was decided to use two different methods: the variance based Sobol method and the regional based Monte Carlo filtering. In order to compare and show two different possibilities that the modelers could apply in their model. These two methods will be described in detail in the next subsections.

3.6.1 Monte Carlo filtering

In MCF the values of the model output variable under analysis are classified into two different realizations, the behavioural output (B) and the non-behavioural output ($\overline{B}$). The behavioural realization is where the output result meets the requirements pre-established by the modeller, i.e., it is in the good behaviour zone of the model. Other outputs outside this zone belong to the non-behavioural domain because they not fulfil the model target (Saltelli, 2002a; Saltelli et al., 2008).

The MCF procedure is simple. First several input values are generated and propagated into the model and, a response is obtained for eah combination of parameters. Then, according to the model target, each output obtained is categorized belonging to the B or $\overline{B}$ group. In the next step the values of the parameters that originates a model output in the B zone are separated from the ones that do not fulfil the model target, creating two subsets for each parameter, the subsets $(X_i|B)$ and $(X_i|\overline{B})$. These subsets are from two different unknown probability density functions, $f_n(X_i|B)$ and $f_n(X_i|\overline{B})$. In order to find the most important parameters $f_n(X_i|B)$ and $f_{\overline{n}}(X_i|\overline{B})$ are compared. If for a given parameter there is a significant difference between the two distributions, it means that the input is important for driving the model to the target behaviour (Saltelli et al., 2008).

For the comparison between the two probability density functions the statistical two sample Kolmogorov-Smirnov (K-S) test is applied. This is a hypothesis test used to quantify the difference between two cumulative distribution functions (CDFs), $F_n(X_i|B)$ and $F_n(X_i|\overline{B})$. The null hypothesis of this test is that the two CDFs are equal (Equation 3.4) and the alternative hypothesis is that they are different (Equation 3.5) (Wu et al., 2017; Saltelli et al., 2008)

$$H_0 : F_n(X_i|B) = F_{\overline{n}}(X_i|\overline{B}) \tag{3.4}$$

$$H_1 : F_n(X_i|B) \neq F_{\overline{n}}(X_i|\overline{B}) \tag{3.5}$$

The two sample K-S test gives two results, the $d_{n,\overline{n}}$ which represents the maximum distance between the two CDFs (Equation 3.6) and the $p - \text{value}$ which is the probability, considering that the null hypothesis is true, to obtain a $d_{n,\overline{n}}$ as large as the one obtained between the two distributions studied.

$$d_{n,\overline{n}}(X_i) = sup \mid F_n(X_i|B) - F_{\overline{n}}(X_i|\overline{B}) \mid \tag{3.6}$$

To classify the parameters, according to the p-value, the strategy represented in the Table 3.2

was used. This classification is recommended by Saltelli et al. (2004) and is used in the work of Wu et al. (2017). The $d_{n,\overline{n}}$ statistic can be used as a sensitive measure to rank the input parameters, but it must be taken into consideration that this value often does not take into account input interactions (Yao, 2015; Wang and Solomatine, 2019). The model interaction structure can be briefly analysed by a bivariate correlation analysis of the filtered samples (i.e., by calculating the correlation coefficient between $f_n(X_i|B)$ of each pair of parameters) (Saltelli et al., 2004, 2008). This analysis will be performed in this work following an example presented in Saltelli et al. (2004). Performing this analysis, when is stated that the inputs interact, it means that they interact in order to obtain a behavioural output (for more information the reader is refereed to Saltelli et al. (2004)).

Table 3.2: Parameter classification in function of the p value.

p value	Classification
$< 1\%$	Critical
$1\% - 10\%$	Important
$> 10\%$	Insignificant

The CDFs can give an important information, about the range of the input space where there is more probability of obtaining an output in the good behaviour zone. In order to explain how this can be visualised trough the CDFs, Figure 3.2 is taken as an example. In this figure is represented a plot of the two CDFs, $F_n(X_i|B)$ and $F_{\overline{n}}(X_i|\overline{B})$, for a model parameter study in this work. In this figure, is possible to analyse that in low values of the input the $F_n(X_i|B)$ curve is steeper than $F_{\overline{n}}(X_i|\overline{B})$. This indicates that there is more probability of obtaining an output in the behavioural zone by using low values of the input range (i.e., between 4.85×10^{-6} and 1.21×10^{-5}, in this example) (Yao, 2015; Saltelli et al., 2008). This is very helpful in a way that can give the modeler information on what range of values of the input should be used, in order to obtain a behavioural response. If the slope of the $F_n(X_i|B)$ is zero it means that if it is used a input value in that range, there will be zero probability of obtaining a behavioural output.

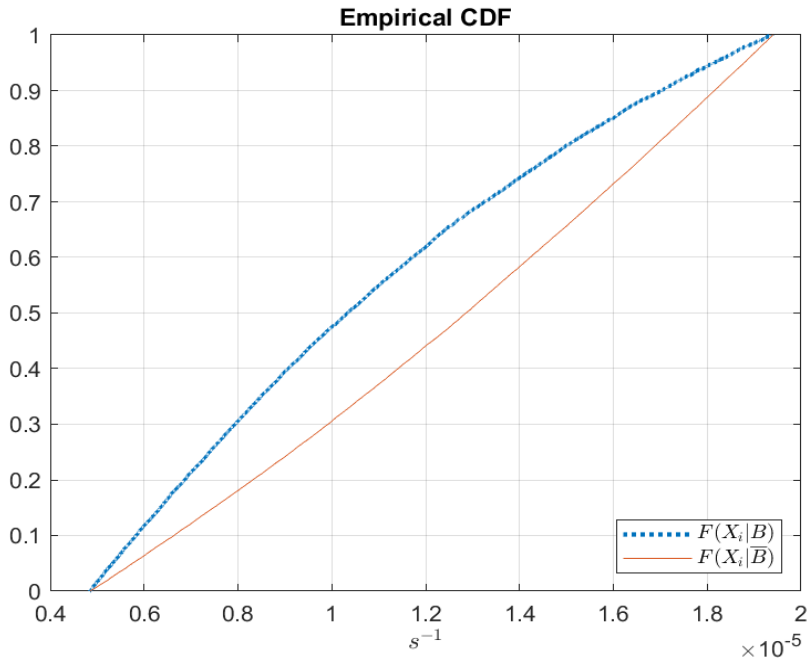


Figure 3.2: Example of a CDFs plot.

3.6.2 Sobol method

The Sobol method is part of the variance based global sensitivity analysis methods and it was developed by the Russian mathematician Ilya Sobol and first publish in Sobol (1990). This method was developed in order to accelerate the computation of the sensitivity indices by using an efficient sample based approach (Wang and Solomatine, 2019). When all the model parameters vary within a determinate range, the uncertainty of the output can be quantified by the output unconditional variance ($V(Y)$). $V(Y)$ can be defined by (3.7), trough the elements of this equation it is possible to study which are the most influential parameters (Ochoa et al., 2016; Saltelli et al., 2008).

$$V(Y) = E_{X_i}(V_{X_i}(Y|X_i)) + V_{X_i}(E_{X_i}(Y|X_i)) \tag{3.7}$$

The meaning of $E_{X_i}(V_{X_i}(Y|X_i))$ and $V_{X_i}(E_{X_i}(Y|X_i))$ is quite difficult to understand and visualise at a glance, especially when one does not have a good statistical background. But in a simple way, and by the explanation given in Ochoa et al. (2016), $E_{X_i}(V_{X_i}(Y|X_i))$ is the expected value over all the possible values of the parameter X_i of the conditional variance of Y under the variation of all the parameters except X_i. On the other hand,$V_{X_i}(E_{X_i}(Y|X_i))$ can be defined as the variance over all possible realizations of the parameter X_i of the conditional expected value of the output Y under the variation of all the parameters expect X_i (Ochoa et al., 2016). These metrics can give useful information of the input importance, because if $E_{X_i}(V_{X_i}(Y|X_i))$ is small and $V_{X_i}(E_{X_i}(Y|X_i))$ is higher, it means that the input is important (Note: This again can be hard to mentally visualize why) (Saltelli et al., 2008). To measure the importance of a given input

(X_i) in relation to the output, the first order sensitivity index (S_i) can be calculated by (3.3). This sensitivity index varies between 0-1, and the higher the value, the more important the parameter is (Saltelli et al., 2008).

Ilya Sobol proposed a way of calculating this index using a Monte-Carlo based approach, evaluating the function $Y = f(X_1, X_2, ..., X_i, ..., X_K)$ multiple times with different input values. The demonstrations made by Sobol can be found in the Appendix A in Section A.1.

These demonstrations also enable the calculation of the total Sensitivity indices (S_{T_i}), which accounts for the total contribution of a given input to the variation of the model output, i.e., it takes into account both the individual contribution of the parameter as well as its interactions with the other inputs. The total sensitivity indices were not explicit introduced by Sobol but rather by Tomshimitsu Homma and Andrea Saltelli in Homma and Saltelli (1996) (Saltelli et al., 2008). In addition to (3.7) $V(Y)$ can be calculated by Equation 3.8 (Saltelli et al., 2008).

$$V(Y) = E_{X_i}(V_{X_i}(Y|X_{\sim i})) + V_{X_i}(E_{X_i}(Y|X_{\sim i})) \tag{3.8}$$

The definition of $E_{X_i}(V_{X_i}(Y|X_{\sim i}))$ and $V_{X_i}(E_{X_i}(Y|X_{\sim i}))$ in (3.8) is very similar to the definition of $E_{X_i}(V_{X_i}(Y|X_i))$ and $V_{X_i}(E_{X_i}(Y|X_i))$, except that for this case the conditional variance and the conditional expected value is calculated by fixing all model inputs expect X_i. Dividing (3.8) by $V(Y)$ the total sensitivity index can be calculated by Equation 3.9 (Saltelli et al., 2008).

$$S_{T_i} = \frac{E_{X_i}(V_{X_i}(Y|X_{\sim i})}{V(Y)} = 1 - \frac{V_{X_i}(E_{X_i}(Y|X_{\sim i}))}{V(Y)} \tag{3.9}$$

Analysing $E_{X_i}(V_{X_i}(Y|X_{\sim i}))$ and $V_{X_i}(E_{X_i}(Y|X_{\sim i}))$ and Equation 3.9 it is possible to verify that this two metrics have a different behaviour that the one verified for $E_{X_i}(V_{X_i}(Y|X_i))$ and $V_{X_i}(E_{X_i}(Y|X_i))$.

Saltelli et al. (2008) describe a Monte-Carlo based numerical procedure, derived from the original method proposed in (Sobol, 1990), to calculate the first order and the total order indices. This method and its mathematical formulation were develop by Saltelli (2002a). The numerical procedure starts by generating a matrix of N lines and $2K$ (K is the number of inputs under studied) columns with the sampling of the model parameters. Then the matrix is divided in two matrices, $A(N, K)$ and $B(N, K)$. Using the notation of Saltelli et al. (2008), these matrices are given by:

$$A = \begin{pmatrix} X_1^{(1)} & X_2^{(1)} & ... & X_i^{(1)} & ... & X_K^1 \\ X_1^{(2)} & X_2^{(2)} & ... & X_i^{(2)} & ... & X_K^2 \\ ... & ... & ... & ... & ... & ... \\ X_1^{(N-1)} & X_2^{(N-1)} & ... & X_i^{(N-1)} & ... & X_K^{N-1} \\ X_1^{(N)} & X_2^{(N)} & ... & X_i^{(N)} & ... & X_K^N \end{pmatrix} \tag{3.10}$$

$$B = \begin{pmatrix} X_{K+1}^{(1)} & X_{K+2}^{(1)} & \dots & X_{K+i}^{(1)} & \dots & X_{2K}^{1} \\ X_{K+1}^{(2)} & X_{K+2}^{(2)} & \dots & X_{K+i}^{(2)} & \dots & X_{2K}^{2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ X_{K+1}^{(N-1)} & X_{K+2}^{(N-1)} & \dots & X_{K+i}^{(N-1)} & \dots & X_{2K}^{N-1} \\ X_{K+1}^{(N)} & X_{K+2}^{(N)} & \dots & X_{K+i}^{(N)} & \dots & X_{2K}^{N} \end{pmatrix} \tag{3.11}$$

Then, another matrix, $C_i(N, K)$, is defined for each of one of the parameters, with all columns of B except the column corresponding to the parameter which is replaced by the corresponding column of matrix A. An example of its form for the ith parameter is represented in (3.12).

$$C_i = \begin{pmatrix} X_{K+1}^{(1)} & X_{K+2}^{(1)} & \dots & X_{i}^{(1)} & \dots & X_{2K}^{1} \\ X_{K+1}^{(2)} & X_{K+2}^{(2)} & \dots & X_{i}^{(2)} & \dots & X_{2K}^{2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ X_{K+1}^{(N-1)} & X_{K+2}^{(N-1)} & \dots & X_{i}^{(N-1)} & \dots & X_{2K}^{N-1} \\ X_{K+1}^{(N)} & X_{K+2}^{(N)} & \dots & X_{i}^{(N)} & \dots & X_{2K}^{N} \end{pmatrix} \tag{3.12}$$

Then, the model output is computed for each line of the matrices (That have a different combination of model input values) A, B and C, individually. In the end, there will be vector of dimension N for which one of the matrices with the model output result (Equation 3.13).

$$y_A = f(A) \qquad y_B = f(B) \qquad y_{C_i} = f(C_i) \tag{3.13}$$

With these three vectors it is possible to obtain the first order and the total sensitivity indices. Trough some demonstrations, the sensitivity indices can be calculated by Equation 3.14. For more information see Appendix A in Section A.2 and (Saltelli, 2002a).

$$S_i = \frac{Ui - E^2(Y)}{V(Y)} \tag{3.14}$$

The elements of Equation 3.14, have the advantage of being possible to approximate their value by Monte Carlo methods. Then, $E(Y)$ can be approximated through the mean value as the statistical estimator, using the values of $y(A)$ given by Equation 3.15 (Note: $y(B)$ can also be used but in Saltelli et al. (2008) the authors choose to use the formula with the values of $y(A)$, being the same notation used in this work) (Saltelli, 2002a; Wu et al., 2012).

$$E(Y) \approx \hat{f}_0 = \frac{1}{N} \sum_{j=1}^{N} y_A^j \tag{3.15}$$

U_i in Equation 3.14 can be approximated by the estimator $\hat{U}_i$. This estimator can be calculated by Equation 3.16 (Saltelli, 2002a; Saltelli et al., 2008).

$$U_i \approx \hat{U}_i = \frac{1}{N} \sum_{j=1}^{N} y_A^j y_{C_i}^j \tag{3.16}$$

The $V(Y)$ is approximated by Equation 3.17, where $\hat{V}$ is calculated by (3.18).

$$V(y) \approx \hat{V} - \hat{f}_0 \tag{3.17}$$

$$\hat{V} = \frac{1}{N}\sum_{j=1}^{N}(y_A^j)^2 \tag{3.18}$$

To calculate the total indices, Homma and Saltelli took advantage of the work done by Sobol. In Homma and Saltelli (1996) it was proposed the estimator $U_{\sim i}$, for the calculation of these indices (3.19), where the value of $U_{\sim i}$ can be approximated by $\hat{U}_{\sim i}$ (Homma and Saltelli, 1996; Saltelli, 2002a).

$$S_{T_i} = 1 - \frac{\hat{U}_{\sim i} - E^2(Y)}{V(Y)} \tag{3.19}$$

In Saltelli et al. (2008) it was proposed to calculate $\hat{U}_{\sim i}$ using Equation 3.20.

$$U_{\sim i} \approx \hat{U}_{\sim i} = \frac{1}{N}\sum_{j=1}^{N} y_B^j y_{C_i}^j \tag{3.20}$$

At first, in this thesis the method presented in Saltelli et al. (2008) was implemented in MATLAB®. Later on, after analysing the existing literature, it was found some modification to this methods proposed by Saltelli, with the aim of improving it. One of the most interesting modification is presented at Wu et al. (2012). This method was developed with the aim of increasing the computing efficiency and to accelerate the convergence of the classic Saltelli numerical procedure. This method was already used in a sensitivity analysis performed on a co-polimerization reaction in Salas et al. (2019). In particular, in this method another matrix is defined, which, contrary to the matrix C_i, has all columns of A except the column corresponding to the parameter which is taken from B. In this thesis, this matrix is represented by the letter D_i. An example of its form for the ith parameter is represented in (3.21).

$$D_i = \begin{pmatrix} X_1^{(1)} & X_2^{(1)} & \dots & X_{k+i}^{(1)} & \dots & X_K^1 \\ X_1^{(2)} & X_2^{(2)} & \dots & X_{k+i}^{(2)} & \dots & X_K^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ X_1^{(N-1)} & X_2^{(N-1)} & \dots & X_{k+i}^{(N-1)} & \dots & X_K^{N-1} \\ X_1^{(N)} & X_2^{(N)} & \dots & X_{k+i}^{(N)} & \dots & X_K^N \end{pmatrix} \tag{3.21}$$

Wu et al. (2012) have also proposed new formulas for the statistical estimators. These are given in Table 3.3, where they can be compared with the original formulation.

Table 3.3: Statistical estimators.

Estimator	Saltelli et al. (2008)	Wu et al. (2012)
$\hat{f}_0$	$\frac{1}{N}\sum_{j=1}^{N} y_A^j$	$\frac{1}{2N}\sum_{j=1}^{N}(y_A^j + y_B^j)$
$\hat{U}_i$	$\frac{1}{N}\sum_{j=1}^{N} y_A^j y_{C_i}^j$	$\frac{1}{2N}\sum_{j=1}^{N}(y_A^j y_{C_i}^j + y_B^j y_{D_i}^j)$
$\hat{U}_{\sim i}$	$\frac{1}{N}\sum_{j=1}^{N} y_B^j y_{C_i}^j$	$\frac{1}{2N}\sum_{j=1}^{N}(y_A^j y_{D_i}^j + y_B^j y_{C_i}^j)$
$\hat{V}_i$	$\frac{1}{N}\sum_{j=1}^{N}(y_A^j)^2$	$\frac{1}{2N}\sum_{j=1}^{N}((y_A^j)^2 + (y_B^j)^2)$

In Homma and Saltelli (1996) it is also stated that better estimates of the first and total order Sobol indices, can be accomplished by using the $\hat{\gamma}$ estimator for calculating the numerator terms in (3.3) and (3.9), as follows:

$$V_{X_i}(E_{X_i}(Y|X_i)) = U_i - \gamma^2 \tag{3.22}$$

$$V_{X_i}(E_{X_i}(Y|X_{\sim i})) = \hat{U}_{\sim i} - \gamma^2 \tag{3.23}$$

This estimator is not used in the method presented in Saltelli et al. (2008), but in Wu et al. (2012) a new formula is proposed, which is given in (3.24).

$$\gamma^2 \approx \hat{\gamma}^2 = \frac{1}{2N}\sum_{j=1}^{N}(y_A^j y_B^j + y_{C_i}^j y_{D_i}^j) \tag{3.24}$$

In this thesis, it will be used the γ to calculate the sensitivity indices, when using the method of Wu et al. (2012). At first glance, one of the disadvantages of the new modification could be the computational effort necessary compared to the one presented in Saltelli et al. (2008). This is because in the later the number of model evaluations is $N(K+2)$ whereas for this new method is $N(2K+2)$. For a $N = 1000$ and with 13 parameters the first method will need 15000 model evaluations and 2nd method 28000 model evaluations. However, it should be noted that this method was also developed to accelerate the convergence of the indices. Thus, it is expected that using the new strategy of Wu et al. (2012) will require a smaller N. Further on, the two methods will be compared by using a simple analytical function as a case study.

It is important to mention some properties of the sensitivity indices in order ensure a correct interpretation of the results of the sensitivity analysis (Saltelli et al., 2008), such as:

- S_i indicates how much, on average, the output variance can be reduced if X_i could be fixed;
- S_{T_i} is always equal or larger than S_i, and when it is equal it means that X_i has no interactions with other inputs;
- If $S_{T_i} = 0$ it means that X_i is a non-influential factor that does not affect the output variance;
- $\sum_i S_i$ is always equal or smaller than one, and $\sum_i S_{T_i}$ is always equal or higher than one;
- $1 - \sum_i S_i$ is an indicator of the presence of interactions in the model.

To assist in the interpretation of the results of the sensitivity indices, a threshold of 0.05 is applied. Therefore, if the first order and the total sensitivity index is less than 0.05 the parameter is considered unimportant, i.e., the output is considered to be independent of that input. This criteria was adopted in this thesis from Bhonsale et al. (2019).

3.7 Settings for GSA

In this section it will be presented some important settings that need to be defined before starting the sensitivity analysis, such as: the model output; the parameters input space; and the sampling method.

3.7.1 Model output

The model output for this GSA will be defined as the error between the experimental data and the model predictions. This is in fact a widely used strategy in GSA instead of calculating the sensitivity of the model output directly (Wang and Solomatine, 2019). In this thesis, the error will be evaluated with respect to the monomer conversion, PDI and M_n. To do so, the normalized root mean square error (NRMSE) represented in (3.25) is used, where $y_{exp,\,i}$ is the ith experimental value, $y_{prev,\,i}$ is the ith value predicted by the model, N_{exp} is the number of experimental points and $\overline{y_{exp}}$ is the average of the experimental values. It was decided to use this error instead of the root mean square error (RMSE), as in Wang and Solomatine (2019), due to the fact that the outputs have very different scales of variation.

$$NRMSE = \frac{\sqrt{\sum_{i=1}^{N_{exp}} \frac{(y_{\mathrm{exp,\,i}} - y_{\mathrm{prev},\,i})^2}{N_{\mathrm{exp}}}}}{\overline{y_{\mathrm{exp}}}} \tag{3.25}$$

Other choices of outputs can be found in the literature for dynamic models. For example in Ochoa et al. (2016) and Qian and Mahdi (2020) the authors used the Sobol method and calculated the total and first order sensitivity indices at several time points, obtaining their variation along time. This approach allows to verify how the importance of the inputs changes with time. However, it was concluded that this methodology would be much more difficult to implement, would require more computational time and the discussion of results would be more difficult. Hence, it was decided that using the NRMSE would be the best alternative. Once the output of the model is defined, is possible to specify the threshold between the behavioural and non-behavioural for the MCF method application. Given that the model will be simulated first, with literature parameters, the NRMSE will be calculated for these simulations. Tanking into account that one of the goals of the SA is to identified critical parameters for model improvements, it makes sense that a behavioural response would correspond to a NRMSE lower that the one obtained with simulations using the original kinetic parameters. Also, note that the threshold for the 3 outputs can be only specified after these first simulations.

3.7.2 Input space

Defining the input space is a difficult part of sampling-based GSA, because the range o variation could be unknown. This is particularly difficult in relatively new reaction such as the one studied in this thesis. In Lyra et al. (2019) the author does a sensitivity analysis of a ARGET-ATRP system by one-factor-at-time, changing the deactivation, activation and the reduction constant between $0.1 \times$ *nominal value* and $10 \times$ *nominal value*. In this work, it was decided to use the same procedure for the kinetic parameters study. In fact, changes in the input space study can severely affect the outcome of the GSA (Qian and Mahdi, 2020). Because at this stage there is no information on real parameter variation, it was decided to test two more intervals with a smaller size (see Table 3.4).

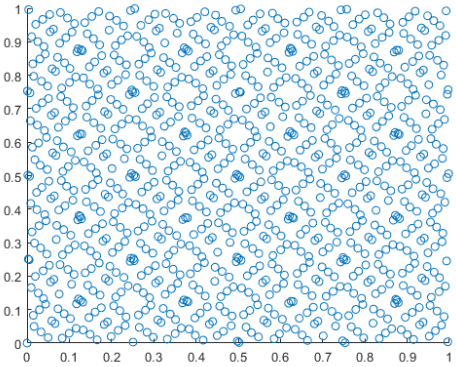
Table 3.4: Parameters input spaces for the sensitivity analysis.

Interval	Lower Bound	Upper bound
1st	$\frac{1}{10}\times$ Nominal value	$10\times$ Nominal value
2nd	$\frac{1}{5}\times$ Nominal value	$5\times$ Nominal value
3rd	$\frac{1}{2}\times$ Nominal value	$2\times$ Nominal value

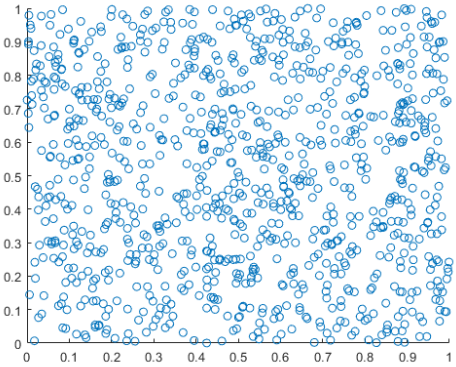
The ranges of variation are not equidistant from the nominal value, as it is usual in a GSA study. This is because it is impossible to do a variation in percentage using the range as in Lyra et al. (2019), without obtaining negative kinetic parameters.

3.7.3 Sampling approach

The sampling for the calculation of the sensitivity indices and MCF is done by using the Sobol quasi-random numbers sequence, as recommended in Saltelli et al. (2008). This quasi-random sequences ensure that the samples fill the input space more uniformly, avoiding gaps and clusters, which occur when random number generators are used. This will typically result in a faster convergence and more stable results (Qian and Mahdi, 2020). A simple example is presented in the Figures 3.3a and 3.3b for two inputs uniformly distributed over 0 and 1, where one is sampled using the Sobol quasi-random numbers and the other is sampled by a random number generator, for this example was decided to sample 10000 points.



(a) By the Sobol quasi-random number generator.



(b) By a random number generator.

Figure 3.3: Input space generated by two different number generators for the same number of points.

Chapter 4

SARA-ATRP kinetic model formulation and numerical simulation

In this chapter the kinetic model proposed in Krys et al. (2017) is described with the model assumptions and the formulation of the mass balances. The reaction scheme is described in Section 4.1 followed by the model assumptions in Section 4.2 and the dynamic model formulation in Section 4.3. Finally, the numerical simulation results are presented and discussed in Section 4.4.

4.1 Reaction scheme

The formulation of the kinetic model of a SARA-ATRP (4.1) is based on the reaction scheme described by Krys et al. (2017). In Krys et al. (2017) it was only considered termination by combination. This is in fact a good approximation used by others modelers in similar studies, as reported for example in Monteiro et al. (2007). This approximation is done because it was demonstrated that acrylate radicals predominantly end by combination (Ribelli et al., 2017), although some chains can still terminate by disproportionation. In Bamford et al. (1969) it was determined that for free-radical polymerization of methyl acrylate (MA) 0.13 of the chains terminate by disproportionation. This reaction was also considered in the kinetic model of Lyra et al. (2019) and Zhou and Luo (2014) of a MA polymerization. In this reaction scheme RX represent the alkyl halide initiator, X^{-} the halogen ion and P_0 is a species formed from a radical self-termination. The description of these kinetic rate constants is given in in Table 4.1.

$$\mathrm{Na_2S_2O_4} \overset{\mathrm{K_S}}{\rightleftharpoons} 2\,\mathrm{Na^+} + {\mathrm{S_2O_4}}^{2-}, \tag{4.1a}$$

$${\mathrm{S_2O_4}}^{2-} \overset{\mathrm{K_{dis}}}{\rightleftharpoons} 2\,{\mathrm{SO_2}}^{\bullet}, \tag{4.1b}$$

$$\mathrm{SO_2}^{\bullet-} + \mathrm{RX} \xrightarrow{\mathrm{k_{a0}^{app}}} \mathrm{R}^{\bullet} + \mathrm{X}^{-} + \mathrm{SO_2}\,, \tag{4.1c}$$

$$\mathrm{SO_2}^{\bullet-} + \mathrm{P}_n\mathrm{X} \xrightarrow{\mathrm{k_{a0p}^{app}}} \mathrm{P}_n{}^{\bullet} + \mathrm{X}^{-} + \mathrm{SO_2}\,, \tag{4.1d}$$

$$\mathrm{SO_2}^{\bullet-} + \mathrm{Cu^{II}X/L} \xrightarrow{\mathrm{k_{red,CuL}^{app}}} \mathrm{Cu^{I}/L} + \mathrm{X}^{-} + \mathrm{SO_2}\,, \tag{4.1e}$$

$$\mathrm{SO_2}^{\bullet-} + \mathrm{Cu^{II}/L} \xrightarrow{\mathrm{k_{red,CuL}^{app}}} \mathrm{Cu^{I}/L} + \mathrm{SO_2}\,, \tag{4.1f}$$

$$\mathrm{RX} + \mathrm{Cu^{I}/L} \underset{\mathrm{k_{d1}}}{\overset{\mathrm{k_{a1}}}{\rightleftharpoons}} \mathrm{R}^{\bullet} + \mathrm{Cu^{II}X/L}\,, \tag{4.1g}$$

$$\mathrm{P}_n\mathrm{X} + \mathrm{Cu^{I}/L} \underset{\mathrm{k_{d1p}}}{\overset{\mathrm{k_{a1p}}}{\rightleftharpoons}} \mathrm{P}_n{}^{\bullet} + \mathrm{Cu^{II}X/L}\,, \tag{4.1h}$$

$$\mathrm{Cu^{I}/L} + \mathrm{X}^{-} \underset{\mathrm{k_{diss2}}}{\overset{\mathrm{k_{ass1}}}{\rightleftharpoons}} \mathrm{Cu^{I}X/L}\,, \tag{4.1i}$$

$$\mathrm{Cu^{II}/L} + \mathrm{X}^{-} \underset{\mathrm{k_{diss1}}}{\overset{\mathrm{k_{ass2}}}{\rightleftharpoons}} \mathrm{Cu^{II}X/L}\,, \tag{4.1j}$$

$$\mathrm{R}^{\bullet} + \mathrm{M} \xrightarrow{\mathrm{k_{add}}} \mathrm{P_1}{}^{\bullet}\,, \tag{4.1k}$$

$$\mathrm{P}_n{}^{\bullet} + \mathrm{M} \xrightarrow{\mathrm{k_p}} \mathrm{P}_{n+1}{}^{\bullet}\,, \tag{4.1l}$$

$$\mathrm{R}^{\bullet} + \mathrm{R}^{\bullet} \xrightarrow{\mathrm{k_{t0}}} \mathrm{P_0}\,, \tag{4.1m}$$

$$\mathrm{R}^{\bullet} + \mathrm{P}_n{}^{\bullet} \xrightarrow{\mathrm{k_{tR}}} \mathrm{D}_n\,, \tag{4.1n}$$

$$\mathrm{P}_m{}^{\bullet} + \mathrm{P}_n{}^{\bullet} \xrightarrow{\mathrm{k_{tc}}} \mathrm{D}_{n+m}\,, \tag{4.1o}$$

$$\mathrm{P}_m{}^{\bullet} + \mathrm{P}_n{}^{\bullet} \xrightarrow{\mathrm{k_{td}}} \mathrm{D}_n + \mathrm{D}_m\,, \tag{4.1p}$$

Table 4.1: Meaning of the rate constants in 4.1.

Nomenclature	description
K_s	dissolution of inorganic salt equilibrium constant
K_dis	dissociation of inorganic salt equilibrium constant
$k_\mathrm{a0}^\mathrm{app}$	apparent rate of initiator activation by $\mathrm{SO_2}^{\bullet}$
$k_\mathrm{a0p}^\mathrm{app}$	apparent rate of dormant chain activation by $\mathrm{SO_2}^{\bullet}$
k_a1	rate constant of initiator activation
k_d1	rate constant of radicals deactivation
k_ass1	rate constant of association of the $\mathrm{Cu^{I}/L}$ halidophilicity equilibrium
k_diss1	rate constant of dissociation of the $\mathrm{Cu^{I}/L}$ halidophilicity equilibrium
k_ass2	rate constant of association of the $\mathrm{Cu^{II}X/L}$ halidophilicity equilibrium
k_diss2	rate constant of dissociation of the $\mathrm{Cu^{II}X/L}$ halidophilicity equilibrium
k_t0	rate constant of termination between two radicals
k_tR	rate constant of termination between a radical and a propagating chain

The $\mathrm{SO_2}^{\bullet}$ is involved in the supplementary activating reaction of the alkyl halide initiator and dormant chains (4.1c, 4.1d) and is a reduction agent of $\mathrm{Cu^{II}X/L}$ and $\mathrm{Cu^{II}/L}$ (Equations 4.1e

and 4.1f). This compound is formed from $Na_2S_2O_4$ by two successive reactions (Equations 4.1a and 4.1b). The rate of the first reaction (Equation 4.1a) is very dependent on the reaction medium, due to the solubility of $Na_2S_2O_4$ changing with the polarity of the solvent. When polar solvents are used (e.g., aqueous solvents), where $Na_2S_2O_4$ is totally soluble, there will be a large concentration of $SO_2{}^{\bullet-}$ in the beginning of the reaction. On the other hand, if non-polar solvents (e.g., organic solvents) are used, the concentration of $SO_2{}^{\bullet-}$ will be constant and low throughout the polymerization (Abreu et al., 2017; Krys et al., 2017; Góis et al., 2014). In this reaction scheme, the direct initiation of growing chains by $SO_2{}^{\bullet}$ is neglected, as it has been proven that this reaction is slow compared to the remaining ones (Krys et al., 2017).

The ATRP equilibrium between (macro)alkyl halides and (macro)radicals is represented by (4.1g) and (4.1h). Equations (4.1i) and (4.1j) represent the halidophicility equilibrium of $Cu^{II}X/L$ and Cu^{I}/L complexes (Krys et al., 2017). The reactions of first monomer addition, monomer addition to propagating chains, termination between two radicals, termination between a radical and a propagating chain and termination between two propagating chains by combination and disproportionation are represented by Equations 4.1k, 4.1l, 4.1m, 4.1n, 4.1o and 4.1p, respectively.

4.2 Model assumptions

For the development of the kinetic model it is first necessary to state the main assumptions in its formulation. The model prediction will be compared to the experimental data obtained in Krys et al. (2017). The data are relative to the polymerization of MA at 30ºC, using a mixture of 90 % of ethanol ($EtOH$) and 10 % of water as solvent and Ethyl α-bromoisobutyrate as the initiator. Taking into account similar works, the main assumptions are:

- isothermical polymerization reaction (Krys et al., 2017);
- the diffusion limitations are not taken into account (Huang et al., 2013; Lyra et al., 2019; Massicotte, 2015);
- the volume variation is not taken into account (Lyra et al., 2019);
- rate constants are independent of polymer chain length (Huang et al., 2013; Massicotte, 2015; Vieira et al., 2015);
- the cooper salt is completely complexed with the ligand at the beggining of the reaction (Lyra et al., 2019);
- the kinetic rate constants and the experimental data for the model validation, are taken from a polymerization performed with a organic solvent. Taking into account the discussion made in the previous section, the concentration of $SO_2{}^{\bullet-}$ is assumed to be constant throughout the reaction. Therefore, apparent rates are used in (4.1c), (4.1d), (4.1e) and (4.1f). Thus, the reactions (4.1a) and (4.1b) can be neglected (Krys et al., 2017).

4.3 Dynamic model formulation

In this section the mathematical expressions of the dynamic model will be presented.

4.3.1 The moments differential equations

The 0th, 1st and 2nd moments differential equations were obtained for the P_n, $\mathrm{P}_n\mathrm{X}$ and D_n species. The development of these equations can be found in the Appendix B. The moment differential equations for the P_n are:

$$\begin{aligned}\frac{\mathrm{d}\lambda_0}{\mathrm{d}t} &= k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_0 + k_{\mathrm{a1p}}\,[\mathrm{Cu^{I}/L}]\mu_0 - k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\lambda_0 \\ &\quad + k_{\mathrm{add}}\,[M][\mathrm{R}] - k_{\mathrm{tR}}\,[\mathrm{R^{\bullet}}]\,\lambda_0 - (k_{\mathrm{tc}} + k_{\mathrm{td}})\lambda_0\,\lambda_0\,,\end{aligned} \tag{4.2}$$

$$\begin{aligned}\frac{\mathrm{d}\lambda_1}{\mathrm{d}t} &= k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_1 + k_{\mathrm{a1p}}\,[\mathrm{Cu^{I}/L}]\mu_1 - k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\lambda_1 + k_{\mathrm{p}}\,[\mathrm{M}]\lambda_0 \\ &\quad + k_{\mathrm{add}}\,[M][\mathrm{R^{\bullet}}] - k_{\mathrm{tR}}[\mathrm{R}]\,\lambda_1 - (k_{\mathrm{tc}} + k_{\mathrm{td}})\lambda_1\,\lambda_0\,,\end{aligned} \tag{4.3}$$

$$\begin{aligned}\frac{\mathrm{d}\lambda_2}{\mathrm{d}t} &= k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_2 + k_{\mathrm{a1p}}\,[\mathrm{Cu^{I}/L}]\mu_2 - k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\,\lambda_2 + k_{\mathrm{p}}\,[\mathrm{M}]\,(\lambda_0 + 2\,\lambda_1) \\ &\quad + k_{\mathrm{add}}\,[M][\mathrm{R}] - k_{\mathrm{tR}}\,[\mathrm{R^{\bullet}}]\,\lambda_2 - (k_{\mathrm{tc}} + k_{\mathrm{td}})\lambda_2\,\lambda_0\end{aligned} \tag{4.4}$$

It follows that the moment differential equations for the $\mathrm{P}_n\mathrm{X}$ are :

$$\frac{\mathrm{d}\mu_0}{\mathrm{d}t} = -k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_0 - k_{\mathrm{a1p}}\,[\mathrm{Cu^{I}/L}]\mu_0 + k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\lambda_0 \tag{4.5}$$

$$\frac{\mathrm{d}\mu_1}{\mathrm{d}t} = -k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_1 - k_{\mathrm{a1p}}\,[\mathrm{Cu^{I}/L}]\mu_1 + k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\lambda_1 \tag{4.6}$$

$$\frac{\mathrm{d}\mu_2}{\mathrm{d}t} = -k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_2 - k_{\mathrm{a1p}}\,[\mathrm{Cu^{I}/L}]\mu_2 + k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\,\lambda_2 \tag{4.7}$$

Finally, the moment differential equations for the D_n are given by:

$$\frac{\mathrm{d}\delta_0}{\mathrm{d}t} = k_{\mathrm{tR}}\,[\mathrm{R^{\bullet}}]\,\lambda_0 + k_{\mathrm{td}}\,\lambda_0\,\lambda_0 + \frac{k_{\mathrm{tc}}}{2}\,\lambda_0\,\lambda_0 \tag{4.8}$$

$$\frac{\mathrm{d}\delta_1}{\mathrm{d}t} = k_{\mathrm{tR}}\,[\mathrm{R^{\bullet}}]\,\lambda_1 + (\,k_{\mathrm{td}} + k_{\mathrm{tc}}\,)\,\lambda_1\,\lambda_0 \tag{4.9}$$

$$\frac{\mathrm{d}\delta_2}{\mathrm{d}t} = k_{\mathrm{tR}}\,[\mathrm{R^{\bullet}}]\,\lambda_2 + (\,k_{\mathrm{td}} + k_{\mathrm{tc}}\,)\,\lambda_2\,\lambda_0 + k_{\mathrm{tc}}\,\lambda_1\lambda_1 \tag{4.10}$$

4.3.2 Non-polymeric species molar balances

For the non-polymeric species (radicals, initiator, terminated radicals, activator, deactivator, $\mathrm{Cu^{II}/L}$, $\mathrm{Cu^{I}X/L}$, monomer, sulfur dioxide and the halogen ion) a molar balance for each one

is considered, resulting in the additional set of ordinary differential equations (4.11).

Radicals, $R^\bullet$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{R}]}{\mathrm{d}t} &= k_{\mathrm{a0}}^{\mathrm{app}}[\mathrm{RX}] + k_{a1}[\mathrm{RX}][\mathrm{Cu^I/L}] \\ &\quad - K_{\mathrm{d1}}[\mathrm{R}^\bullet][\mathrm{Cu^{II}X/L}] - k_{\mathrm{add}}[\mathrm{M}][\mathrm{R}^\bullet] \\ &\quad - 2\,k_{t0}[\mathrm{R}^\bullet]^2 - k_{\mathrm{tR}}[\mathrm{R}^\bullet]\lambda_0\end{aligned} \tag{4.11a}$$

Initiator, RX

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{RX}]}{\mathrm{d}t} &= -k_{\mathrm{a0}}^{\mathrm{app}}[\mathrm{RX}] - k_{a1}[\mathrm{RX}][\mathrm{Cu^I/L}] \\ &\quad + k_{\mathrm{d1}}[\mathrm{R}^\bullet][\mathrm{Cu^{II}X/L}]\end{aligned} \tag{4.11b}$$

Terminated rad., P_0

$$\frac{\mathrm{d}[\mathrm{P_0}]}{\mathrm{d}t} = k_{\mathrm{t0}}[\mathrm{R}^\bullet]^2 \tag{4.11c}$$

Activator, Cu^I/L

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{Cu^I/L}]}{\mathrm{d}t} &= k_{\mathrm{red,X-CuL}}^{\mathrm{app}}[\mathrm{Cu^{II}X/L}] + k_{\mathrm{red,CuL}}^{\mathrm{app}}[\mathrm{Cu^{II}/L}] \\ &\quad - k_{a1}[\mathrm{RX}][\mathrm{Cu^I/L}] + k_{\mathrm{d1}}[\mathrm{R}^\bullet][\mathrm{Cu^{II}X/L}] \\ &\quad - k_{\mathrm{a1p}}[\mathrm{Cu^I/L}]\mu_0 + k_{\mathrm{d1p}}[\mathrm{Cu^{II}X/L}]\lambda_0 \\ &\quad - k_{\mathrm{ass1}}[\mathrm{Cu^I/L}][\mathrm{X^-}] + k_{\mathrm{diss1}}[\mathrm{Cu^IX/L}]\end{aligned} \tag{4.11d}$$

Deactivator,Cu^IX/L

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{Cu^{II}X/L}]}{\mathrm{d}t} &= -k_{\mathrm{red,X-CuL}}^{\mathrm{app}}[\mathrm{Cu^{II}X/L}] + k_{a1}[\mathrm{RX}][\mathrm{Cu^I/L}] \\ &\quad - k_{\mathrm{d1}}[\mathrm{R}^\bullet][\mathrm{Cu^{II}X/L}] + k_{\mathrm{a1p}}[\mathrm{Cu^I/L}]\mu_0 \\ &\quad - k_{\mathrm{d1p}}[\mathrm{Cu^{II}X/L}]\lambda_0 + k_{\mathrm{ass2}}[\mathrm{Cu^{II}/L}][\mathrm{X^-}] \\ &\quad - k_{\mathrm{diss2}}[\mathrm{Cu^{II}X/L}]\end{aligned} \tag{4.11e}$$

Cu^{II}/L

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{Cu^{II}/L}]}{\mathrm{d}t} &= -k_{\mathrm{red,CuL}}^{\mathrm{app}}[\mathrm{Cu^{II}/L}] - k_{ass2}[\mathrm{Cu^{II}/L}][\mathrm{X^-}] \\ &\quad + k_{\mathrm{diss2}}[\mathrm{Cu^{II}X/L}]\end{aligned} \tag{4.11f}$$

Cu^IX/L

$$\frac{\mathrm{d}[\mathrm{Cu^IX/L}]}{\mathrm{d}t} = k_{\mathrm{ass1}}[\mathrm{Cu^I/L}][\mathrm{X^-}] - k_{\mathrm{diss1}}[\mathrm{Cu^IX/L}] \tag{4.11g}$$

Monomer, M

$$\frac{\mathrm{d}[\mathrm{M}]}{\mathrm{d}t} = -k_{\mathrm{add}}[\mathrm{R}][\mathrm{M}] - k_{\mathrm{p}}[\mathrm{M}]\lambda_0 \tag{4.11h}$$

Sulfur dioxide, SO_2

$$\frac{\mathrm{d}[\mathrm{SO_2}]}{\mathrm{d}t} = k_{\mathrm{a0}}^{\mathrm{app}}[\mathrm{RX}] + k_{\mathrm{a0p}}^{\mathrm{app}}\mu_0 + k_{\mathrm{red,X-CuL}}^{\mathrm{app}}[\mathrm{Cu^{II}X/L}] + k_{\mathrm{red,CuL}}^{\mathrm{app}}[\mathrm{Cu^{II}/L}] \tag{4.11i}$$

Halogen ion, X^-

$$\frac{\mathrm{d}[\mathrm{X^-}]}{\mathrm{d}t} = k_{\mathrm{a0}}^{\mathrm{app}}[\mathrm{RX}] + k_{\mathrm{a0p}}^{\mathrm{app}}\mu_0 + k_{\mathrm{red,X-CuL}}^{\mathrm{app}}[\mathrm{Cu^{II}X/L}] - k_{\mathrm{ass1}}[\mathrm{Cu^{I}/L}][X^-] + k_{\mathrm{diss1}}[\mathrm{Cu^{I}X/L}] - k_{\mathrm{ass2}}[\mathrm{Cu^{II}/L}][\mathrm{X^-}] + k_{\mathrm{diss2}}[\mathrm{Cu^{II}X/L}] \tag{4.11j}$$

Therefore, the dynamic model of this SARA-ATRP system is formed by of 19 ordinary differential equations (ODEs).

4.4 Numerical simulation

For the kinetic model simulation, the same kinetic parameters and initial conditions as in (Krys et al., 2017) were used. These values are given in Tables 4.2 and 4.3, respectively. Also, in the simulations performed in Krys and Matyjaszewski (2017) it was considered that the initial concentration of X^- is the same as the one of the deactivator ($[\mathrm{CuBr_2}]_0$). The same consideration is used in this simulation.

Table 4.2: Kinetic parameters reported in Krys et al. (2017).

Parameter	Value
$k_{\mathrm{a0}}^{\mathrm{app}}$	1.20×10^{-5} $\mathrm{s^{-1}}$
$k_{\mathrm{a0p}}^{\mathrm{app}}$	9.70×10^{-6} $\mathrm{s^{-1}}$
$k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$	6.2×10^{-5} $\mathrm{s^{-1}}$
$k_{\mathrm{red,CuL}}^{\mathrm{app}}$	2.3×10^{-4} $\mathrm{s^{-1}}$
k_{a1}	1.30×10^{5} $\mathrm{M^{-1}}$ s-1
k_{a1p}	6.30×10^{3} $\mathrm{M^{-1}}$ $\mathrm{s^{-1}}$
k_{d1}	3.70×10^{7} $\mathrm{M^{-1}}$ $\mathrm{s^{-1}}$
k_{d1p}	3.70×10^{7} $\mathrm{M^{-1}}$ $\mathrm{s^{-1}}$
k_{p}	1.47×10^{4} $\mathrm{M^{-1}}$ $\mathrm{s^{-1}}$
k_{add}	1.38×10^{3} $\mathrm{M^{-1}}$ $\mathrm{s^{-1}}$
k_{t}	2.45×10^{8} $\mathrm{M^{-1}}$ $\mathrm{s^{-1}}$
k_{tR}	1.00×10^{8} $\mathrm{M^{-1}}$ $\mathrm{s^{-1}}$
k_{t0}	8.00×10^{8} $\mathrm{M^{-1}}$ $\mathrm{s^{-1}}$
k_{ass1}	1.00×10^{9} $\mathrm{M^{-1}}$ $\mathrm{s^{-1}}$
k_{diss1}	3.10×10^{7} $\mathrm{s^{-1}}$
k_{ass2}	1.00×10^{9} $\mathrm{M^{-1}}$ $\mathrm{s^{-1}}$
k_{diss2}	6.10×10^{4} $\mathrm{s^{-1}}$

Table 4.3: Initial conditions.

Initial condition	Value
Target DP	222.000
$[M]_0$	7.400 M
$[RX]_0$	0.033 M
$[\mathrm{CuBr_2}]_0$	0.003 M

The constant k_t is defined by (4.12) (Young and Lovell, 2011). The rationale behind this equation is that some chains will terminate by combination and others by disproportionation. Thus, to calculate the constants k_tc and k_td, it is necessary to multiply k_t by the corresponding fraction (see (4.13)).

$$k_\mathrm{t} = k_\mathrm{td} + k_\mathrm{tc} \tag{4.12}$$

As mentioned before, the disproportionation reaction is considered in this work, using the fraction of combination of 0.87 calculated in Bamford et al. (1969), for a free-radical polymerization of MA.

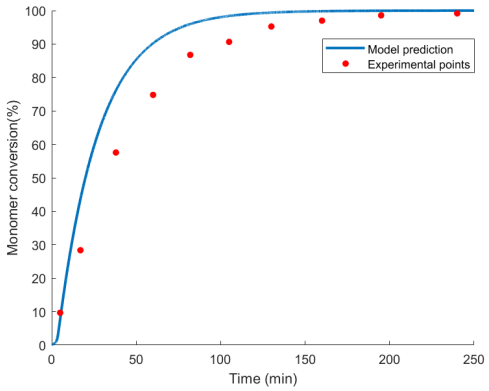
$$\begin{cases} k_\mathrm{tc} = 0.87 k_\mathrm{t} \\ k_\mathrm{td} = (1 - 0.87) k_\mathrm{t} \end{cases} \tag{4.13}$$

For instance, thesimulations made by Lyra et al. (2019) and Zhou and Luo (2014) for the MA ATRP, a 0.90 combination fraction was used, which is very close to the one calculated by Bamford et al. (1969).

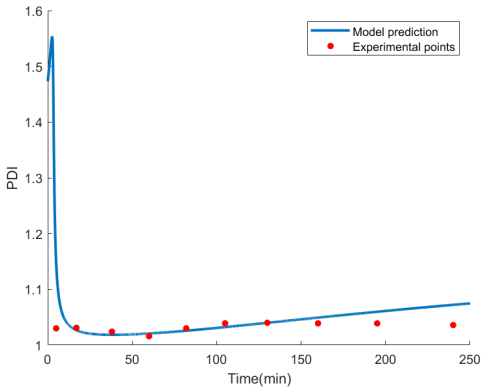
The kinetic model was implemented in MATLAB®, and was solved using the ode15s solver. This solver is specific for stiff models. A stiff system is when one of the solutions of the ODE varies slowly while another varies rapidly. This implies that the numerical method use a very small integration step to obtain good results (Moler, 2003). This kinetic model is predicted to be stiff due to the a priori knowledge of the reaction. For example, it is expected that the concentration of radicals from the initiator and the initiator will decrease sharply at the beginning of the reaction, while the monomer concentration will gradually decrease over time.

In order to compare the model prediction with the experimental data obtained in Krys et al. (2017), the profile of the monomer conversion, PDI and M_n along the reaction time were obtained. These profiles are presented in the Figures 4.1a, 4.1b and 4.1c, respectively. A simulation time of 250 min was considered, because the last experimental point is available at 240 min. The experimental points are given in Appendix C in Table C.3.

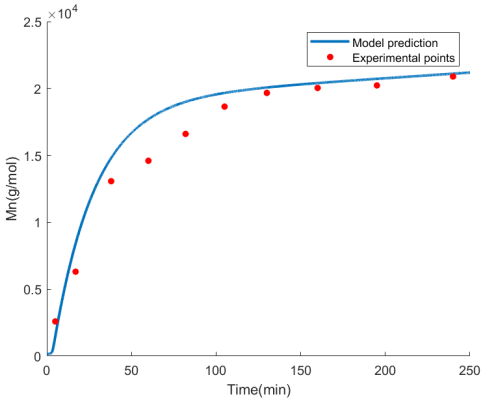
From observing these Figures, it is possible to note that there is not a perfect match between the model prediction and the experimental data. Through a visual comparison between the 3 plots it appears that this error is more pronounced for the monomer conversion and M_n profiles. For the PDI, the model is able to predict the value well with the exception of the 1st and last two points. This difference may be due to not being considered limitations by diffusion, as these are more



(a) Monomer conversion vs time.



(b) PDI vs time.



(c) M_n vs time.

Figure 4.1: Profiles obtained along time.

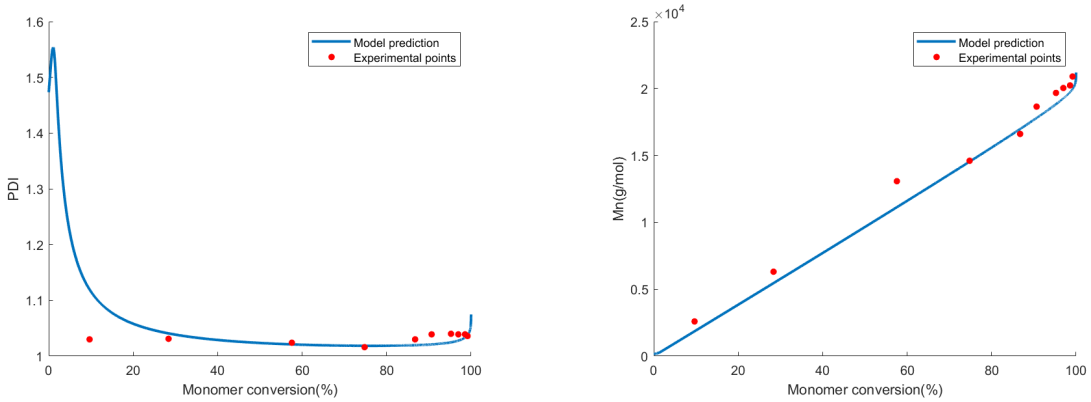
pronounced at the end of the reaction, when the medium becomes more viscous. Nevertheless, it is also verified that the 4th and 5th point are those with the largest deviation from the model prediction in the M_n profile. Also, it is possible to visualize that this two points show some discrepancy in relation to the evolution of the M_n along time. This could indicate that these points could be affected by experimental errors.

The normalized root mean square error (NRMSE) was calculated using Equation 3.25. This type of error was chosen instead of the more widely used root mean square error (RMSE) because the values of the monomer conversion, the M_n and the PDI are on different scales.

Table 4.4: Error between model prediction and experimental data.

	Conversion	PDI	M_n
NRMSE	0.1351	0.0386	0.1051

The results of the NRMSE values confirm the conclusions drawn from the visual analysis of the profiles in Figure 4.1. The PDI presents the lowest error and by a significant margin compared to the other outputs. Also, it follows that the monomer conversion shows the largest error. Figures 4.2a and 4.2b show the profiles of PDI and M_n in function of the monomer conversion.



(a) PDI vs monomer conversion.

(b) M_n vs monomer conversion.

Figure 4.2: Profiles obtained along the monomer conversion.

It can be concluded that the model provides relatively good predictions of the PDI and the M_n in function of the monomer conversion, with the exception of the first experimental point in the PDI profile. This is in accordance to the simulations results reported in Krys et al. (2017). It is worth to note that the main objective of the model is not to predict well in this zone, since it corresponds to the initial stage of reaction where the monomer conversion is very low (5%). Also to be noted, that the M_n has a linear evolution, as expected in reversible deactivation radical polymerization reactions (Matyjaszewski and Davis, 2003).

In order to improve the model ability to predict the conversion and M_n, there are, two options. The first option is to increase the complexity of the model by considering fewer assumptions. Another possibility is an optimization problem process, where the values of the kinetic parameters are estimated with the aim of reducing the error between the model prediction and the experimental data. In this work was decided to pursue this last option.

This can be in fact a good opportunity to demonstrate the application of SA, as this model is constituted by 17 parameters, and most certainly not all of these will have the same importance to the monomer conversion and the M_n. Trough the SA results, it will be possible to only choose the important parameters to be estimated in the optimization, similarly to the work of Salas et al. (2019), reducing the complexity of the kinetic estimation problem and the risks of overfitting the model to this set of data.

Chapter 5

Global sensitivity analysis results

In this chapter, the sensitivity analysis results obtained with the two GSA methods, the Monte Carlo filtering (Section 5.2) and the Sobol method (Section 5.3), are presented. Also, in Section 5.4 a comparison between these two results is presented. One emphasizes that in order to reduce the computational time of this sensitivity analysis, the MATLAB® parallel computing toolbox was used in both methods.

5.1 Simplifications

5.1.1 Parameters studied

For this sensitivity analysis it was decided to study not all the model kinetic parameters. The main reason behind this decision has to do with the total number of model runs necessary to perform the Sobol method. As said before, the number of runs will depend on the number of parameters, being $N(K + 2)$ when using the method in Saltelli et al. (2008) or $N(2K + 2)$ if the method of Wu et al. (2012) is used. Obviously, a higher number of model parameters, will increase the total number of model runs and consequently the computational time of the analysis.Therefore the parameters that were not included are k_{ass1}, k_{diss1}, k_{ass2} and k_{diss}. These parameters correspond to the halidophilicity equilibrium that is not usually accounted in most of the ATRP kinetic models found in the literature. Nevertheless, the model was tested without these reactions and no significant differences were obtained. This proves that choosing these kinetics constants it is the best option in order to reduce the number of model simulations and, consequently, the computational time of the analysis.

5.1.2 Notation

As said in Chapter 3, three different input spaces will be studied. This is because the true range of variation of the parameters under study is unknown, and different input spaces can affect the

results of the GSA. The three input space intervals will be refereed simply as the 1st, 2nd and 3rd interval accordingly to the order they are given in the Table 3.4.

5.2 Monte Carlo filtering

The Monte Carlo filtering method was implemented in MATLAB® following an example described in Saltelli et al. (2008). The MATLAB® Statistical and Machine learning toolbox was used to apply the two sample K-S statistical test which is available in this toolbox.

Following the explanation given in the Section 3.6 of Chapter 3, the behavioural and non-behavioural output will be defined as follow:

- $B_x \rightarrow$ Have a NRMSE less than 0.1351;
- $\overline{B}_x \rightarrow$ Have a NRMSE higher than 0.1351;
- $B_{PDI} \rightarrow$ Have a NRMSE less than 0.0386;
- $\overline{B}_{PDI} \rightarrow$ Have a NRMSE higher than 0.0386;
- $B_{Mn} \rightarrow$ Have a NRMSE less than 0.1051;
- $\overline{B}_{Mn} \rightarrow$ Have a NRMSE higher than 0.1051.

5.2.1 Choosing the number of samples (*N*)

To apply the MCF method it is first necessary to choose the number of samples to be generated (N). In the case of MCF, N, corresponds to the number of model simulations. Recall that the two sample Kolmogorov-Smirnov test (K-S test) gives two different outputs, the $d_{n,\overline{n}}$ and the p-value. In order to choose the correct N, it was decided to check from which value of N the $d_{n,\overline{n}}$ remained constant or when its variation was minimum. This approach follows a similar study reported in Wu et al. (2017).

The output $d_{n,\overline{n}}$ was chosen instead of the p-value because this last tends to decrease as the number of samples increase, except for the model parameters which have no influence on the output. This is mainly due to the fact that if the size of the two distributions generated by the model simulations for each parameter, $f_n(X_i|B)$ and $f_{\overline{n}}(X_i|\overline{B})$, are larger, there will be a greater probability that the two sample K-S test will find a difference between them. Thus, decreasing the p-value, which tends to 0 as the number of N increases (Thiese et al., 2016). This problem is discussed in Lin et al. (2013), where is stated that small p-values can be an artifact of a large sample size and the interpretation of this value has to be critical and no conclusions should be drawn from this number alone. This is commonly addressed as the p-value problem (Lin et al., 2013). In this thesis, the two results obtained from the K-S test, the $d_{n,\overline{n}}$ and the p-value, will

be analysed. The p-value will be used to classify the parameters, following the classification in Saltelli et al. (2004), and the $d_{n,\overline{n}}$ for ranking the parameters, working as a sensitivity measure.

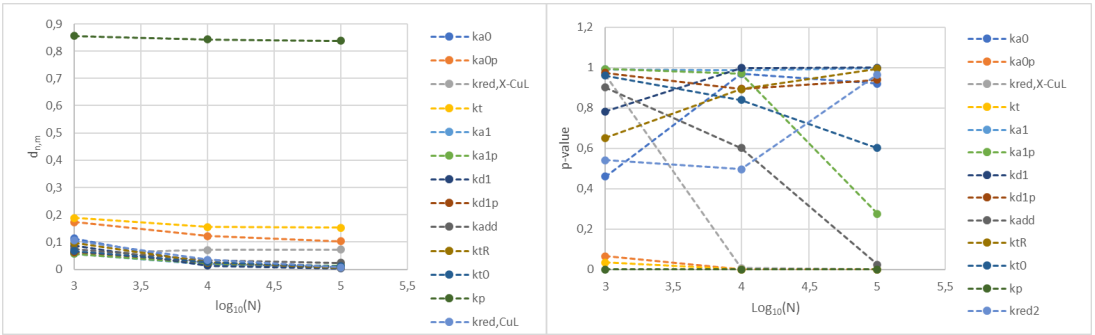
As mentioned in Chapter 3, due to the difficulty in defining the correct sampling interval for each parameter 3 different intervals will be studied, represented in Table 3.4. To choose the number of N to apply the MCF method, it was decided only to study the variation of $d_{n,\overline{n}}$ along N, using the 1st interval inputs range. This is that an initial analysis demonstrated that using the inputs space of the 1st interval will require a larger N for $d_{n,\overline{n}}$ to converge. In order to maintain consistency, in order to make a uniform comparison between this intervals, the number of N concluded in this study will be the one used for the application of the MCF method when using the 2nd and 3rd interval inputs range.

It was observed that by sampling the parameters within the range of the 1st interval and then simulate the model with these values, the number of outputs obtained in the B was much smaller than the outputs in $\overline{B}$. Table 5.1 shows the results for 1000 model evaluations where there is a big difference between the outputs with the expected behaviour and those without it, especially for M_n output. This is can be due to the fact that very extreme values of the parameters are explored, which inevitably cause the model to have an off-target response.

Table 5.1: Number of outputs off-target and in-target with 1000 model evaluations.

	Conversion	PDI	Mn
B	59	55	8
$\overline{B}$	941	945	992

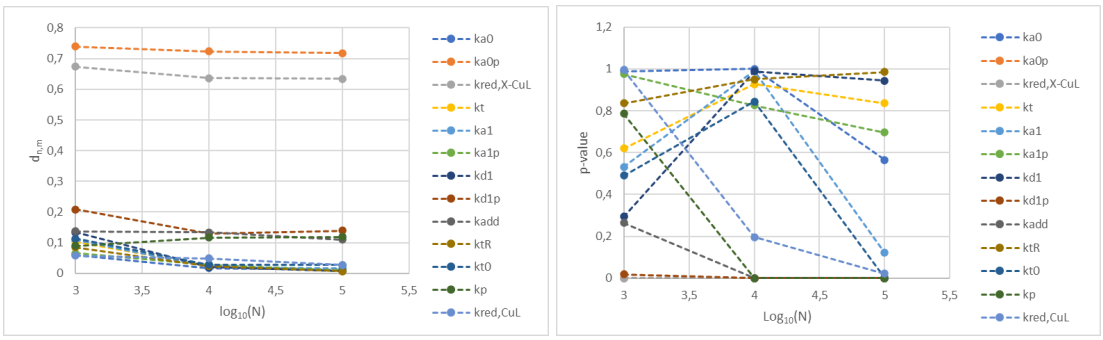
In order decide the number of N, the MCF was aplied for $N = 1000$, $N = 1000$ and $N = 10000$. The results of $d_{n,\overline{n}}$ and p-value in function of the $log_{10}(N)$ are presented in the Figures 5.1a to 5.3b, for the three outputs. The p-value profile is also studied to illustrate the explanation above.



(a) $d_{n,\overline{n}}$. **(b)** p-value.

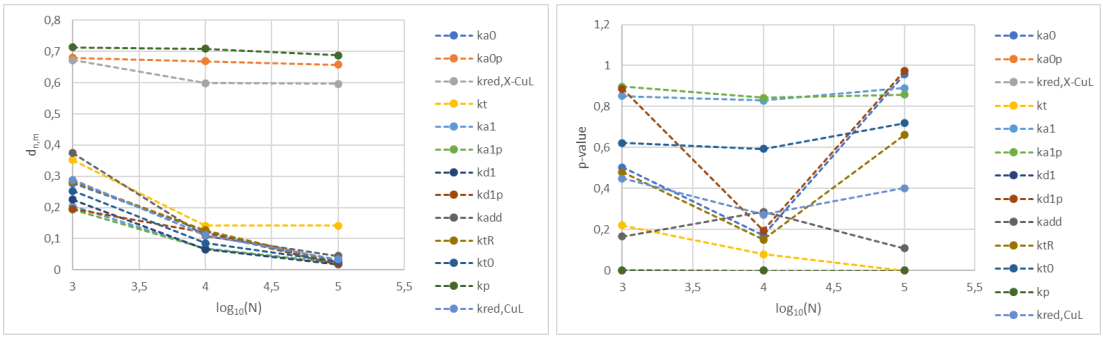
Figure 5.1: Two sample Kolmogorov-Smirnov test statistics in function of $log_{10}(N)$ for the monomer conversion output.

From these figures it is possible to observe that in most of the cases the $d_{n,m}$ values remain constant or presents small variations between $N = 10000$ and $N = 100000$. The only clear



(a) $d_{n,\overline{n}}$. **(b)** p-value.

Figure 5.2: Two sample Kolmogorov-Smirnov test statistics in function of $log_{10}(N)$ for the PDI output.



(a) $d_{n,\overline{n}}$. **(b)** p-value.

Figure 5.3: Two sample Kolmogorov-Smirnov test statistics in function of $log_{10}(N)$ for the M_n output.

exception, is in Figure 5.3a, where it is possible to verify that some inputs, with a small $d_{n,\overline{n}}$ still varied considerable between $N = 10000$ and $N = 100000$. On the other hand, in most of the cases the p-value does not converge within the number of N used. It can be also verified, that this value decreases as the number of N increases, possibly changing the classification of the parameter, following the classification in Table 3.2. Nevertheless, other pattern of variation can be observed, such as in the case of k_{d1p}, in Figure 5.3b. In this case, the p-value decreases between $N = 1000$ and $N = 10000$ but increases between $N = 10000$ and $N = 100000$. This clearly shows that the value of N should not be chosen according to the p-value. Based on these data, it was decided to use a $N = 10000$. The reason for this decision is because for most parameters the $d_{n,\overline{n}}$ shows a minimum variation between $N = 10000$ and $N = 100000$. The only exception is in Figure 5.3b, where for some parameters the $d_{n,\overline{n}}$ value still varies between $N = 10000$ and $N = 100000$. The parameters for which this occurs have a small $d_{n,\overline{n}}$. Therefore, they are considered insignificant parameters. Thus, it is not mandatory to obtain the most correct $d_{n,\overline{n}}$ value for these parameters. Also, using a $N = 100000$ could increase the number of inputs considered important/critical due to the p-value problem. Remark that if number of parameters are considered critical/important, no meaningful conclusions can be drawn from this sensitivity analysis. This problem can be more problematic if it is considered the 2nd and 3rd interval parameter range, where it is expected a faster convergence of the $d_{n,\overline{n}}$ value.

5.2.2 Monomer conversion output results

The results of the two sample K-S test for the monomer conversion output are presented in the Table 5.2, 5.3 and 5.4 for the 1st, 2nd and 3rd interval, respectively.

Table 5.2: 1st interval monomer conversion MCF result.

Parameter	$d_{n,\overline{n}}$	Classification
$k_{\mathrm{a0}}^{\mathrm{app}}$	0.0210	Insignificant
$k_{\mathrm{a0p}}^{\mathrm{app}}$	0.1226	Critical
$k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$	0.0713	Critical
$k_{\mathrm{red,CuL}}^{\mathrm{app}}$	0.0354	Insignificant
k_{a1}	0.0190	Insignificant
k_{a1p}	0.0209	Insignificant
k_{d1}	0.0141	Insignificant
k_{d1p}	0.0246	Insignificant
k_{p}	0.8435	Critical
k_{add}	0.0327	Insignificant
k_{t}	0.1558	Critical
k_{tR}	0.0247	Insignificant
k_{t0}	0.0264	Insignificant

Table 5.3: 2nd interval monomer conversion MCF result.

Parameter	$d_{n,\overline{n}}$	Classification
$k_{\mathrm{a0}}^{\mathrm{app}}$	0.0177	Insignificant
$k_{\mathrm{a0p}}^{\mathrm{app}}$	0.1051	Critical
$k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$	0.0678	Critical
$k_{\mathrm{red,CuL}}^{\mathrm{app}}$	0.0157	Insignificant
k_{a1}	0.0162	Insignificant
k_{a1p}	0.0212	Insignificant
k_{d1}	0.0179	Insignificant
k_{d1p}	0.0661	Critical
k_{p}	0.7736	Critical
k_{add}	0.0322	Insignificant
k_{t}	0.1538	Critical
k_{tR}	0.0322	Insignificant
k_{t0}	0.0200	Insignificant

Table 5.4: 3rd interval monomer conversion MCF result.

Parameter	$d_{n,\overline{n}}$	Classification
$k_{\mathrm{a0}}^{\mathrm{app}}$	0.0169	Insignificant
$k_{\mathrm{a0p}}^{\mathrm{app}}$	0.1435	Critical
$k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$	0.0970	Critical
$k_{\mathrm{red,CuL}}^{\mathrm{app}}$	0.0157	Insignificant
k_{a1}	0.0149	Insignificant
k_{a1p}	0.0181	Insignificant
k_{d1}	0.0154	Insignificant
k_{d1p}	0.0126	Insignificant
k_{p}	0.7348	Critical
k_{add}	0.0419	Critical
k_{t}	0.2219	Critical
k_{tR}	0.0192	Insignificant
k_{t0}	0.0192	Insignificant

From these results it is possible to verify that the same parameters are critical, according to the p-value, in all 3 intervals (k_{p}, k_{t}, $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$). The only exception is the 2nd

interval where k_{d1p} is considered critical and in the 3rd interval where k_{add} is also classified as critical. Also, for all the intervals the parameter ranking, based on the $d_{n,m}$, was always the same. Therefore, it follows that k_{p} is the most influential parameter, with the largest $d_{n,m}$, followed by k_t, $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{red,X-CuL}^{app}$. For the second interval, k_{d1p} appears as the 5th most influential parameter. In the 3rd interval k_{add} is also the 5th most influential parameter.

The result of k_{p} being the most influential parameter was expected. It is the propagation rate constant related to the monomer addition to the growing chain reaction, which is in the monomer mass balance equation and in the propagation rate equation for a ATRP system. Although, in a controlled polymerization system it is expected that there will be a low termination of polymeric chains, k_{t} appears as the second most influential parameter. This may be due to the fact that the variation in the value of this constant will have an effect on the concentration of propagating chains, thus changing the rate of propagation (See Equation 2.5), and consequently the monomer conversion.

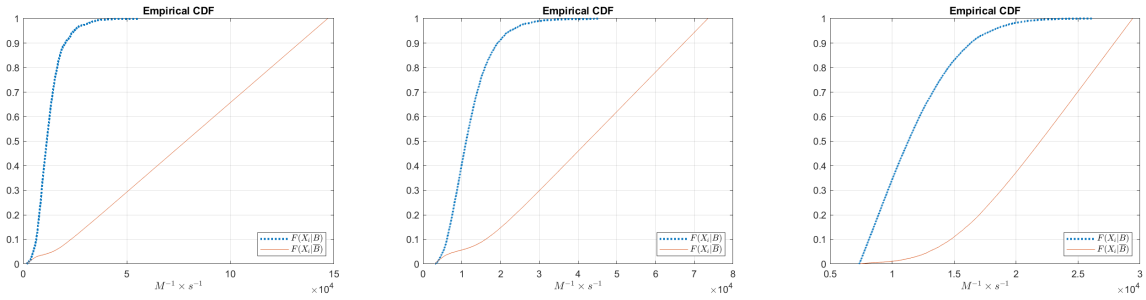
Regarding the specific parameters of the SARA-ATRP system, $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$, from the kinetic model it is possible to verify that these constants will directly affect the concentration of dormant chains, activator and deactivator. These are all elements of Equation 2.5, thus affecting the monomer conversion. The reason for $k_{\mathrm{a0p}}^{\mathrm{app}}$ being more influential than $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ may be due to the fact that $k_{\mathrm{a0p}}^{\mathrm{app}}$ is the constant of supplemental activation of dormant chains, directly producing propagating chains. Also, it does makes sense that k_{d1p} receives the critical classification for the same reason as $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$. Taking into account that the ATRP equilibrium is the main reaction in this polymerization system, (see Chapter 2) it should make sense that k_{d1p} has a higher sensitivity index than the one of $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$. The importance of k_{add} can be due to the fact that this parameter, similarly to the k_{p}, appears directly in the monomer mass balance equation (see Chapter 4), but their relative importance is only acknowledged in the 3rd interval.

In Figures 5.4,5.5 ,5.6, 5.7 and 5.8 are presented the two CDFs, for the three intervals and for the parameters that are classified as critical.

From these figures, it is possible to observe the bigger difference between the two CDFs for the constant k_{p}. This is in agreement with what was concluded previously, by the statistical two sample K-S test, that k_{p} is the most important parameter.

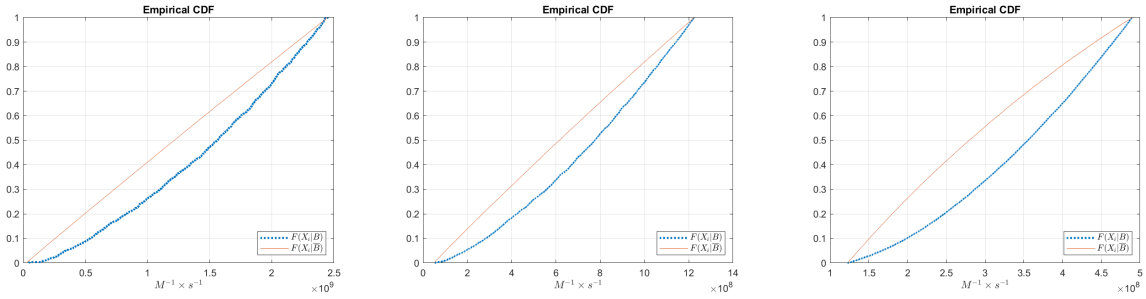
In an individual analysis of the results obtained for each parameter and taking into account the slopes of the CDFs, it is possible to verify there is a higher probability of obtaining a behavioral output by using low values of k_{p}, $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$, and high values of k_{t}. In fact with a certain range of k_{p} values it is impossible to obtain a behavioural response.

Finally, it can be also verified that only in the 2nd interval the CDFs for k_{d1p} present some difference between them. Nevertheless, in this figure the two CDFs are very similar and no meaningful conclusion can be taken on which interval of k_{d1p} values there is more probability



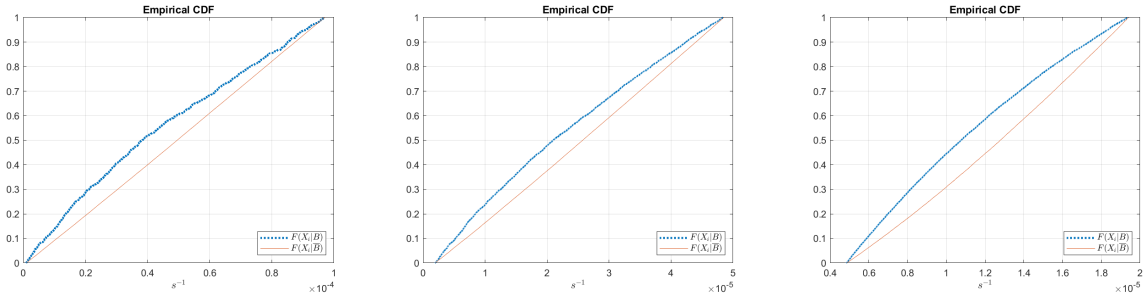
(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.4: CDFs of k_{p} for the monomer conversion.



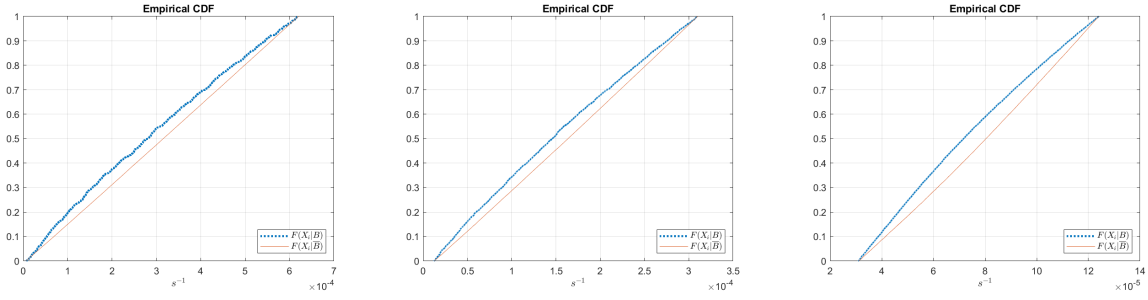
(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.5: CDFs of k_{t} for the monomer conversion.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

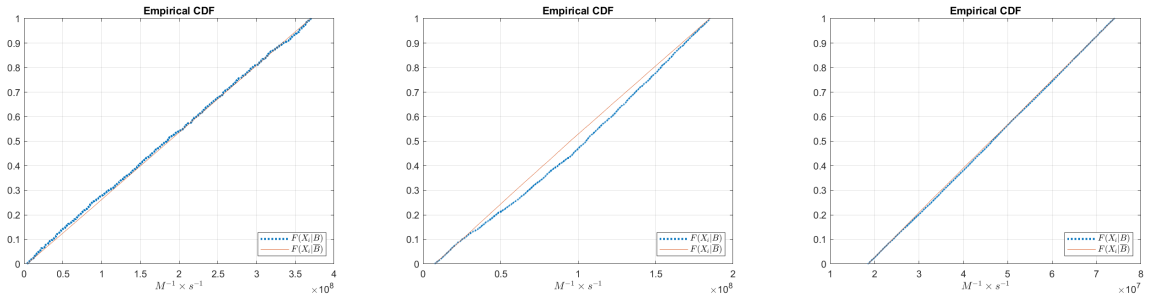
Figure 5.6: CDFs of $k_{\mathrm{a0p}}^{\mathrm{app}}$ for the monomer conversion.



(a) 1st interva. **(b)** 2nd interval. **(c)** 3rd interval.

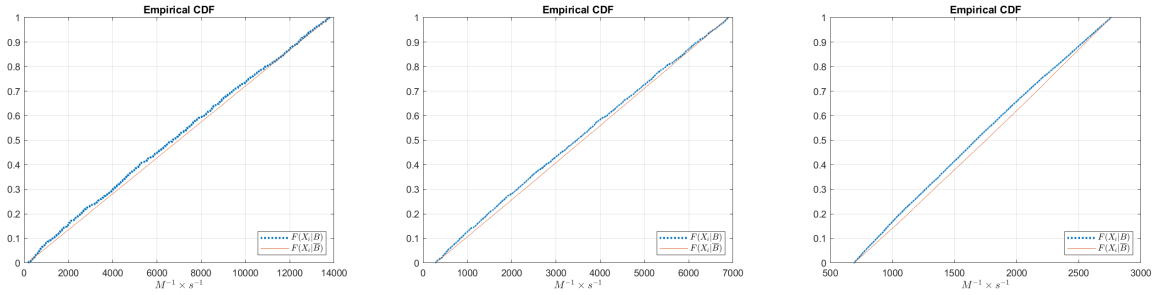
Figure 5.7: CDFs of $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ for the monomer conversion.

of obtaining a behavioural output. From the Figure 5.9c is possible to verify that by using low values of k_{add} there will be a higher probability of obtaining a behavioural output.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.8: CDFs of k_{d1p} for the monomer conversion.



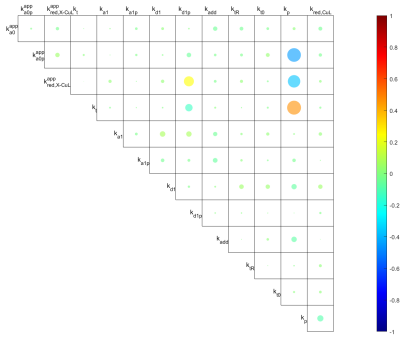
(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.9: CDFs of k_{add} for the monomer conversion.

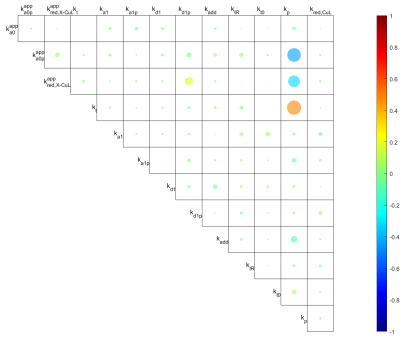
As for the correlation analysis, a graphical representation of the empirical correlation matrix was plotted, similarly to what was done by Saltelli et al. (2004). The result for the monomer conversion is in the Figures 5.10a, 5.10b and 5.10c for the 3 intervals. An approximation of these figures is given in the Appendix D in Figures D.1, D.2 and D.3 for the 1st, 2nd and 3rd interval, respectively. In Figure 5.10, the higher the correlation coefficient, the larger the circle (in absolute value and between -1 and 1). When the coefficient is closer to 1 (positive correlation) the color of the circle will be closer to red. If the coefficient is closer to -1 (negative correlation), the color will approach the colour blue.

From the Figures, it is possible to observe that all the highest values of the correlation coefficient involve the k_{p} constant, suggesting that this parameter has a great interaction with the remaining inputs. In particular with those parameters that have obtained the critical classification by the K-S two sample test. The highest correlation of k_{p} is with k_{t} with a value of 0.4664 for the 1st interval, 0.4868 for the 2nd interval and 0. 3968 for the 3rd interval, followed by $k_{\mathrm{a0p}}^{\mathrm{app}}$ with -0.4485 for the 1st interval, -0.4441 for the 2nd interval and -0.2918 for the 3rd interval. The 3rd highest is the with $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$, obtaining a value of -0.3757 for the 1st interval, -0.3329 for the 2nd interval and -0.2062 for the 3rd interval. The decrease of the correlation values along the interval size can be justified by the fact that in the larger interval very extreme values are explored, which can make the interactive behaviour between parameters more obvious.

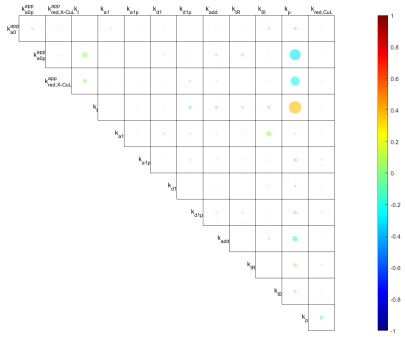
It can also be observed that k_{d1p} interacts with $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ with a correlation coefficient of



(a) 1st interval.



(b) 2nd interval.



(c) 3rd interval.

Figure 5.10: Graphical representation of the empirical correlation matrix for the monomer conversion.

0.1756. Strangely enough, in the 3rd interval a correlation coefficient between these two parameters is also detected with a higher coefficient value of about 0.2444. Nevertheless, neither in the K-S two sample test or by a comparison between the two CDFs of k_{d1p} it is possible to observe any influence of this parameter to drive the model output to the good behaviour.

The sign of the correlation coefficients can also be useful to qualitatively understand the interaction between the set of parameters (for more information see Saltelli et al. (2004) and Saltelli et al. (2008)). However, this topic will not be discussed in this work due to the complexity of the model.

5.2.3 PDI output results

The results for the K-S two sample test for the PDI output are given in the Tables 5.5, 5.6 and 5.7.

When one analyzes the values obtained in the K-S two sample test it can be observed that in contrast to the monomer conversion, the results show a variation depending on the interval studied, in particular between the 2nd and the 3rd interval. These differences can be observed in

Table 5.5: 1st interval PDI MCF result.

Parameter	$d_{n,\overline{n}}$	Classification
$k_{\mathrm{a0}}^{\mathrm{app}}$	0.0166	Insignificant
$k_{\mathrm{a0p}}^{\mathrm{app}}$	0.7236	Critical
$k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$	0.6367	Critical
$k_{\mathrm{red,CuL}}^{\mathrm{app}}$	0.0491	Insignificant
k_{a1}	0.0200	Insignificant
k_{a1p}	0.0286	Insignificant
k_{d1}	0.0204	Insignificant
k_{d1p}	0.1293	Critical
k_{p}	0.1171	Critical
k_{add}	0.1293	Critical
k_{t}	0.0249	Insignificant
k_{tR}	0.0236	Insignificant
k_{t0}	0.0280	Insignificant

Table 5.6: 2nd interval PDI MCF result.

Parameter	$d_{n,\overline{n}}$	Classification
$k_{\mathrm{a0}}^{\mathrm{app}}$	0.0259	Insignificant
$k_{\mathrm{a0p}}^{\mathrm{app}}$	0.5414	Critical
$k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$	0.3754	Critical
$k_{\mathrm{red,CuL}}^{\mathrm{app}}$	0.0120	Insignificant
k_{a1}	0.0186	Insignificant
k_{a1p}	0.0332	Insignificant
k_{d1}	0.0151	Insignificant
k_{d1p}	0.1831	Critical
k_{p}	0.1396	Critical
k_{add}	0.2033	Critical
k_{t}	0.0089	Insignificant
k_{tR}	0.0090	Insignificant
k_{t0}	0.0620	Critical

Table 5.7: 3rd interval PDI MCF result.

Parameter	$d_{n,\overline{n}}$	Classification
$k_{\mathrm{a0}}^{\mathrm{app}}$	0.0742	Critical
$k_{\mathrm{a0p}}^{\mathrm{app}}$	0.0776	Critical
$k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$	0.0685	Critical
$k_{\mathrm{red,CuL}}^{\mathrm{app}}$	0.0150	Insignificant
k_{a1}	0.0188	Insignificant
k_{a1p}	0.0970	Critical
k_{d1}	0.0120	Insignificant
k_{d1p}	0.2238	Critical
k_{p}	0.0380	Critical
k_{add}	0.6053	Critical
k_{t}	0.0973	Critical
k_{tR}	0.0212	Insignificant
k_{t0}	0.2214	Critical

the parameter ranking of importance, based on the value of $d_{n,\overline{n}}$, that changes significantly in the last interval, but also in the number of parameters considered critical: 5 in the 1st interval and 9 in the 3rd interval. As mentioned before, it is expected that using the input range of the 2nd and 3rd interval will require a smaller number of N for the $d_{n,\overline{n}}$ to remain constant. This could potentially increase the number of parameters that are considered important/critical due to the p-value problem. Also, using different input spaces for the GSA can affect the result obtained, as mentioned earlier in this work.

The SARA-ATRP specific constants, $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$, are the most critical constants to

drive the model output to the good behaviour in the 1st and 2nd interval based on the $d_{n,\bar{n}}$ value, but their influence decreases sharply in the 3rd interval, with its ranking going down to the 5th and 6th position, respectively. From a mechanistic point of view, the change in the value of these kinetic parameters causes a variation in the concentration of propagating chains, dormant chains, activator and deactivator. Note that the concentration of dormant chains and deactivator are terms in the PDI equation for an ATRP reaction (see equation 2.6). For instance, the increase of these kinetics parameters will increase the concentration of propagating chains, leading to an increase of the termination and uncontrolled propagation. On the contrary, if the value of the parameters are smaller, the concentration of propagating chains will be lower, decreasing the PDI. These results can be in conflict with the conclusions of Krys et al. (2017) where it is demonstrated by simulation, with a one factor variation, that the the SARA-ATRP specific constants do not change the PDI meaningfully. However, it should be taken into account that in this work one analyzed the error between the experimental and model prediction of the PDI over time, whereas in Krys et al. (2017) it is the PDI vs monomer conversion profile that it is studied. The GSA was also tested considering PDI vs monomer conversion profile and significant differences in the results were obtained, this could be explored in future works. With respect to the difference between the parameters ranking obtained using the 2nd or 3rd intervals input space. As mentioned in Chapter 3, differences in the parameters input space can severely affect the outcome of the sensitivity analysis, being this a expected result. At last, another SARA-ATRP parameter appears with the critical classification in the 3rd interval, the $k_{\mathrm{a0}}^{\mathrm{app}}$. This can be the result of interactions between inputs, such as k_{add} and k_{t0}, because these parameters are also important in this interval and all are kinetic rate constants of reactions that occur in the early stages of the polymerization (see reaction scheme 4.1).

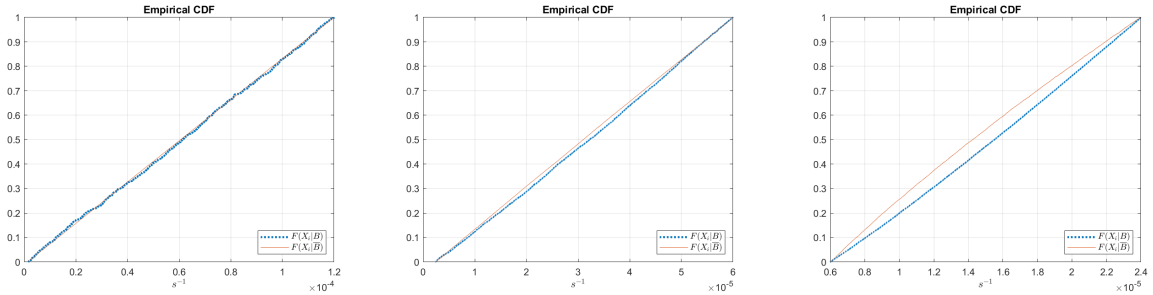
With respect to the k_{d1p} constant, it changes its ranking along the intervals but always with a small variation, being the 3rd in the 1st interval, 4th in the 2nd interval and 2nd in the 3rd interval. This parameter was expected to be critical because it corresponds to the deactivation rate constant of propagating chains of the ATRP equilibrium that is the key to get a low PDI. Also, k_{d1p} appears directly in the PDI equation for an ATRP system (see Equation 2.6). The decrease and increase of this constant has the same effect as the variation of the SARA-ATRP specific parameters. A polymer scientist can find this result questionable because if the equilibrium reaction is the responsible for the success of the ATRP reaction, then k_{d1p} must be the most important parameter for this specific output. However, the results obtained using the MCF method do not explicit explain which parameters are important for the model output, but rather the ones that are critical in order to get a behavioural output. Then, a different definition of a behavioural output can affect the results. If, for instance the behavioral output was specified for o a PDI lower than 1.10, then k_{d1p} could be the most important parameter. Another important point that could affect the results, is the fact that in this GSA we consider the NRMSE and not PDI directly. Another kinetic constant from the ATRP equilibrium (k_{a1p}), is classified as critical only in the 3rd interval being considered less important than k_{d1p} through the analysis of the

$d_{n,\overline{n}}$ value. The greater sensitivity of the PDI with respect to k_{d1p} in comparison with k_{a1p}, was also detected in the local sensitivity analysis perform in Lyra et al. (2019).

The k_{add} and k_{t0} parameters appear to be the most important parameters in the 3rd interval, being k_{add} the 1st and k_{t0} the 2nd. Actually, k_{t0} increases its ranking as the interval decreases, being insignificant in the 1st interval and the 6th more important parameter in the 2nd interval. On the other hand, k_{add} has the same ranking in the 1st and 2nd interval, the 3rd position. This result is unexpected. As mentioned before, these are parameters of reactions that occur in the polymerization initial stage. A subsequent application of the MCF method was performed where the first experimental point was not considered in the calculation of the NRMSE. To obtain these results the model was simulated using 10000 model evaluations and the 3rd interval input space. The results show a significant decrease on the ranking and classification for both the k_{add} and k_{t0}, where k_{add} was classified important and k_{t0} insignificant. This results clearly show that the importance of these two parameters are linked to the first experimental point.

The parameter k_{p} obtained the critical classification in all 3 intervals, being the 5th most important in the 1st and 2nd interval and 9th in the 3rd interval. This kinetic constant appears in the PDI equation (Equation 2.6), so its importance for this output is expected. For example, an increase in the value of k_{p} will lead to a higher rate of monomer addition, which will have a negative impact on the micro structural control of the polymers, increasing the PDI (Al-Harthi et al., 2006). Finally, the kinetic constant k_{t} appears as the 4th most important parameter only in the 3rd interval. Analysing Equation 2.6, it is possible to see that in fact k_{t} can directly affect some variables in this equation, such as: the concentration of dormant chains; the monomer conversion; and the degree of polymerization.

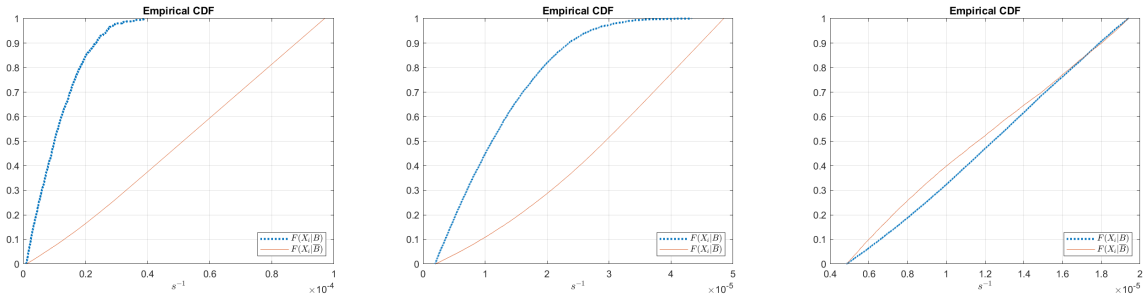
The CDFs of each parameter that receive the critical classification for the PDI are represented from the Figure 5.11 to 5.19 for the 1st, 2nd and 3rd interval.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

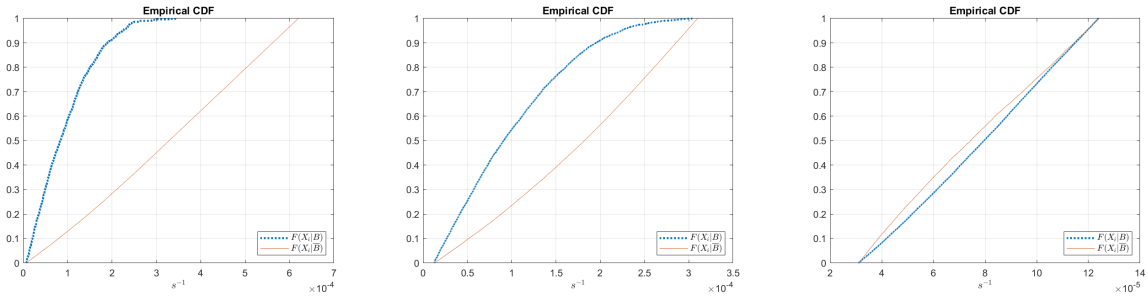
Figure 5.11: CDFs of $k_{\mathrm{a0}}^{\mathrm{app}}$ for the PDI.

One can observe that for most of the parameters the profile of the two curves, $F_n(X_i|B)$ and $F_{\overline{n}}(X_i|\overline{B})$, is quite different in the 3rd intervalb when compared to the profiles in the 1st and 2nd interval. This confirms what was observed in the K-S two sample test that in the 3rd interval the importance of the parameters is different of the one verified with the 1st and 2nd intervals.



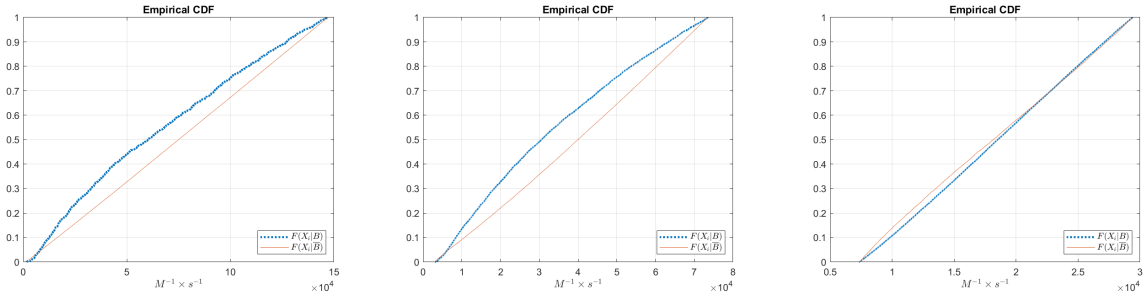
(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.12: CDFs of $k_{\mathrm{a0p}}^{\mathrm{app}}$ for the PDI.



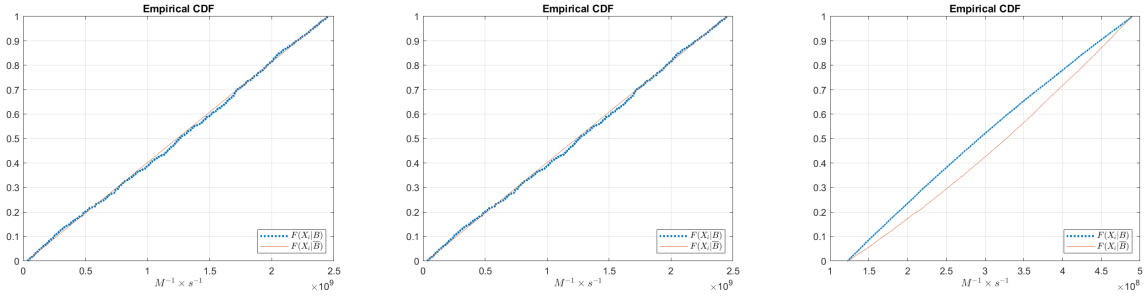
(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.13: CDFs of $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ for the PDI.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

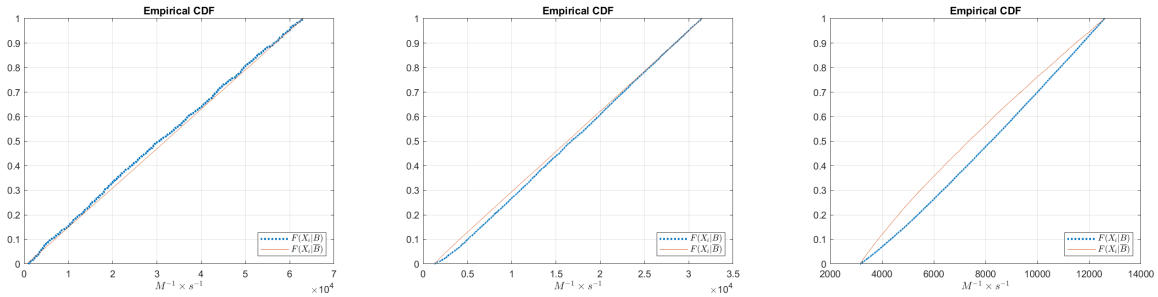
Figure 5.14: CDFs of k_{p} for the PDI.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

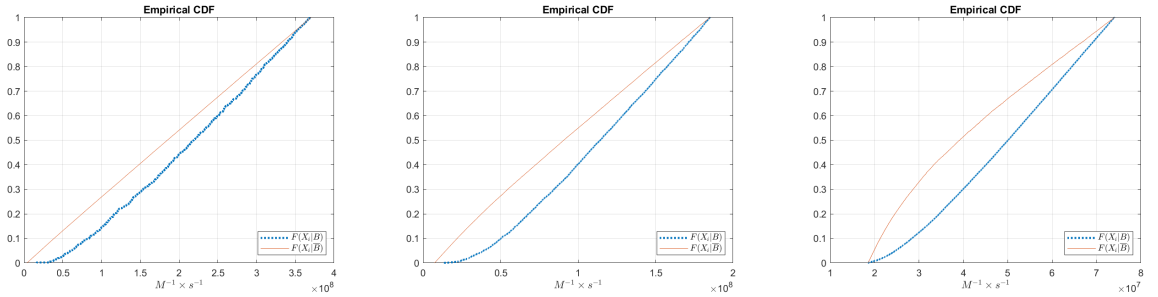
Figure 5.15: CDFs of k_{t} for the PDI.

Analysing Figure 5.11, it is possible to see that for the parameter $k_{\mathrm{a0}}^{\mathrm{app}}$, only a significant difference between the two CDFs is verified in the 3rd interval. Observing Figure 5.11c, it is



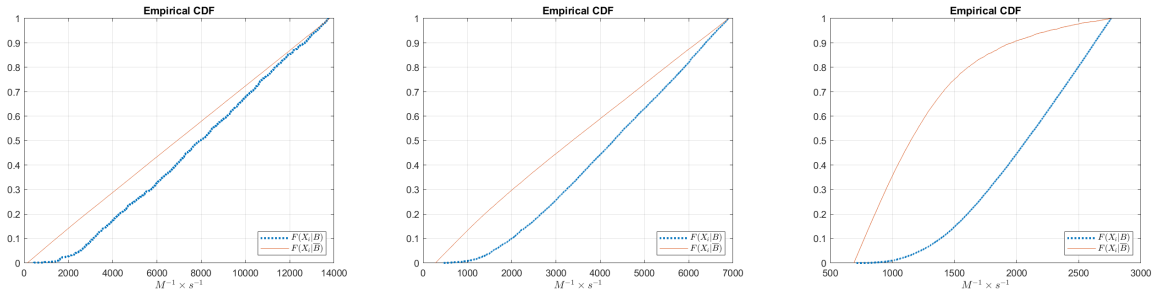
(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.16: CDFs of k_{a1p} for the PDI.



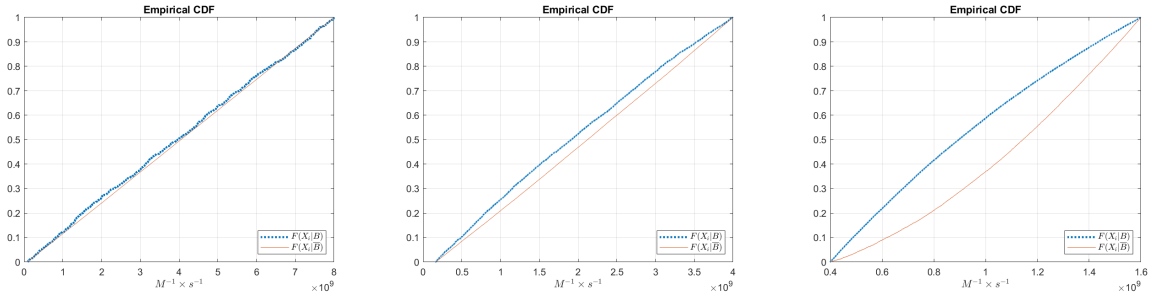
(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.17: CDFs of k_{d1p} for the PDI.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.18: CDFs of k_{add} for the PDI.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.19: CDFs of k_{t0} for the PDI.

possible to verify that by using higher values for this constant, there is more probability of obtaining a behavioural output. In Figures 5.12 and 5.13, corresponding to the parameters, $k_{\mathrm{a0p}}^{\mathrm{app}}$

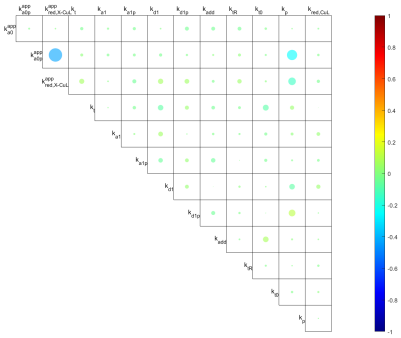
and $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$, is clearly visible the similarities between this two figures in all the intervals studied. For the 1st and 2nd interval the probability of obtaining a behavioural output by using low values of these kinetic constants is higher. However, in the 3rd interval low values of this parameter should not be used. Also, for the CDF curves in the 1st and 2nd interval is possible to see that from values of 4×10^{-5} s^{-1} for $k^{\mathrm{app}}_{\mathrm{a0p}}$ and 3×10^{-4} s^{-1} for $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$ there is zero probability of obtaining a behavioural output.

Figure 5.14 shows that for the 1st and 2nd interval low values of k_{p} should be used in order to have a better chance of obtaining a behavioural output. The exception is the extreme lower values of the intervals, where the two CDFs are overlapped. However, similarly to the result obtained for $k^{\mathrm{app}}_{\mathrm{a0p}}$ and $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$, the CDFs for the 3rd interval indicates that using low values of k_{p} it leads to an increase in the probability of obtaining a non-behavioural output. For the parameters k_{t} and k_{a1p}, as well as to $k^{\mathrm{app}}_{\mathrm{a0}}$ only have a significant difference between the two CDFs in the 3rd interval, as it can be seen in the Figures 5.15 and 5.16, there is an higher probability of obtaining a behavioural output by using low values of k_{t} and high values of k_{a1p}.

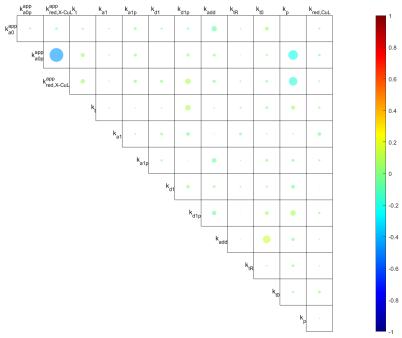
Both parameters k_{d1p} and k_{add} present very similar CDFs in the 1st and 2nd interval (Figures 5.17 and 5.18). Here, there is a higher probability of obtaining a good behavioral result when high values of the input range are used, for both parameters. This probability is almost zero if the values are very close to the interval lower bound. In the 3rd interval there are considerable differences between the two parameters. Here, the CDFs of k_{d1p} are similar to the ones obtained for the 1st and 2nd interval. Contrarily, in the Figure 5.18 bigger differences between $F_n(X_i|B)$ and $F_{\overline{n}}(X_i|\overline{B})$ are detected for k_{add} when compared to the one observed in 1st and 2nd intervals. This clearly shows the increase of importance of this parameter as the interval decreases. The parameter k_{t0} only presents a significant difference in the 2nd and 3rd interval, and lower values of this input will have more probability of producing a behavioural output.

The graphical representation of the empirical correlation matrix was obtained for the PDI output for the 1st, 2nd and 3rd interval. These results are represented in the Figures 5.20a, 5.20b and 5.20c, respectively. An approximation of these figures are in the Appendix D in the Figures D.4, D.5 and D.6 for the 1st, 2nd and 3rd interval, respectively.

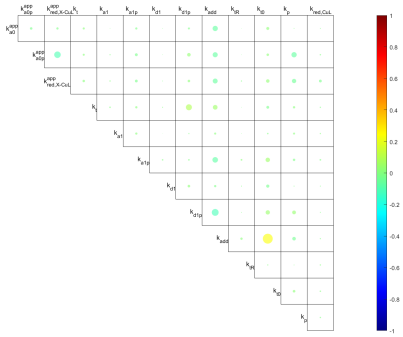
From Figure 5.20 one can make the same conclusions as in the case of the result obtained in K-S two sample test and by the analysis of the CDFs. There is a big difference between the results obtained in 1st and 2nd interval and the results of the 3rd interval. In the 1st and 2nd interval, the biggest correlation appears with the two most important parameters, $k^{\mathrm{app}}_{\mathrm{a0p}}$ and $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$, being the value -0.4309 for the 1st and -0.4546 for the 2nd interval. Another substantial correlation appears with $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$ and k_{p}, being the value of the correlation -0.2543 for the 1st and -0.2261 for the 2nd interval. For the 3rd interval the biggest correlation is with the two most important parameters, as in 1st and 2nd interval. However, in this case the parameters k_{add} and k_{t0} have a significant smaller value of correlation, only 0.2375. In this interval, there is no other substantial correlation between the parameters.



(a) 1st interval.



(b) 2nd interval.



(c) 3rd interval.

Figure 5.20: Graphical representation of the empirical correlation matrix for the PDI.

5.2.4 M_n output results

The results for the M_n for the three intervals are presented in the Tables 5.8, 5.9 and 5.10.

The results obtained for the output M_n are very similar to those obtained for the monomer conversion. Here, the same 4 parameters are significant to obtain the good model response ($k^{\text{app}}_{\text{a0p}}$, $k^{\text{app}}_{\text{red,X-CuL}}$, k_{t} and k_{p}). In fact, the same conclusion was presented in Flores-Tlacuahuac et al. (2003), where in a sensitivity study of a controlled polymerization, by a local sensitive technique, the most influential parameters that were identified are the same for the monomer conversion and M_n output. For a RDRP, M_n depends directly from the monomer conversion (Equation 2.3). If the rate of propagation is lower, the addition of monomers to the propagating chains will be slowed down. Consequently, the polymers growth will be delayed, changing the M_n time profile.

Analysing the results in more detail, it can be seen that the ranking based on $d_{n,\bar{n}}$ value in 1st and 2nd interval is the same, being k_{p} the most important parameter followed by $k^{\text{app}}_{\text{a0p}}$, $k^{\text{app}}_{\text{red,X-CuL}}$ and k_{t}. In fact, based on the p-value, k_{t} only receives the important classification in the 1st interval. k_{add} is classified as important in the 2nd interval, and is classified as insignificant in the remaining intervals. In the 2nd interval the parameter $k^{\text{app}}_{\text{a0}}$ is also classified as important,

Table 5.8: 1st interval M_n MCF result.

Parameter	$d_{n,\overline{n}}$	Classification
$k_{\mathrm{a0}}^{\mathrm{app}}$	0.1234	Insignificant
$k_{\mathrm{a0p}}^{\mathrm{app}}$	0.6689	Critical
$k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$	0.5986	Critical
$k_{\mathrm{red,CuL}}^{\mathrm{app}}$	0.1110	Insignificant
k_{a1}	0.0696	Insignificant
k_{a1p}	0.0685	Insignificant
k_{d1}	0.0661	Insignificant
k_{d1p}	0.1200	Insignificant
k_{p}	0.7086	Critical
k_{add}	0.1097	Insignificant
k_{t}	0.1412	Important
k_{tR}	0.1266	Insignificant
k_{t0}	0.0857	Insignificant

Table 5.9: 2nd interval M_n MCF result.

Parameter	$d_{n,\overline{n}}$	Classification
$k_{\mathrm{a0}}^{\mathrm{app}}$	0.0655	Important
$k_{\mathrm{a0p}}^{\mathrm{app}}$	0.4580	Critical
$k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$	0.3476	Critical
$k_{\mathrm{red,CuL}}^{\mathrm{app}}$	0.0423	Insignificant
k_{a1}	0.0397	Insignificant
k_{a1p}	0.0577	Insignificant
k_{d1}	0.0234	Insignificant
k_{d1p}	0.0613	Insignificant
k_{p}	0.6337	Critical
k_{add}	0.0675	Important
k_{t}	0.1398	Critical
k_{tR}	0.0605	Insignificant
k_{t0}	0.0524	Insignificant

Table 5.10: 3rd interval M_n MCF result.

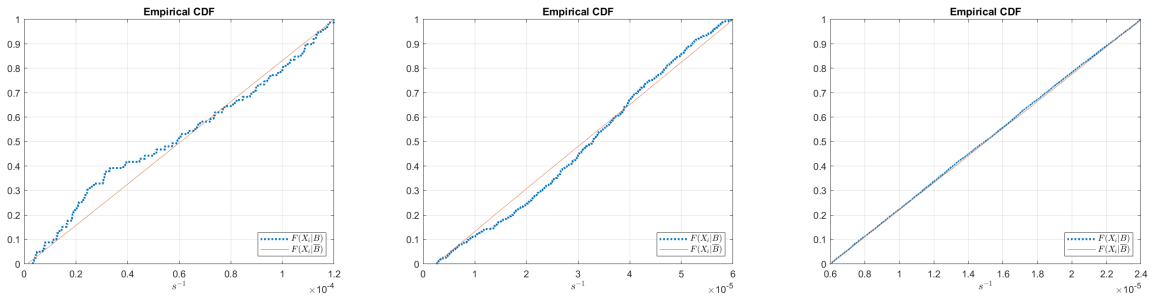
Parameter	$d_{n,\overline{n}}$	Classification
$k_{\mathrm{a0}}^{\mathrm{app}}$	0.0117	Insignificant
$k_{\mathrm{a0p}}^{\mathrm{app}}$	0.1827	Critical
$k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$	0.1208	Critical
$k_{\mathrm{red,CuL}}^{\mathrm{app}}$	0.0091	Insignificant
k_{a1}	0.0113	Insignificant
k_{a1p}	0.0149	Insignificant
k_{d1}	0.0075	Insignificant
k_{d1p}	0.0260	insignificant
k_{p}	0.6692	Critical
k_{add}	0.0109	Insignificant
k_{t}	0.1987	Critical
k_{tR}	0.0135	Insignificant
k_{t0}	0.0091	Insignificant

being the last in the ranking. In the 3rd interval, k_{p} continues to be most important parameter but k_{t} becomes the 2nd most influential parameter followed by $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$.

The variation in the value of k_{p}, will influence the rate of the monomer addition to the propagating chains (Al-Harthi et al., 2006). Thus, it affects the evolution of M_n along the time. The value of k_{t} will influence the concentration of propagating chains that will affect the number of active sites, consequently affecting as well the evolution of M_n. Regarding the SARA specific kinetic constants, $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$, the importance of these two parameters to the M_n can be due to their influence on the concentration of propagating and dormant chains through the

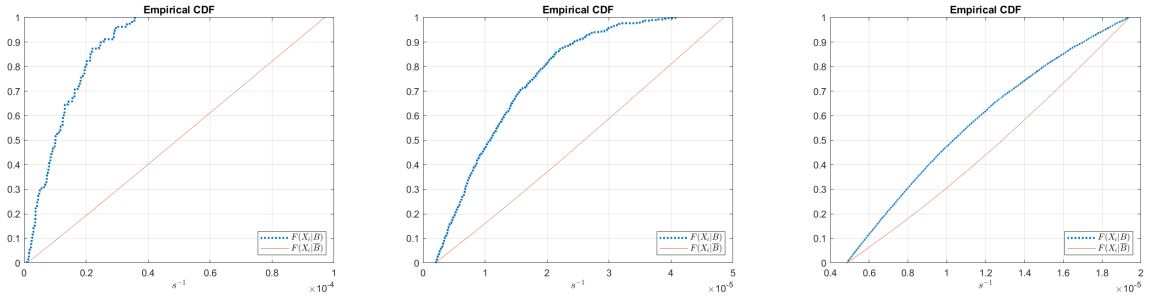
direct activation of dormant chains, by $k_{\mathrm{a0p}}^{\mathrm{app}}$, or indirectly by the regeneration of the activator ($\mathrm{Cu^I/L}$) with the kinetic constant of $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$. The reason for the importance of k_{add} in the 2nd interval was studied and the conclusion was the same as the one obtained for the PDI output. The importance of this parameter is related to the first experimental point. If this point is removed, the classification of k_{add} will be insignificant. The importance of $k_{\mathrm{a0}}^{\mathrm{app}}$ can be due to some interaction with k_{add}, because they are kinetic rate constants of reactions that occur in the early stages of the polymerization.

The CDFs for all the parameters that get the critical/important classification for all the 3 intervals are presented in the Figures 5.21 to 5.26. At first glance, it is possible to verify that for the M_n output the CDFs of $F(X_i|B)$ are not as smoother as the ones obtained for the monomer conversion and the PDI. The reason for this is that for the M_n the number of outputs that meet the objective are smaller than for the monomer conversion and the M_n, as it can be seen in the Table 5.1. The CDFs of this output have a less smoother profile. This effect is more visible in the 1st interval because is the one where is more difficult to obtain behavioural outputs (see Table 5.1.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

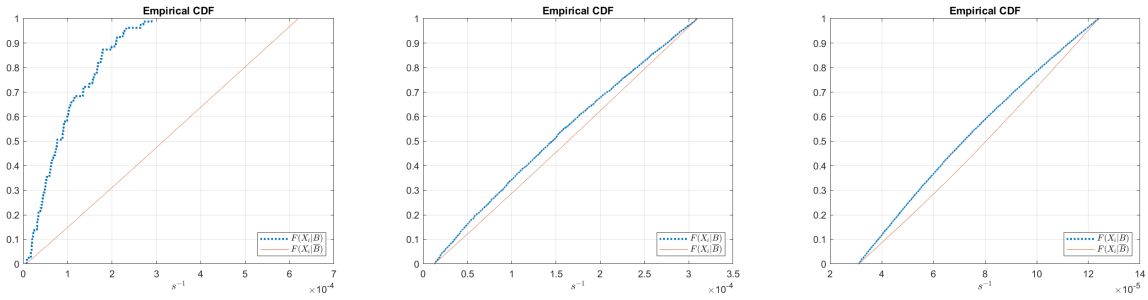
Figure 5.21: CDFs of $k_{\mathrm{a0}}^{\mathrm{app}}$ for the M_n.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

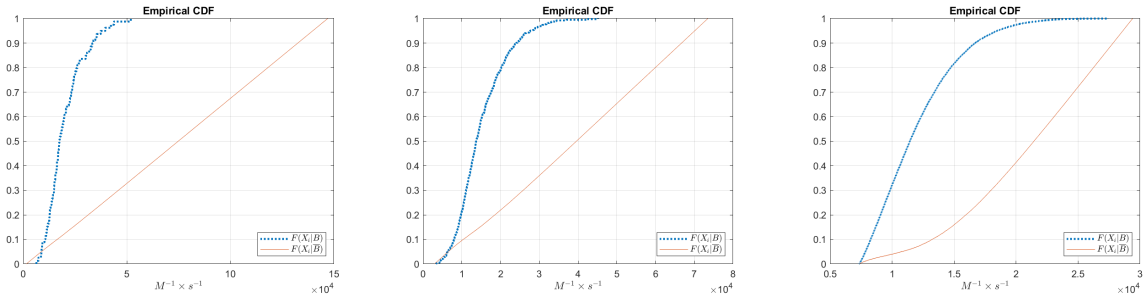
Figure 5.22: CDFs of $k_{\mathrm{a0p}}^{\mathrm{app}}$ for the M_n.

For the parameter $k_{\mathrm{a0}}^{\mathrm{app}}$ (Figure 5.21) it is possible to verify that in the 2nd interval no meaningful conclusion can be taken from the CDFs. Nevertheless, the CDFs on the 1st interval appear to have a bigger difference compared to the ones on the 2nd interval. However, because of the p-value the parameter is classified as insignificant. From Figures 5.22 and 5.23 it can be seen that for the remaining SARA-ATRP parameters, $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$, there is a higher



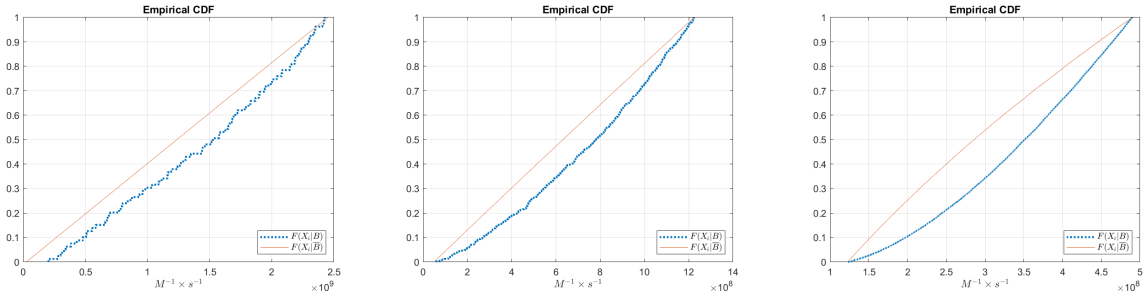
(a) 1st interval. **(b)** 2nd interval **(c)** 3rd interval

Figure 5.23: CDFs of $k^{\text{app}}_{\text{red,X-CuL}}$ for the M_n.



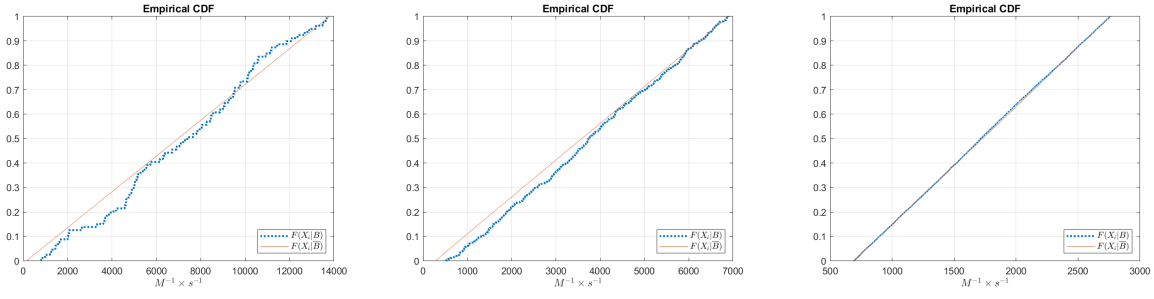
(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.24: CDF of k_{p} for the M_n.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.25: CDFs of k_{t} for the M_n.



(a) 1st interval. **(b)** 2nd interval. **(c)** 3rd interval.

Figure 5.26: CDFs of k_{add} for the M_n.

probability of generating a behavioural output by using low values of this parameters, in all the intervals. However, if values of $k^{\text{app}}_{\text{a0p}}$ higher than 4×10^{-5} s^{-1} are used, there will be zero

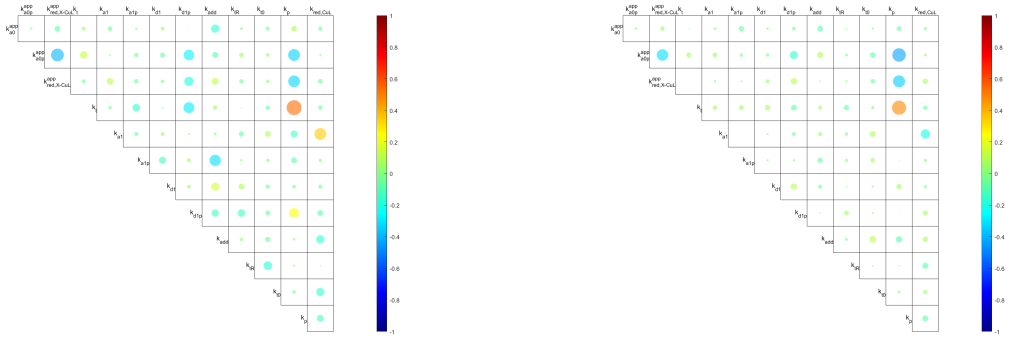
probability of obtaining a good output in the 1st and 2nd interval. The same can be concluded for $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$ in the 1st interval, but from values higher than 3×10^{-4} s^{-1}. As for k_{p}, in Figure 5.24, it can be observed that in the 1st and 2nd interval low values of this parameter will lead to a higher probability of obtaining a behavioural output, except for the ones closer to the lower bound. From these figures, it is possible to concluded that above a given value of the parameter, when the slope of the $F_{n,\overline{n}}(X_i|B)$ CDF is zero, there is no probability of obtaining a good model output. This value seems to change accordingly to the interval under study. Figure 5.25 shows that there is more probability of obtaining a behavioural output by using high values of k_{t}.

Finally, in Figure 5.26, for k_{add}, it seems that in the 1st interval there is a difference between the two CDFs and not just in 2nd interval as detected by the two sample K-S test. This result was further studied by performing the two sample K-S test with more model evaluations, 100000 and 200000. With 100000 evaluations the p-value obtained was 0.1078, being almost classified as a important parameter. Using 200000 model evaluations, the parameter receives the critical classification with a p-value of 0.0054. This clearly shows how the result of the p-value, and consequently the classification of the model parameters, can change due to a higher N. This parameter could be also classified as important, due to the influence of the first experimental point, similarly to the case of the PDI output.

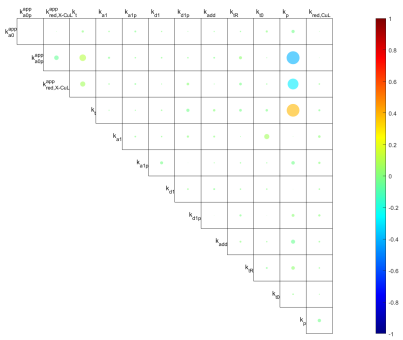
As for the other model outputs, a graphical representation of the empirical correlation matrix for the M_n is presented in the Figures 5.27a, 5.27b and 5.27c for the 1st, 2nd and 3rd interval, respectively. An approximation of these figures is given in the Appendix D in the Figures D.7, D.8 and D.6, respectively.

Analysing the Figure 5.27, it is easy to notice the similarities between these results and the ones obtained for the monomer conversion, where the main correlations are with k_{p}. On the other hand, some differences between the two outputs are visible. At first, in the 1st and 2nd interval there is a high correlation between the SARA rate constants ($k^{\mathrm{app}}_{\mathrm{a0p}}$ and $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$), which is not obtained in the monomer conversion output. Also, in the 1st interval the graphical representation of the M_n presents more circles with a larger size compared to the one obtained for the monomer conversion. This can be an indicator, that in this interval, the interactions between parameters are more pronounced than the ones obtained for the monomer conversion.

For the 1st interval the highest correlation is between k_{p} and k_{t}, followed by the correlation between the SARA-ATRP constants, with values of 0.5352 and -0.3163, respectively. Other important correlations are with k_{p} and two other constants, $k^{\mathrm{app}}_{\mathrm{a0p}}$ and $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$ the correlation values are -0.337 and -0.3317, respectively. In the 2nd interval the important correlations are the same, expect for the correlation between the SARA-ATRP parameters, which is smaller. The values of these correlations are 0.4830, -0.4363, -0.3496 and -0.3163, respectively. In the last interval, the most important correlations are between k_{p} and $k^{\mathrm{app}}_{\mathrm{a0p}}$, followed by the one with k_{t} and $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$, with the values of -0.4039, 0.3899 and -0.2826, respectively.



(a) 1st interval. **(b)** 2nd interval.



(c) 3rd interval.

Figure 5.27: Graphical representation of the empirical correlation matrix for the M_n.

5.3 Sobol method

5.3.1 Methods comparison

As mentioned before, the method to calculate the sensitivity indices presented in Saltelli et al. (2008) and its extension in Wu et al. (2012), will be compared using a simple function given by (5.1). This equation makes possible to calculate the first order index analytically (see Saltelli et al. (2008), pages 174-176).

$$y = X_1 + X2 + X3 \tag{5.1}$$

The values of the parameters X_1, X_2 and X_3 follow a uniform distributions within the intervals presented in (5.2).

$$X_1 \sim U(0.5, 1.5) \qquad X_2 \sim U(1.5, 4.5) \qquad X_3 \sim U(4.5, 13.5) \tag{5.2}$$

The first order sensitivity indices, calculated analytically in Saltelli et al. (2008), for the three inputs, are presented in Table 5.11.

The two methods were implemented in MATLAB® and the first order indices were calculated

Table 5.11: First order sensitivity indices true values of the parameters in Equation 5.1.

Parameter	S_i
X_1	0.011
X_2	0.099
X_3	0.890

for N = 10, 100, 1000 and 10000 using the two computational procedures. The results were obtained for each parameter. In order to compare the two methods, the result of the sensitivity indices was plotted in function of $log_{10}(N)$. In these plots, is also represented the first order sensitivity index calculated analytically. The results are represented in the Figures 5.28, 5.29 and 5.30.

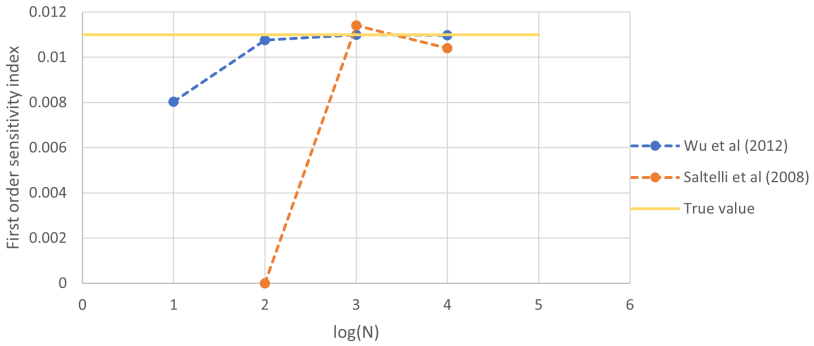


Figure 5.28: Results for the parameter X_1.

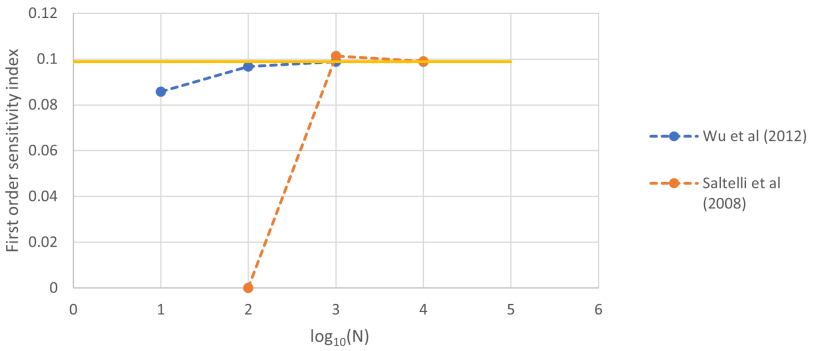


Figure 5.29: Results for the parameter X_2.

It should be noted that the results for $N = 10$ are not represented for the method of Saltelli et al. (2008). This is because the values calculated for X_1 and X_2 are negative. This can happen due to numerical errors as it is discussed in Saltelli et al. (2008). In an example presented in Saltelli et al. (2008) of an infection model sensitivity analysis (page 169), two negative sensitivity indices are obtained, with values smaller than -1×10^{-2}. Despite these results, the GSA is carried out considering that these inputs are unimportant because the sensitivity indices are negative, but almost zero. In this case, the values obtained for X_1 and X_2 are -0.50 and -0.17,

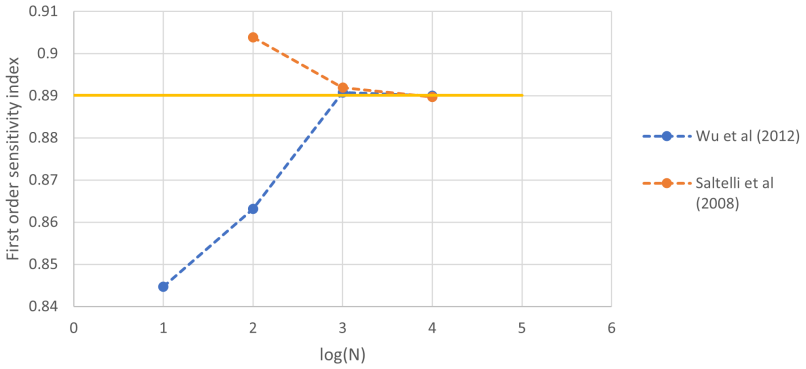


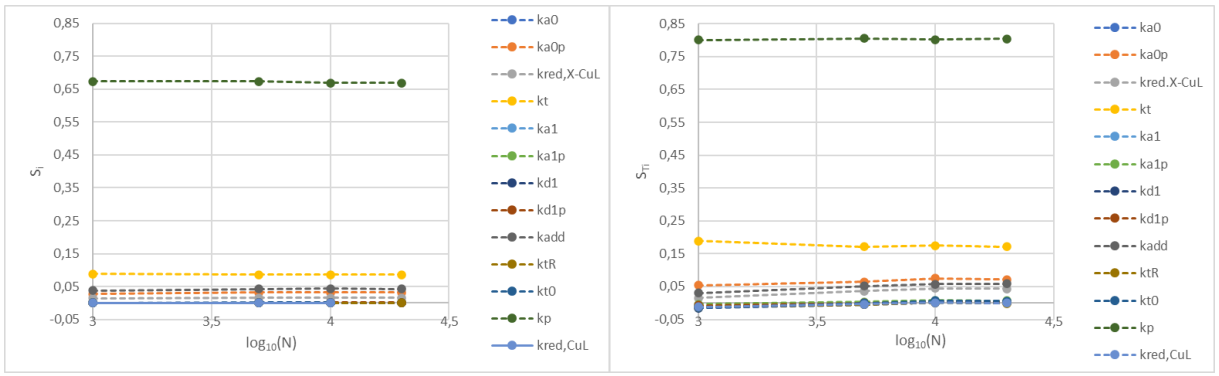
Figure 5.30: Results for the parameter X_3.

respectively, smaller than the ones obtained in the book example. Therefore, it was decided that these values should not be represented and it was concluded that the method of Saltelli et al. (2008) requires a larger N to give acceptable results for this specific function. With N = 100 the method gives negative sensitivity indices, but closer to zero.

In general, the method of Wu et al. (2012) has a better performance, does not give negative sensitivity indices and provides values closer to the ones calculated analytically, even for a smaller N. This method gives the correct parameter ranking even using a $N = 10$. For the example under analysis, the only exception is for the case of the parameter X_3, in which the values obtained by the Saltelli method are closer to the real value with a smaller number of N. Also, the Wu et al. (2012) was used to performed a GSA in a polymerization kinetic model in Salas et al. (2019). Hence, it was decided to use this method. It is worthy to mention that, before applying the proposed method to the SARA-ATRP kinetic model, an implementation test was made to verify if the method was properly implemented in MATLAB® for the calculation of the first and total order sensitivity indices. Further details on this preliminary test are given in Appendix E.

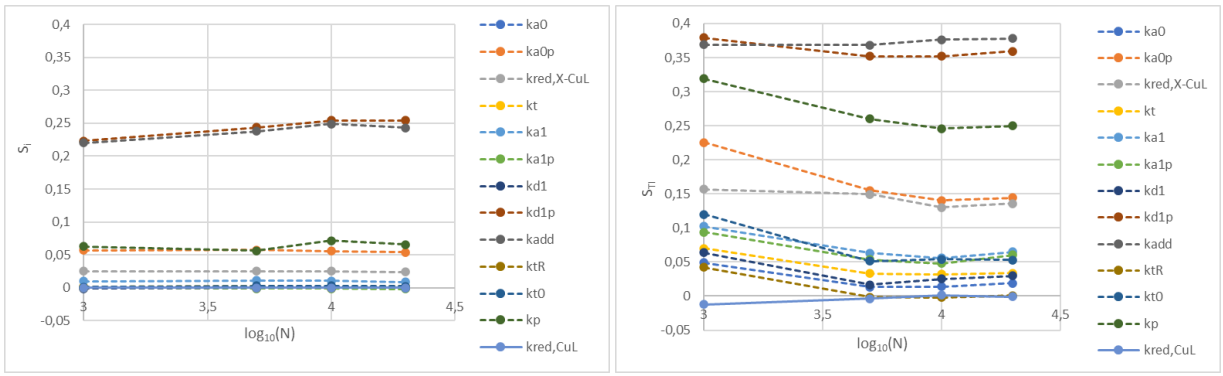
5.3.2 Choosing the number of samples (N)

The first and total order sensitivity indices were calculated for $N = 1000$, 5000, 10000 and 20000. In a preliminary study, it was observed that the values obtained with the 1st interval needed a higher N to reduce the numerical errors of the method, indicating that this interval will need a large sample size. The N selected for this interval will be the one used for the 2nd and 3rd intervals, to ensure a uniform comparison. The results are presented in a plot of the variation of the sensitivity indices in function of $log_{10}(N)$ (Figures 5.31, 5.32, 5.33). Another important point is the computational time needed to calculate these sensitivity indices, which increases with N. This factor is also taken into account in this work and the CPU time for all the N are presented in the Table 5.12.



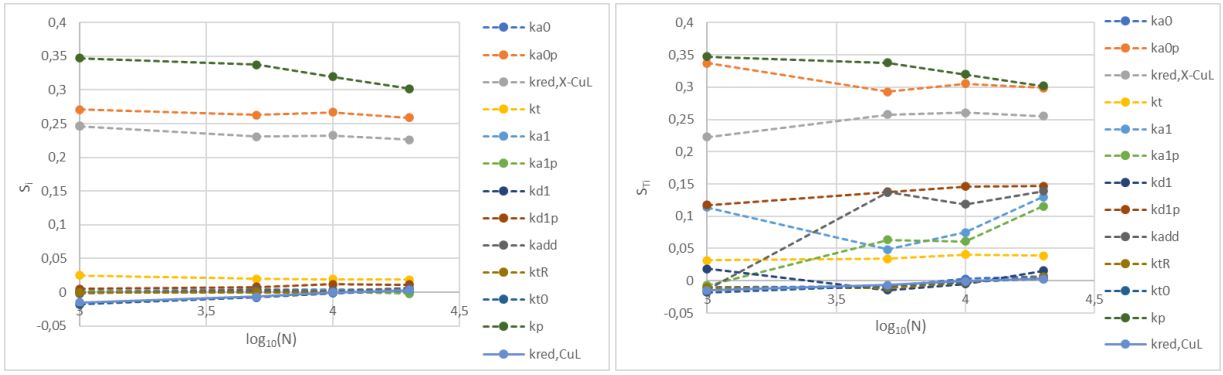
(a) First order indices. **(b)** Total order indices.

Figure 5.31: First and total order indices in function of $log_{10}(N)$ for the monomer conversion.



(a) First order indices. **(b)** Total order indices.

Figure 5.32: First and total order indices in function of $log_{10}(N)$ for the PDI.



(a) First order indices. **(b)** Total order indices.

Figure 5.33: First and total order indices in function of $log_{10}(N)$ for the M_n.

From these figures it follows that the first order sensitivity indices have a more stable behaviour along N, in comparison with the total order sensitivity indices. This is observed for all the 3 outputs. It is more difficult to see these differences in the monomer conversion graphs, due to the scale issues. Therefore, the conclusions will be taken from the total order sensitivity indices where it follows that it is necessary a larger sample size to have stable results. In Figures 5.31 and 5.32 it is possible to see that for most of the parameters the total indices have a stable behaviour for $N = 5000$, and most of them are almost constant for $N = 10000$. The parameters

Table 5.12: CPU time to calculate the first and total indices in function of N.

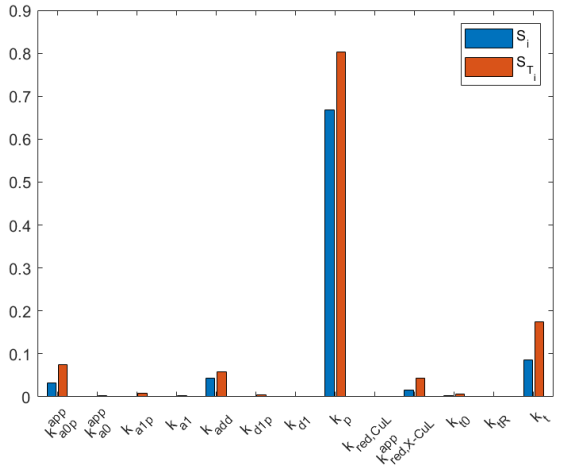
N	Time (s)
1000	880
5000	2404
10000	4610
20000	8334

k_{a1} and k_{a1p} for the M_n output (Figure 5.33b) are the only two exceptions where their values for $N = 10000$ and 20000 are quite different. Another problem that was taken into account, but is not clearly visible in these figures, was the numerical errors that result in negative indices. It was verified that, in general, increasing the sample size decreases the magnitude of these numerical errors (i.e., less negative values of sensitivity indices close to zero). This remark also reported in the work of Yao (2015). As mentioned before, in Saltelli et al. (2008) the sensitivity analysis is pursued even with very small negative values, close to zero. Therefore, in this work it was decided to limit of how negative can be a sensitivity index. A given tolerance of 10^{-2} was specified. It was found that the sensitivity indices smaller than 10^{-2} were only obtained with $N < 10000$. Hence, it was decided to use $N = 10000$, because the results do not change considerably with $N = 20000$ except for the two parameters described above, and all the negative numbers are bigger than 10^{-2}. As for the computational cost it takes approximately more than 1 hour with $N = 20000$ than with $N = 10000$ to calculate the sensitivity indices. Therefore, it was decided that for the specific application and by the results obtained in the previous figures was not worth the extra CPU time.

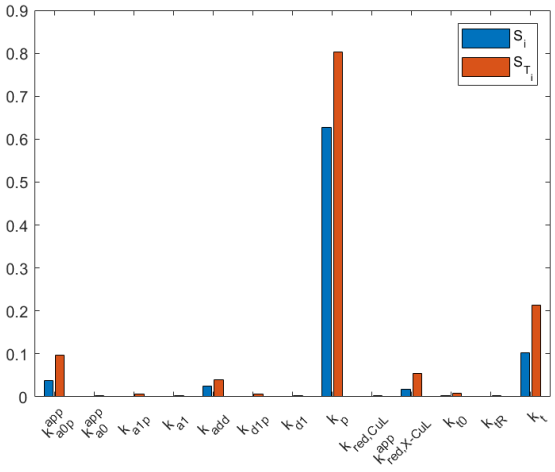
5.3.3 Monomer conversion output results

The first and total sensitivity indices obtained for the monomer conversion are in the Figures 5.34a, 5.34b and 5.34c for the 1st, 2nd and 3rd interval, respectively.

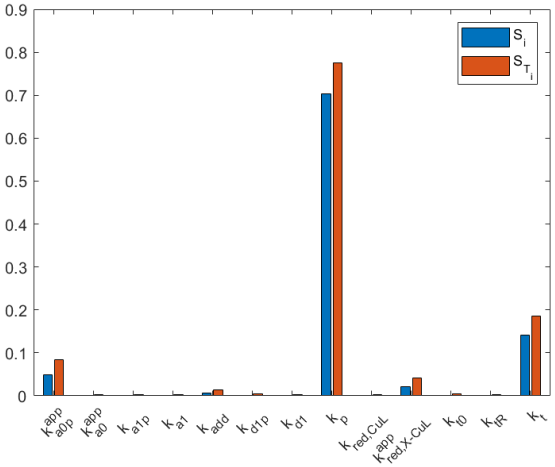
The results obtained for the 3 intervals very similar. Nevertheless, small differences can be detected. The parameters k_{p} and k_{t} are always the two most important parameters with the higher first order and total sensitivity indices. k_{p} features the highest indices by a large margin. This, denotes the importance of this kinetic constant for this output in comparison to the other parameters. The importance of the other kinetic constants that can be considered important to the output, using the threshold of 0.05 defined previously, change in function of the interval considered. These parameters are only important because of the value of the total order sensitivity index. In the 1st interval the parameter $k_{\mathrm{a0p}}^{\mathrm{app}}$ has the 3rd highest total sensitivity index with a value of 0.0753 followed by k_{add} with 0.0573. It is interesting to note that the first order sensitivity index of k_{add} is higher than the sensitivity index of $k_{\mathrm{a0p}}^{\mathrm{app}}$. This indicates that k_{add} contributes the most for the variation of the output, but $k_{\mathrm{a0p}}^{\mathrm{app}}$ interacts more with the other parameters and its total effect (accounting the interactions and its effect alone) can induce more



(a) 1st interval results.



(b) 2nd interval results.



(c) 3rd interval results.

Figure 5.34: Sensitivity indices for the monomer conversion.

variability in the output. In the 2nd interval, the 3rd most important parameter is $k_{\mathrm{a0p}}^{\mathrm{app}}$ followed by $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ with a total sensitivity index of 0.0962 and 0.0541, respectively. At last, for the 3rd interval the only parameter that is considered important to the output, besides k_{p} and k_{t}, is $k_{\mathrm{a0p}}^{\mathrm{app}}$ with a total sensitivity index of 0.0834.

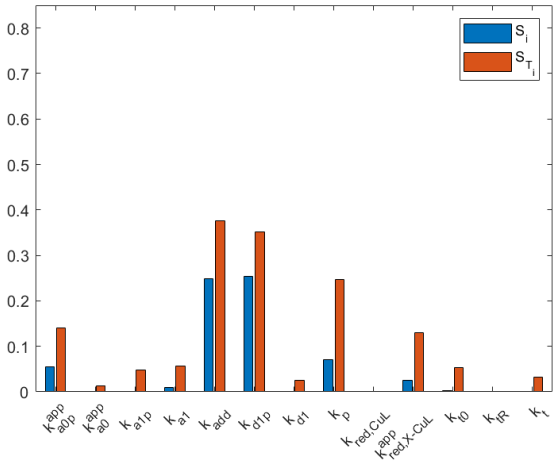
Also, it is interesting that in the 2nd interval there is a stronger interaction between the parameters, with $1 - \sum_i S_i$ approximately 0.1854, while for the 1st and 3rd interval this value is respectively 0.1500 and 0.0746. Observing the important parameters individually, the total order index is always bigger in the case of the 2nd interval. This was unexpected, it was thought that interactions would be more pronounce in the 3rd interval. That is, because the bigger interval includes very extreme values of the parameters, which may lead to a more pronounced interaction between the kinetic constants.

It is possible to observe a quite good agreement between the results obtained by this method and by MCF, where k_{p}, k_{t} and $k_{\mathrm{a0p}}^{\mathrm{app}}$ are considered important parameters for this output (considering

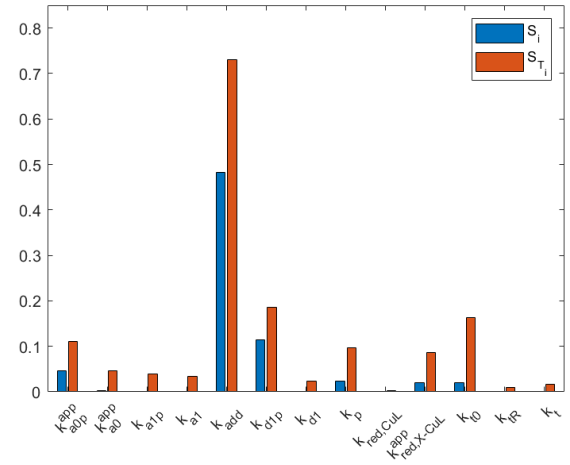
the calculation of the first order and total order sensitivity indices) and critical to obtain the correct behaviour of the model for all the 3 intervals. The kinetic constant $k^{\text{app}}_{\text{red,X-CuL}}$ is always considered critical using the MCF in all the intervals, but the same can not be concluded from the sensitivity indices. In fact, this parameter is always in the top of the ranking (in 4th or 5th place depending on the interval). However, due to the threshold of 0.05 it is only considered important to the output in the 2nd interval. The k_{add} constant is only considered critical in the MCF method results when using the 3rd interval while their importance its only acknowledge in the 1st interval based on the sensitivity index value. In fact, the value of the total order sensitivity index in the 3rd interval has a considerable difference with respect to the limit proposed of 0.05, being this value approximately 0.0136.

5.3.4 PDI output results

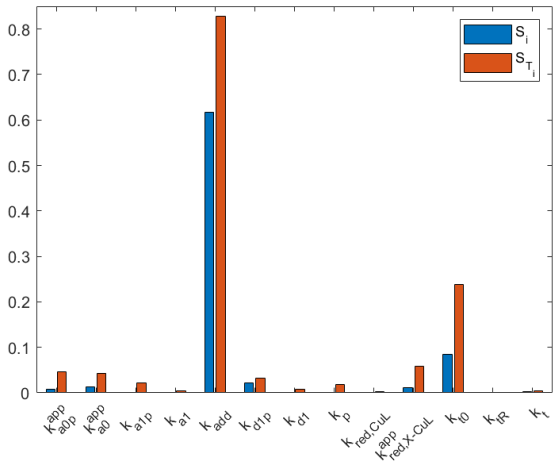
The results for the PDI are presented in the Figures 5.35a, 5.35b and 5.35c for the 1st, 2nd and 3rd interval, respectively.



(a) 1st interval results.



(b) 2nd interval results.



(c) 3rd interval results.

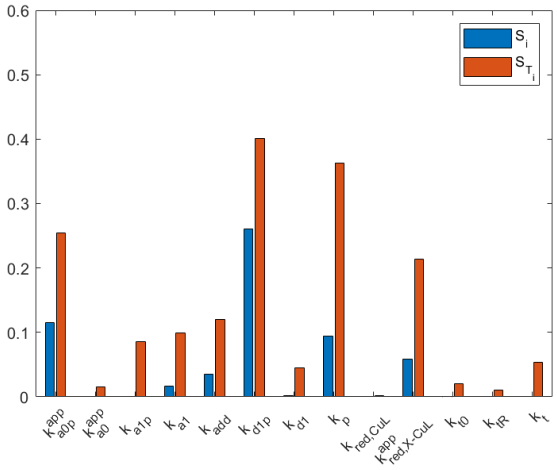
Figure 5.35: Sensitivity indices for the PDI output.

In contrast to what has been observed for the case of the monomer conversion output, a significant difference can be observed between the intervals, especially between the 1st and the other two intervals. The 1st interval is the one where the values of the indices are more distributed between the parameters. In this case, the ranking of the parameters changes if it is considered the first or the total sensitivity order index. The parameter k_{d1p} is the one with the highest first order sensitivity index of 0.2581 followed by k_{add} with a value of 0.2487. However, considering the total order sensitivity index, k_{add} has the highest index with 0.3764 and k_{d1p} is the second in the ranking with 0.3764. This is an indication that k_{add} interacts more with the other parameters. These kinetic constants are followed by k_{p} with a first order and total order sensitivity index of 0.0714 and 0.2461. Then, in 4th place appears $k_{\mathrm{a0p}}^{\mathrm{app}}$ with 0.0554 and 0.1505, respectively. The other parameters that are considered important are $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$, k_{a1} and k_{t0} with a total index of 0.1303, 0.0561 and 0.0540, respectively. The importance of k_{a1} and k_{t0} are mainly due to interactions with the other parameters given that there is still a significant difference between the first order and total order indices. This interactive behaviour can be with k_{add}, because these constants are relevant during the early stage of the reaction. In the 2nd interval, k_{add} is the parameter that affects the most the variability of the output and by a large margin of 0.4818 and 0.7299 for the values of the first and total order indices, followed by k_{d1p} with the indices of 0.1145 and 0.1857, respectively. The other parameters classified as important are k_{t0}, $k_{\mathrm{a0p}}^{\mathrm{app}}$, k_{p} and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ with a total order sensitivity index of 0.1636, 0.1103, 0.0964 and 0.0872, respectively. In the 3rd interval, the importance of k_{add} is even more evident, with a first and total order sensitivity index of 0.6173 and 0.828, with a considerable gap to the second classified, k_{t0}, with sensitivity indices of 0.0840 and 0.2340, respectively. At last, $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ has a total order index of 0.0576. In this interval the number of parameters that have sensitivity indices above the threshold of 0.05 decreases considerably, and strangely the parameters k_{d1p} and $k_{\mathrm{a0p}}^{\mathrm{app}}$ lose their importance to the output in favour of k_{add} and k_{t0}.

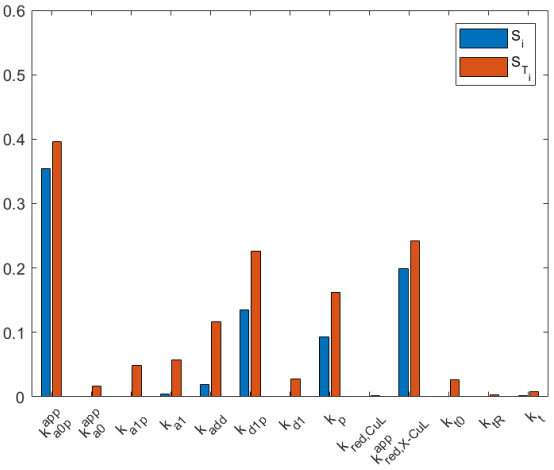
In this interval the biggest interactions between parameters are found in the 1st interval with a $1 - \sum_i S_i$ of 0.3311, followed by the 2nd and 3rd interval with 0.2904 and 0.240, respectively. This reinforces the ideia stated in the discussion of the monomer conversion results, that a bigger interval will induce a more interactive behaviour between the model inputs.

Comparing the results obtained from the two methods used for the global sensitivity analysis, some similarities can be found, such as the increase in the importance of k_{add} and k_{t0} as the size of the interval decreases, while for the parameters $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}$ the opposite effect is observed. Taking into account a ranking based on the $d_{n,\bar{n}}$, different results are found between the two methods.

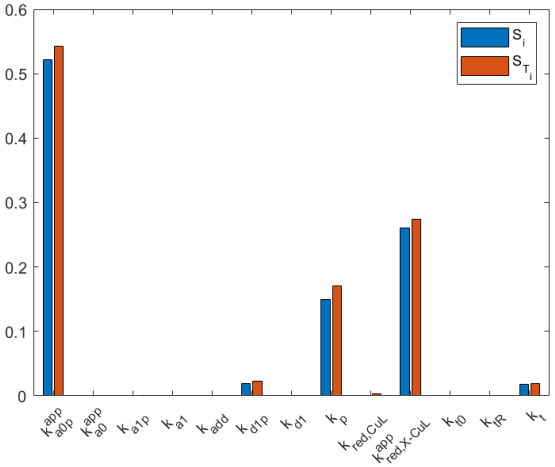
When using the MCF method, the importance of k_{add} and k_{t0} comes as an unexpected result. Thus, to investigate the reason for this result, the calculation of the sensitivity indices was repeated this time without considering the first experimental point. The results for each interval are represented in the Figures 5.36a, 5.36b and 5.36c.



(a) 1st interval results.



(b) 2nd interval results.



(c) 3rd interval results.

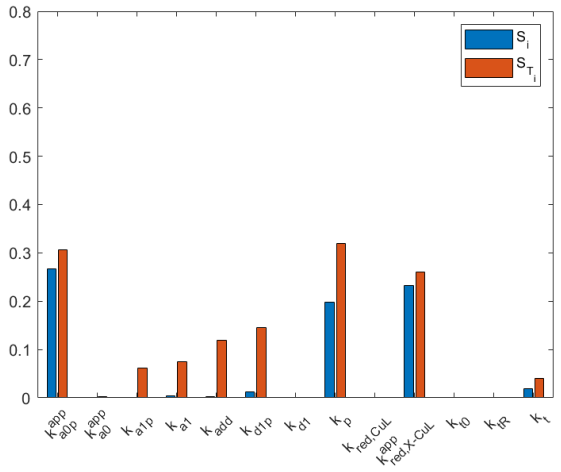
Figure 5.36: Sensitivity indices for the PDI output without the first experimental point.

From these new results it is possible to conclude that the importance of the parameters k_{add} and k_{t0} are highly linked to the first experimental point, due to the fact that the values of the sensitivity indices decrease substantially when this point is not considered in the calculation. Here, k_{add} is only considered important to the output in the 1st interval because of interaction effects. A plausible reason for the importance of k_{add} using the first experimental point is that, it is possible to notice that in at the time of this point, the model predicts a much higher PDI than the one that is determined experimental (Figure 4.1b). In this analysis the NRMSE is considered as the sensitivity output and, therefore, it is normal that a bigger fraction of this value is affected by the error at this point. Furthermore, because this experimental points corresponds to the initial stage of the reaction it is normal to be more affected by k_{add} and k_{t0}. These results reveal also the importance of the SARA kinetic constants ($k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$) to this output. This was also detected by the MCF method, where the SARA-ATRP kinetic constants are the two most important for the PDI in the 2nd and 3rd interval. Despite the difference between the two results, one major similarity can be found, which is the decreasing in the importance of

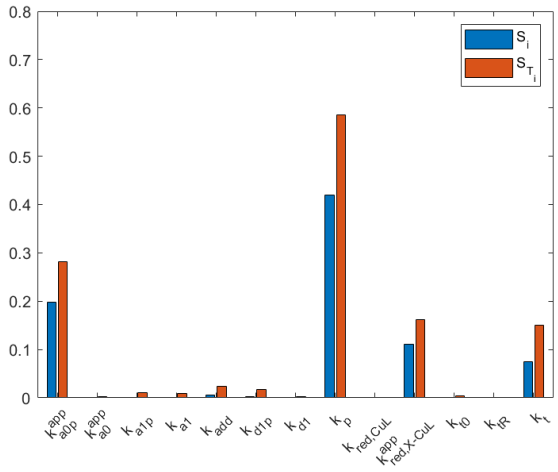
k_{d1p} as the range of the parameter variation is decreased. One major conclusion that is taken from these results, is that substantial differences can be found when considering different sets of experimental points for the PDI.

5.3.5 M_n output results

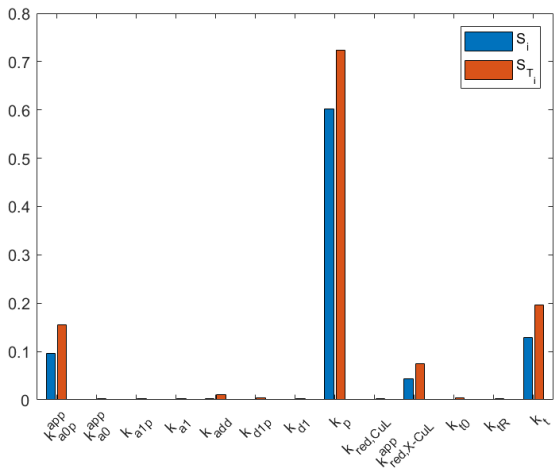
The results for the M_n output are presented in the Figures 5.37a, 5.37b and 5.37c for the 1st, 2nd and 3rd interval, respectively.



(a) 1st interval results.



(b) 2nd interval results.



(c) 3rd interval results.

Figure 5.37: Sensitivity indices for the M_n output.

It is possible to observe some differences specially between the 1st and the 2nd or 3rd interval, similarly to what was observed in the case of the PDI output. On the 1st interval k^{app}_{a0p} has the highest first order sensitivity index with a value of 0.2667 , followed by k_{p} with 0.3195. But if the total order sensitivity index is considered, then k_{p} is the most important parameter and $k^{\mathrm{app}}_{\mathrm{a0p}}$ the 2nd, with sensitivity of 0.3195 and 0.3056, respectively. In this interval, if the first order index is considered, then only one more parameter, $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$, is above the previously

defined threshold, with a value of 0.2667 and with a total order sensitivity index of 0.2605. All other parameters that are considered important are due to the interaction with the other kinetic constants. These parameters are k_{d1p}, k_{add}, k_{a1} and k_{a1p} with total order sensitivity indices of 0.1459, 0.1187, 0.0748 and 0.0612, respectively. In the 2nd interval there is a significant increase in the sensitivity indices of the parameters k_{t} and k_{p}. On contrary, there is a decrease of these indices for the $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$ and $k^{\mathrm{app}}_{\mathrm{a0p}}$. On this interval, k_{p} has the largest first order and total order indexes with values of 0.4200 and 0.5854. This parameter is followed by $k^{\mathrm{app}}_{\mathrm{a0p}}$, $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$, and k_{t} with values of the first order sensitivity index of 0.1973, 0.1100 and 0.0744, and total order sensitivity indices of 0.2813, 0.1623 and 0.1496, respectively. In the 3rd interval, k_{p} and k_{t} continue to increase their importance to the variability of the output. In this interval k_{t} overtakes the SARA-ATRP parameters in the ranking and moves into the 2nd place. In this interval k_{p} has the highest first and total order sensitivity indices with a values of 0.6021 and 0.7244, being followed by k_{t} and $k^{\mathrm{app}}_{\mathrm{a0p}}$ with a first order sensitivity indices of 0.1286 and 0.1965 and total sensitivity indices of 0.1965 and 0.1548. Finally, appears $k^{\mathrm{app}}_{\mathrm{red,X-CuL}}$, but it is only considered important, in this interval, due to the interactions effects, having a total order index of 0.0751.

As for the PDI output, the bigger fraction of interactions is found in the 1st interval with a $1 - \sum_{i=1}^{K} S_i$ of 0.2648, being the values for the 1st and 3rd interval of 0.1877 and 0.1279. These results reinforce the idea that a larger input space should lead to more interactions between the model inputs.

Comparing the results obtained from the two methods it is possible to observed similarities in the parameters that are considered critical or important to this output. One important aspect is that in the two methods the importance of the SARA kinetic constants decrease as the interval of variation decreases, in favour of k_{t}. One exception is k_{p}, which in the MCF method maintains its importance across the 3 intervals, while using the Sobol method the importance of this parameter increases through the intervals. One major difference that can be seen is the number of parameters that are considerable critical in the 1st interval using the MCF method results versus the ones that are considered important using the sensitivity indices. However, it should be taken into account that using the MCF method it is very difficult to observe the interaction structure of the model. For instance, the $d_{n,\overline{n}}$ used for the ranking of the parameter importance may not cover the interactions between inputs, and the parameters k_{d1p}, k_{add}, k_{a1} and k_{a1p} are only considered important due to the total order sensitivity index. Hence, their importance is mainly attributed to interactions that may not be detected by using the MCF method.

5.4 Methods comparison

Comparing the results obtained for the 3 outputs by the MCF and Sobol method some similarities can be found specially for the monomer conversion results, where basically the same conclusion can be drawn using either method. Nevertheless, some major differences can be visualized, such as: the ranking of the parameters in the PDI output; and the difference between the number of parameters classified as critical/important in the 1st interval results for the M_n output. These differences are normal and expected, because as stated in Wang and Solomatine (2019), these methods are based on different theories, concepts and principles, as explained in Chapter 3.

One factor that may be important to a modeler that intends to perform GSA in his/her model is the simplicity of the method and their ease of implementation. From a personal point of view, the MCF method was found to be very simple to understand and to implement in comparison to the Sobol method, which is based on more complex principles (Wang and Solomatine, 2019). Thus, the MCF method can be an easy tool to be applied for a preliminary GSA, giving extra-information by the analysis of the CDFs (i.e., input range where there is more probability of obtaining a behavioural output). Nevertheless, the Sobol method uses a more sophisticated and statistical apparatus, that results in a better estimation of the sensitivity index (Wang and Solomatine, 2019) and a better detection of inputs interactions. These interactions are not always detected by the MCF method, as verified in the 1st interval result for the M_n output. This has the implication that after using the MCF method it is necessary to use another GSA method to fully assess the input relevance (Saltelli et al., 2008). The reason for this, is that it is only possible to state that an input is unimportant if their single effect and interactions are fully analysed.

The computational effort was higher for the Sobol method. This method needed a larger sampling size (N) in order to converge in comparison to the MCF method. To make matters worse, the total number of model runs for the Sobol method, using the numerical procedure in Wu et al. (2012), is $N(2K + 1)$ while it is only N for the MCF method. This may make the use of the Sobol method very impractical, specially in complex models that require a larger CPU time to obtain the results.

Chapter 6

Parameter estimation problem

6.1 Problem formulation

The parameter estimation problem is formulated as an optimization problem, using the same objective function (J) as in Massicotte (2015) for the kinetic constants estimation of a ARGET-ATRP reaction. This objective function is defined in (6.1).

$$\begin{aligned} \min_{k} \quad & J(k) = \sum_{i=1}^{3} \frac{w_i}{J_{i\ ,max}} \sum_{j=1}^{10} \left(\frac{y_{i,j}}{y_{exp\ i,j}} - 1 \right)^2 \\ \text{subject to} \quad & k_L \leq k \leq k_U, \end{aligned} \tag{6.1}$$

Where $y_{i,j}$ is obtained by solving the kinetic model. The objective function (6.1) was written using a similar notation as in Amador et al. (2018), where k represents the vector of decision variables of the optimization problem. The $y_{exp\ i,j}$ and $y_{i,j}$ are the experimental and predicted values by the model for i outputs, respectively. In this case there are three decision variables which are the monomer conversion, PDI and M_n.The subscript j represents the time instant index. Here, $max(j) = 10$ because this is the number of measures taken from the experimental work. The subscripts U and L on k represent the upper and lower bound as in Amador et al. (2018).

The objective function uses the weight sum method. This method is characterized by using a objective function, which is the sum of the individual criteria, where each one is multiplied by a weight (w_i), that reflects the user preferences for each term. For instance, the author of Massicotte (2015) gave the same importance for all the terms (i.e., $w_i = 1$).

In this work different weights will be used for the three terms. First, it is necessary to define the ranking of the weights, that is, to define which output will have the highest and lowest correspondent weight. For this thesis, it was decided to choose the monomer conversion term to be the one with the highest weight. The reason for this option, was due to the fact that this

is the output where the model presents the worst prediction, with the highest NRMSE, being the main motivation to perform this parameter estimation problem. In addition, the method used to measure the monomer conversion (by nuclear magnetic resonance) is very precise, so the experimental data is very reliable. As for the term with the lowest weight, it was decided to choose the term related to M_n, because the experimental data appears to have some errors, namely in the 4th and 5th point (see Chapter 4). Therefore, the weights ranking will be 1st for the monomer conversion, 2nd for the PDI and for the 3rd for the M_n. The weights will be defined iteratively, by choosing w_i, between 0-1 and with $\sum_{n=1}^{i} w_i = 1$, that brings the NRMSE of the conversion closer to the one obtained in the PDI. One problem that should be accountefor choosing the correspondent weights, the 3 objective functions terms must be normalized to be on the same scale, in order to capture the users preferences. To normalize these terms, the same method as in Marler and Arora (2010) will be used, by dividing each of them by their maximum possible value ($J_{i,\ max}$).

To solve this parameter estimation problem the particle swarm optimization (PSO) algorithm was used. The main reason behind this decision is that the model used in this work is not constituted by explicit algebraic expressions, but rather by an ODE system. In this case, it is recommended the use of derivative-free algorithms like PSO (Rios and Sahinidis, 2013). The PSO algorithm was first introduced in Eberhart and Kennedy (1995). This algorithm was inspired by the swarm intelligence of some animals such as fishes and birds, because they have the capability of sharing information between their group, and that gives them a great survival advantage (For more information about the behavior of these swarms of animals and the reason to be transposed to an optimization algorithm the reader is referred to Almeida and Leite (2019)). These types of algorithms are called meta-heuristics, because they mimic successful optimization strategies found in nature. The word heuristic appears, because it is not guaranteed that the exact best solution will be found using these methods, but a good and useful approximation of the optimum is frequently obtained (Schwaab et al., 2008).

6.2 Optimization results

One important decision that needs to be made is which parameters are the decision variables for the optimization problem, taking into account the results of the global sensitivity analysis. This parameter estimation problem is mainly motivated by the bigger error calculated between the experimental points and the predicted conversion and M_n, compared to the one obtained for the PDI. Consequently, it is necessary to choose the kinetic parameters that affect the most the conversion and the M_n, but have a smaller effect on the PDI. The results from the global sensitivity analysis for the conversion and M_n, by the two methods used, are conclusive regarding the most important parameters for these outputs (see Chapter 4). These are k_{p}, that is by far the most important parameter, followed by k_{t}, and then $k_{\mathrm{a1p}}^{\mathrm{app}}$, and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$. These last two parameters are very important regarding their effect on the PDI, as demonstrated by the results

of the two methods (i.e., MCF and Sobol method). Taking this into account, it was decided to only estimate k_{p} and k_{t}.

The limits of the 3rd interval used in the global sensitivity analysis were chosen as the upper and lower bound for the optimization problem. One reason for this selection is that the values explored in this interval are all closer to the ones referenced in the Krys et al. (2017). Another reason, is that a larger interval allow to explore parameter values that may not be physically acceptable.

By defining the upper and lower bound, it was possible to calculate $J_{i,max}$. Various values of k_{p} and k_{t} were sampled, using the Sobol quasi-random sequence, within the interval defined by the upper and lower bound. Then the objective function is evaluated, with these different parameter values, and the maximum value of each term of the objective function was defined as $J_{i,max}$. The results obtained are in Table 6.1, for a total of 10000 objective function evaluations. As it can be seen, the maximum value of each objective function is quite different from each other. This. shows the importance of the normalization of the different objective terms in this specific case.

Table 6.1: $J_{i,max}$ with 10000 objective function evaluations.

O.F	$J_{i\ ,max}$
$J_{x,max}$	4.3827
$J_{PDI,max}$	0.0314
$J_{Mn,max}$	2.4321

As mentioned before, the values of w_i are adjusted by trial and error until the optimal solution gives the most similar NRMSE between the conversion and the PDI. This optimization problem was solved using the MATLAB® global optimization toolbox. This toolbox includes a built-in PSO algorithm, based on the original algorithm by Eberhart and Kennedy (1995) but with some modification by Pedersen (2010) and Mezura-Montes and Coello (2011).To illustrate the approach taken to calculate the weights, the diagram in Figure F.10 in the appendix F summarizes the procedure used. The algorithm described in Figure F.10 is initialized with $w_x = 0.5$, and terminates when the criteria are satisfied. The value of the weights obtained in presented in Table 6.2. As it can be seen, the criterion can only met if a large weight for the monomer conversion output is used than for the PDI and M_n terms in the objective function. Using the weights in Table 6.2 the values of k_{t} and k_{p} obtained are presented in Table 6.3.

6.3 Kinetic model simulations with the new parameters

The kinetic model was simulated again, but this time with the new parameter values obtained from the solution of the parameter estimation problem. In Figures 6.1a, 6.1b and 6.1c are

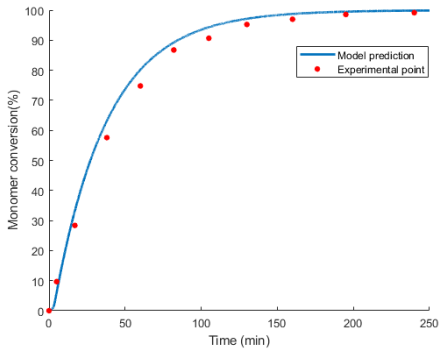
Table 6.2: w_i values.

output	w_i
w_x	0.90
w_{PDI}	0.07
w_{Mn}	0.03

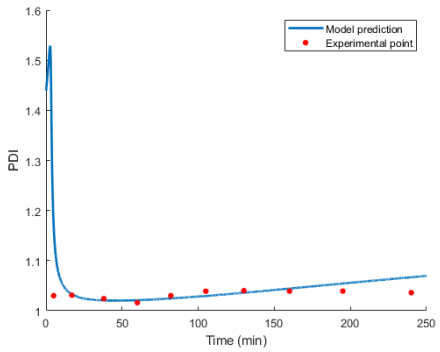
Table 6.3: Optimal parameter values.

Parameter	Value	Variation from the parameters in Krys et al. (2017)
k_t	3.90×10^8 $\mathrm{M^{-1}s^{-1}}$	59.2 %
k_p	1.28×10^4 $\mathrm{M^{-1}s^{-1}}$	-13.0 %

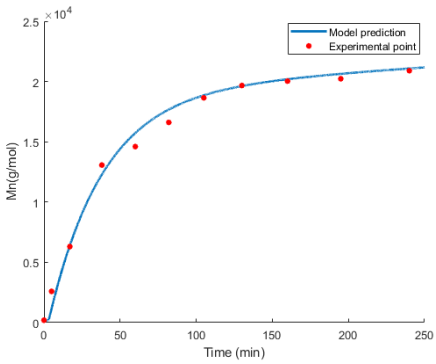
represented the profiles of the monomer conversion, PDI and M_n along the reaction time.



(a) Monomer conversion vs time.



(b) PDI vs time.



(c) M_n vs time.

Figure 6.1: Time profiles with the new parameter values.

It is clear that the use of the new parameter values decrease the error between the model prediction and the experimental data of the monomer conversion and M_n. As for the PDI profile no major difference can be visualized in comparison to the profile in Figure 4.1b. The NRMSE for these new simulations are given in Table 6.4.

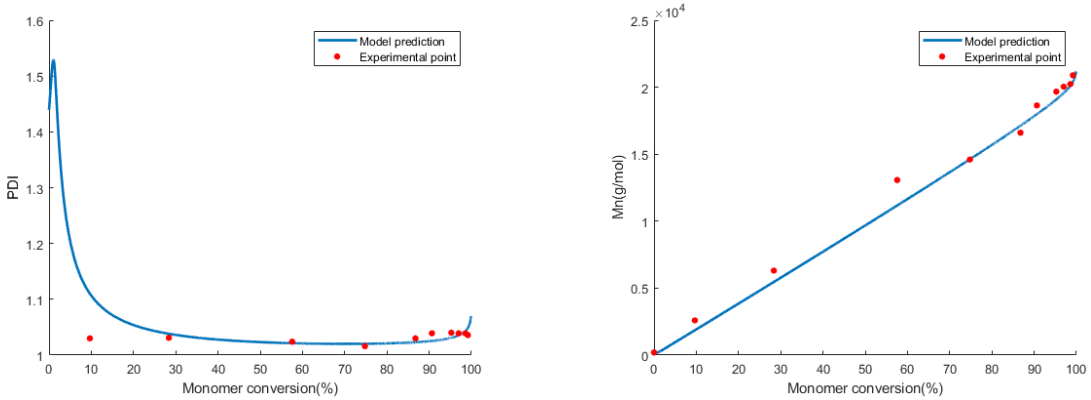
From the NRMSE values in Table 6.4 it follows that the model predictive capability improved with respect to the monomer conversion and the M_n. Regarding the PDI, a very small increase of

Table 6.4: NRMSE for the new simulations.

Output	NRMSE
Conversion	0.0479
PDI	0.0450
M_n	0.0481

about 0.0064 in NRMSE is observed. This increase can be due to the weight of the corresponding objective function term being much lower comparatively to the one used for the monomer conversion. The same does not occur with the M_n, because this output is related to the monomer conversion, as it was verified with the global sensitivity analysis study and by Equation 2.3.

The M_n and PDI profiles as a function of monomer conversion are presented in the Figure 6.2a and 6.2b Comparing these profiles to the ones obtained before the optimization, it is hard to find differences that can be visually detected. More specifically, in the PDI it is possible to see that the model prediction are closer to 6th and 7th experimental points, comparing to the model with the parameters from Krys et al. (2017).



(a) PDI vs monomer conversion.

(b) M_n vs monomer conversion.

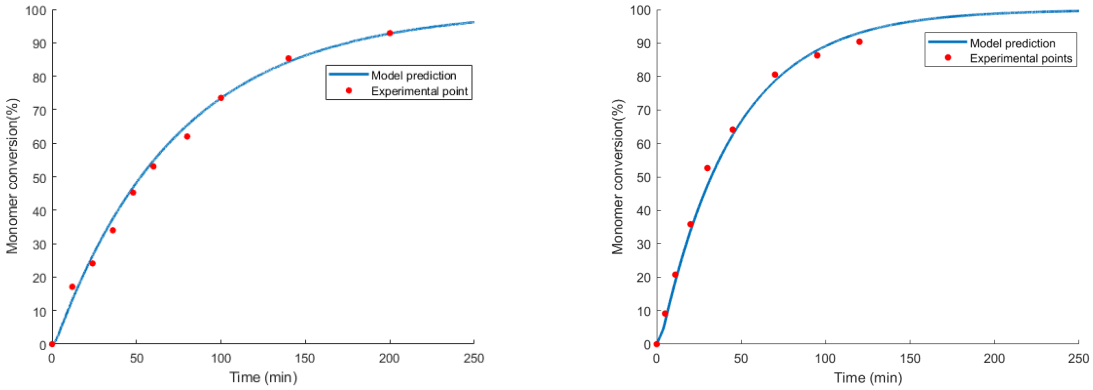
Figure 6.2: Monomer conversion profiles with the new parameter values.

Another problem that needs to be addressed is the risk of overfitting the model to this data. This occurs when the model can reproduce very well the training data (i.e., in this case the data used for the parameter estimation problem), but has a poor predictive capability. In order to assess this aspect, the model prediction was tested with 2 new data sets available in Krys et al. (2017). These 2 experiments were tested with different initial conditions (Table 6.5).

As it can be seen, the differences between these two new experiences and the the original are on the target DP_n in the case of the 2nd experiment, and in the initial concentration of the deactivator in the 3rd experiment. The monomer conversion, PDI and M_n profiles along the reaction time were obtained for the 2 new experiments and compared with the experimental points (6.3a to 6.5b). The experimental points of the 2nd and 3rd experience are presented in Appendix C in Table C.4 and C.5.

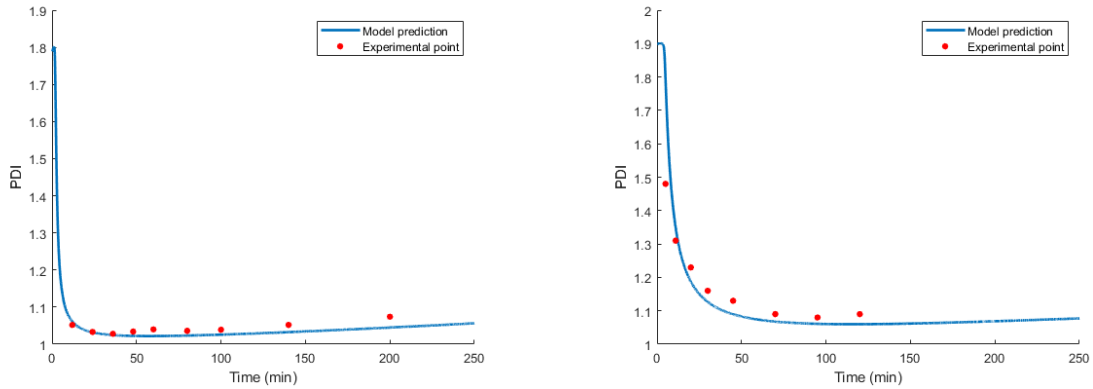
Table 6.5: Initial conditions for the 3 experiments.

Initial condition	1st experiment	2nd experiment	3rd experiment
Target DP	222.000	1000.000	222.000
$[M]_0$	7.400 M	7.400 M	7.400 M
$[RX]_0$	0.033 M	0.033 M	0.033 M
$[CuBr_2]_0$	0.003 M	0.003 M	0.0003 M
NºExperimental points	10	9	8



(a) 2nd experiment. **(b)** 3rd experiment.

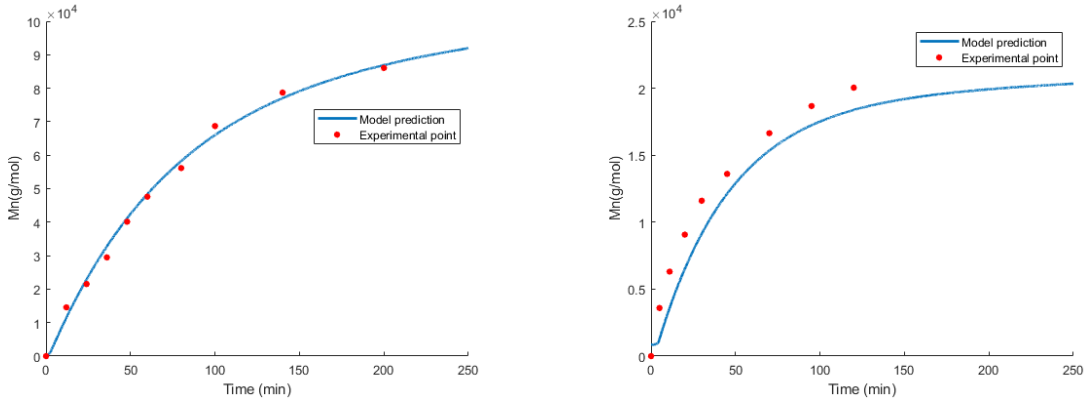
Figure 6.3: Monomer conversion vs time profile.



(a) 2nd experiment. **(b)** 3rd experiment.

Figure 6.4: PDI vs time profile.

From the observation of these profiles it can concluded that, in general, the model does a good prediction of the data in both experiments. From this preliminary analysis, the model is not over-fitting to the original experimental data. It should be noted that, in the future, is recommended the use of more data to verify this result. One clear exception is the M_n profile represented in Figure 6.5b, where it is possible to observe that the experimental value is always higher than the predicted one. One reason for this discrepancy could be the experimental measurement error of M_n by SEC. However this difference is observed for all the points. Note, that in the first instances of the M_n profile, it is possible to observe that the M_n remains constant for some time



(a) 2nd experiment. **(b)** 3rd experiment.

Figure 6.5: M_n vs time profile.

(this can also be visible in the PDI profile), which delays the predicted growing of the M_n. This can explain the difference between experimental and predicted values. The simulations profiles in Krys et al. (2017) also show a similar deviation and a initial constant behaviour of M_n at the beginning of the profile.

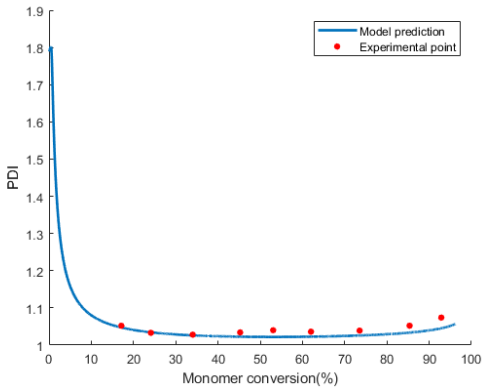
This could indicate that for a smaller concentration of deactivator (this is the case in which the concentration of deactivator is decreased by a factor of 10) it is necessary to consider additional reactions in the formulation of the kinetic model . Nevertheless, further experiments need to be made and no definitive conclusions should be drawn from these tests alone. Table 6.6 presents the NRMSE results calculated for the 3 outputs of the 2 experiences The high value of NRMSE

Table 6.6: NRMSE obtain for the 3 experiments.

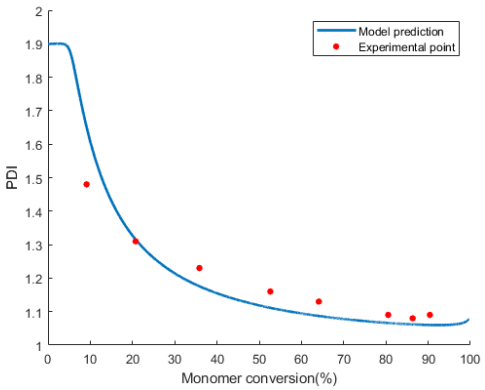
	NRMSE		
	1st experience	2nd experience	3rd experience
Conversion	0.0479	0.0436	0.0485
M_n	0.0450	0.0425	0.0972
PDI	0.0481	0.0147	0.1659

obtained in the simulation of the 3rd experiment came as a surprise, because in Figure 6.4b, it seems that the model predicts well the experimental data. A possible explanation for this error is the first experimental point, where the model predicts a value of 1.7965 and the experimental point is 1.486. If this point is neglected the NRMSE decreases to 0.0290. The larger NRMSE also confirms the model lack of ability in predicting the M_n for this experiment. Nevertheless, good results can be found for 2nd experiment, with smaller NRMSE values. The biggest difference is in the PDI where for 1st experiment the error is 0.0450 and for the second only 0.0147. Again this is explained by the fact that in the 1st experiment the model does not predict well the 1st experimental point. Note that this does not occur with the simulation for the 2nd experiment for which the NRMSE is small. The profiles for the PDI and M_n in function of the monomer

conversion for the 2nd and 3rd experiments are given in Figures 6.6a to 6.7b.

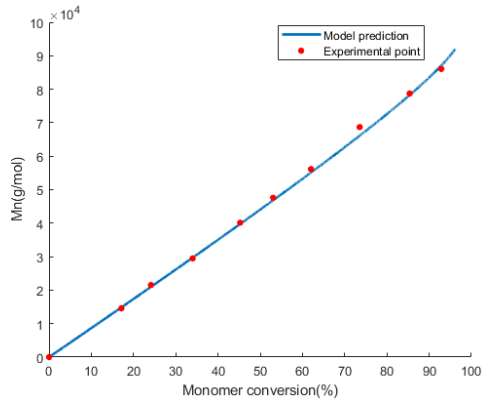


(a) 2nd experiment.

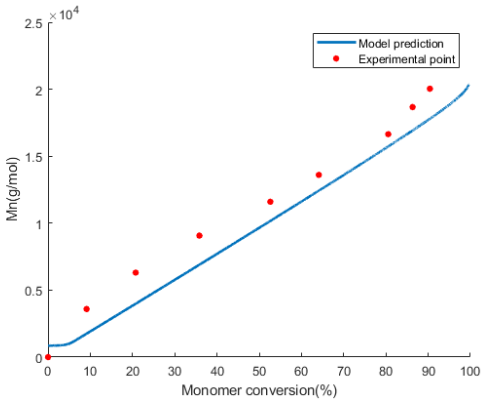


(b) 3rd experiment.

Figure 6.6: PDI vs monomer conversion profile.



(a) 2nd experiment.



(b) 3rd experiment.

Figure 6.7: M_n vs monomer conversion profile.

These profiles show a good agreement with the experimental points of the 1st experiment. The same can not be said for the 3rd experiment where considerable deviations between the experimental points and the model predictions are observed. On the PDI vs monomer conversion profile these errors are more pronounced, in the 3rd, 4th and 5th experimental point. Nevertheless, the model has a good prediction of the PDI at high monomer conversions. As for the M_n profile as in the case of M_n vs time profile, the error is systematic along the monomer conversion. This result was also reported in Krys et al. (2017).

From these results it is possible to conclude that the kinetic model, in general, predicts quite well the experimental data. In the future, further developments can be made to formulate a more accurate model. This can be achieved by considering, for example, the diffusion limitations that are more pronounced at high monomer conversion (Zhou and Luo, 2014), reactor volume variation Vieira et al. (2015), chain-length dependence of the kinetic parameters (Lyra et al., 2019), etc. There is a wide range of possible improvements that improve the model. But it is important to test the model with more data, because this is the only way to be certain that the model is good enough and is not affected by experimental and modeling errors. The sentence "on god we trust all others must bring data" said by the statistician Edward Deming, must be taken seriously.

Chapter 7

Conclusions and Future work

7.1 Conclusions

The main goal of this thesis was to perform a global sensitivity analysis (GSA) on an already validated kinetic model of a SARA-ATRP system proposed in Krys et al. (2017). First of all, the kinetic model formulation for the SARA-ATRP system was reviewed. Then the model was simulated with the kinetic parameters given in Krys et al. (2017), and the results were compared with experimental data. From this analysis it was concluded that the model predictions presented some deviations from the experimental points, specially in the case of the monomer conversion and the M_n outputs. Therefore, to improve the model predictive ability a parameter estimation problem was formulated, to obtain new kinetic parameter values. In order to do this, only the model important parameters identified through the GSA study, were considered as the decision variables of the parameter estimation problem.

From the application of the MCF method it was concluded that the specification of the number of model simulations (N) can not be done based on the studying of the convergence of the p-value, because of the commonly know p-value problem (the p-value decreases as the number of samples increases). Another alternative criteria, used in this work, is the $d_{n,\bar{n}}$. The MCF method identifies the most important parameters in order to reduce the current model error in the three model outputs. For the monomer conversion, the results were very similar for the 3 input space intervals studied. It was always k_{p} the kinetic parameter with the highest rank, followed by k_{t}, $k_{\mathrm{a0p}}^{\mathrm{app}}$ and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$. The same can not be said for the PDI result. Here, significant differences in the parameter ranking were obtained for each input space interval under study, and in the number of parameters that receive the critical classification. Here, the parameters that always receive the critical classification are $k_{\mathrm{a0p}}^{\mathrm{app}}$, $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$, k_{d1p}, k_{add} and k_{p}. The M_n output leads to similar results to the ones obtained for the monomer conversion. These two outputs are related by Equation 2.3 and therefore these results are expected. Then, the parameters that are classified as critical/important are the same as the ones in the monomer conversion results in

the 3 intervals. A difference between the ranking order of k_{t}, k_{a0p}^{app} and $k_{\mathrm{red,X-CuL}}^{app}$ is observed along the different intervals studied.

As for the Sobol method results, two numerical procedures to calculate the first and total order sensitivity indices were compared: the method presented in Saltelli et al. (2008) and a new derivation made by Wu et al. (2012). In general, using the method of Wu et al. (2012) the values of the true sensitivity indices are closer when the N is smaller. In addition, this method also decreases the occurrence of numerical errors (i.e., negative sensitivity indices) in comparison to the numerical procedure in Saltelli et al. (2008). For the monomer conversion, the results are similar to the ones obtained using the MCF method. Using the threshold of 0.05, only k_{p}, k_{t} and $k_{\mathrm{a0p}}^{\mathrm{app}}$ (being this last only important due to interaction effects) are considered important in all the intervals. Similarly to the MCF method, there are significant differences between the PDI results in the three input space intervals studied. Only k_{add}, k_{t0} and $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$ are considered important in all the three intervals. This result, of k_{add} and k_{t0} being classified as important was intriguing because these are kinetic constants of the reactions that occur in the early stages of the polymerization. Then, the Sobol method was performed considering the NRMSE without the first experimental point. The results showed a considerable decrease in the importance of k_{add} and k_{t0} to the PDI output, concluding that their importance is linked to the first experimental point. For the M_n output the kinetic constants that obtained the important classification in all the 3 input space intervals study were $k_{\mathrm{a0p}}^{\mathrm{app}}$, $k_{\mathrm{red,X-CuL}}^{\mathrm{app}}$, k_{p} and k_{t}. These results are very similar to the ones obtained for the monomer conversion, in particular for the 3rd input space interval.

From this work, it was conclude that the MCF method is a simpler and easier method to understand and to implement, with a low computational cost, in comparison to the Sobol method. Nevertheless, due to the fact that the interactions are not always accounted in the $d_{n,\overline{n}}$ value, it is necessary to use another GSA method (e.g., such as Sobol), to fully conclude the sensitivity analysis.

New values for the parameters k_{p} and k_{t} were obtained by solving a kinetic estimation problem. The simulations results show an increase in the model ability to predict the monomer conversion and the M_n. To prevent the risk of overfitting the model to the original set of data, the model was tested with other sets of experimenal data. The model show a relatively good predictive ability. One exception is in predicting the M_n for the 3rd experience. This is also observed in the simulations results of Krys et al. (2017). Nevertheless, there is a need to test the model with more experimental data to make more assertive conclusions.

7.2 Future work

Regarding the kinetic model formulation, further improvements can be made, such as:

- consideration of diffusion limitations and comparison of the results obtained with the ones of this work. A similar study was done in Zhou and Luo (2014);
- consideration of reaction medium volume variation;
- collecting more experimental data to use in new kinetic parameters estimation problems in order to improve the model predictive ability;
- another modeling alternative could be the development of a hybrid model, which involve the integration of mechanistic and a data-driven model. The data could provide information about aspects that are not present in the deterministic model (Sansana et al., 2021);
- estimation of the Arrhenius equation constants for the model kinetic rate constants, to formulate the model with the temperature dynamics.

Also, the kinetic model can be used in the future in several studies, such as:

- the application of the Kalman filter to estimate the concentration of $SO_2^{\bullet}$ (see reaction scheme 4.1) along the reaction time, which is not possible to measure experimentally.
- the scale-up of the SARA-ATRP reactor system;
- the study of control strategies, including advanced process control.

In relation to the global sensitivity analysis, further works can be done, such as:

- to consider correlations between the inputs. In the GSA realized in this thesis, the inputs correlations were not considered. In fact, to use the Sobol method it is necessary to assume that the parameters are independent of each other (see Appendix A). The consideration of correlation in the inputs affects the outcome of the sensitivity analysis and therefore the conclusions drawn form this study can be different. In Xie et al. (2019) two types of sensitivity analysis in two different biochemical reactions. One of them considered the inputs independent and the other one comprises the correlations. Comparing the results of the two cases, significant differences in the ranking and in the values of the sensitivity indices were detected. In the future, a similar study must be performed with this kinetic model, in order to conclude and investigate the influence of input correlation and their effect in the final GSA results;
- test other sensitivity analysis method such as the Fourier amplitude sensitivity test and the elementary effects method, to compare with the results obtained in this thesis;
- plan new laboratory experiences in order to obtaine more accurate values of the kinetic parameters but with special focus on the parameters, that were identified as important in the Global sensitivity analysis study;

- perform a sensitivity analysis considering the kinetic model initial conditions as the inputs.
- perform a study applying the Fisher matrix method to determine the most important parameters and to asses if the problem is identifiable, and to forecast the precision of experiments in the context of design of experiments.

Chapter 8

Bibliography

Abreu, C. M., Fu, L., Carmali, S., Serra, A. C., Matyjaszewski, K., and Coelho, J. F. (2017). Aqueous SARA-ATRP using inorganic sulfites. *Polymer chemistry*, 8(2):375–387.

Abreu, C. M., Mendonça, P. V., Serra, A. C., Popov, A. V., Matyjaszewski, K., Guliashvili, T., and Coelho, J. F. (2012). Inorganic sulfites: Efficient reducing agents and supplemental activators for atom transfer radical polymerization. *Acs Macro Letters*, 1(11):1308–1311.

Abreu, C. M. R. (2018). *Developments in Reversible Deactivation Radical Polymerization: New Ecofriendly Catalytic Systems and Vinyl Chloride (co) Polymerization Methods*. PhD thesis, Universidade de Coimbra.

Al-Harthi, M., Soares, J. B., and Simon, L. C. (2006). Mathematical modeling of atom-transfer radical polymerization using bifunctional initiators. *Macromolecular theory and simulations*, 15(3):198–214.

Almeida, B. S. G. and Leite, V. C. (2019). Particle swarm optimization: A powerful technique for solving engineering problems. *Swarm Intelligence-Recent Advances, New Perspectives and Applications*.

Amador, A., Fernandes, F. P., Santos, L. O., Romanenko, A., and Rocha, A. M. A. (2018). Parameter estimation of the kinetic α-pinene isomerization model using the MCSFilter algorithm. *International Conference on Computational Science and Its Applications*, pages 624–636.

Asteasuain, M. (2018). Deterministic approaches for simulation of nitroxide-mediated radical polymerization. *International Journal of Polymer Science*, 2018.

Azzini, I., Listorti, G., Mara, T., and Rosati, R. (2020). Uncertainty and sensitivity analysis for policy decision making. *An Introductory Guide. Joint Research Centre, European Commission, Luxembourg, Technical report*.

Bamford, C., Dyson, R., and Eastmond, G. (1969). Network formation iv. the nature of the termination reaction in free-radical polymerization. *Polymer*, 10:885–899.

Bhonsale, S., Muñoz López, C. A., and Van Impe, J. (2019). Global sensitivity analysis of a spray drying process. *Processes*, 7(9):562.

Bolker, B. M., Grasselli, M. R., and Holmes, E. (2021). Sensitivity analysis of an integrated climate-economic model. *SIAM Journal on Financial Mathematics*, 12(2):SC–44.

Boyer, C., Corrigan, N. A., Jung, K., Nguyen, D., Nguyen, T.-K., Adnan, N. N. M., Oliver, S., Shanmugam, S., and Yeow, J. (2016). Copper-mediated living radical polymerization (atom transfer radical polymerization and copper (0) mediated polymerization): from fundamentals to bioapplications. *Chemical reviews*, 116(4):1803–1949.

Campolongo, F. and Braddock, R. (1999). Sensitivity analysis of the image greenhouse model. *Environmental modelling & software*, 14(4):275–282.

Cariboni, J., Gatelli, D., Liska, R., and Saltelli, A. (2007). The role of sensitivity analysis in ecological modelling. *Ecological modelling*, 203(1-2):167–182.

Coessens, V., Pintauer, T., and Matyjaszewski, K. (2001). Functional polymers by atom transfer radical polymerization. *Progress in polymer science*, 26(3):337–377.

Eberhart, R. and Kennedy, J. (1995). A new optimizer using particle swarm theory. pages 39–43.

Fischer, H. (1997). The persistent radical effect in "living" radical polymerization. *Macromolecules*, 30(19):5666–5672.

Fischer, H. (1999). The persistent radical effect in controlled radical polymerizations. *Journal of Polymer Science Part A: Polymer Chemistry*, 37(13):1885–1901.

Fischer, H. (2001). The persistent radical effect: a principle for selective radical reactions and living radical polymerizations. *Chemical reviews*, 101(12):3581–3610.

Flores-Tlacuahuac, A., Saldívar-Guerra, E., and Guerrero-Santos, R. (2003). Dynamic modelling, nonlinear parameter fitting and sensitivity analysis of a living free-radical polymerization reactor. In *Computer Aided Chemical Engineering*, volume 16, pages 21–39. Elsevier.

Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361.

Góis, J. R., Konkolewic, D., Popov, A. V., Guliashvili, T., Matyjaszewski, K., Serra, A. C., and Coelho, J. F. (2014). Improvement of the control over SARA ATRP of 2-(diisopropylamino) ethyl methacrylate by slow and continuous addition of sodium dithionite. *Polymer Chemistry*, 5(16):4617–4626.

Guliashvili, T., Mendonça, P. V., Serra, A. C., Popov, A. V., and Coelho, J. F. (2012). Copper-mediated controlled/"living" radical polymerization in polar solvents: Insights into some relevant mechanistic aspects. *Chemistry–A European Journal*, 18(15):4607–4612.

Homma, T. and Saltelli, A. (1996). Importance measures in global sensitivity analysis of non-linear models. *Reliability Engineering & System Safety*, 52(1):1–17.

Huang, Y., Zhou, Y.-N., Zhang, Q., and Luo, Z.-H. (2013). Modeling of the atom transfer radical polymerization for preparing novel fluorosilicone diblock copolymers in a semi-batch reactor. *Journal of Applied Polymer Science*, 130(5):3473–3481.

Hungenberg, K.-D. and Wulkow, M. (2018). *Modeling and Simulation in Polymer Reaction Engineering: A Modular Approach*. John Wiley & Sons.

Kato, M., Kamigaito, M., Sawamoto, M., and Higashimura, T. (1995). Polymerization of methyl methacrylate with the carbon tetrachloride/dichlorotris-(triphenylphosphine) ruthenium (ii)/methylaluminum bis (2, 6-di-tert-butylphenoxide) initiating system: possibility of living radical polymerization. *Macromolecules*, 28(5):1721–1723.

Konkolewicz, D., Wang, Y., Krys, P., Zhong, M., Isse, A. A., Gennaro, A., and Matyjaszewski, K. (2014). Sara atrp or set-lrp. end of controversy? *Polymer Chemistry*, 5(15):4396–4417.

Krys, P., Fantin, M., Mendonça, P. V., Abreu, C. M., Guliashvili, T., Rosa, J., Santos, L. O., Serra, A. C., Matyjaszewski, K., and Coelho, J. F. (2017). Mechanism of supplemental activator and reducing agent atom transfer radical polymerization mediated by inorganic sulfites: experimental measurements and kinetic simulations. *Polymer chemistry*, 8(42):6506–6519.

Krys, P. and Matyjaszewski, K. (2017). Kinetics of atom transfer radical polymerization. *European Polymer Journal*, 89:482–523.

Lin, M., Lucas Jr, H. C., and Shmueli, G. (2013). Research commentary—too big to fail: large samples and the p-value problem. *Information Systems Research*, 24(4):906–917.

Lyra, E. P. et al. (2019). Strategies for reducing catalyst concentration in atom transfer radical polymerization: a kinetic study approach by mathematical modeling and simulation= estratégias para a redução da concentração de catalisadores na polimerização radicalar por transferência de átomo: uma abordagem de estudo cinético por modelagem matemática e simulação.

Marler, R. T. and Arora, J. S. (2010). The weighted sum method for multi-objective optimization: new insights. *Structural and multidisciplinary optimization*, 41(6):853–862.

Massicotte, E. (2015). Modelling and experimental investigation of arget/aget atom transfer radical polymerization systems. Master's thesis.

Mastan, E., Li, X., and Zhu, S. (2015). Modeling and theoretical development in controlled radical polymerization. *Progress in Polymer Science*, 45:71–101.

Mastan, E. and Zhu, S. (2015). Method of moments: A versatile tool for deterministic modeling of polymerization kinetics. *European Polymer Journal*, 68:139–160.

Matyjaszewski, K. (1998). *Overview: fundamentals of controlled/living radical polymerization. In Controlled Radical Polymerization*. ACS Publications.

Matyjaszewski, K. (2000). *Comparison and classification of controlled/living radical polymerizations. In Controlled/Living Radical Polymerization*. ACS Publications.

Matyjaszewski, K. (2012). Atom transfer radical polymerization (atrp): current status and future perspectives. *Macromolecules*, 45(10):4015–4039.

Matyjaszewski, K. and Davis, T. P. (2003). *Handbook of radical polymerization*. John Wiley & Sons.

Mezura-Montes, E. and Coello, C. A. C. (2011). Constraint-handling in nature-inspired numerical optimization: past, present and future. *Swarm and Evolutionary Computation*, 1(4):173–194.

Mishra, V. and Kumar, R. (2012). Living radical polymerization: A review. *J. Sci. Res*, 56:141–176.

Moler, C. (2003). Stiff differential equations. `https://www.mathworks.com/company/newsletters/articles/stiff-differential-equations.html`. Accessed: 2021-06-10.

Monteiro, M. J., Guliashvili, T., and Percec, V. (2007). Kinetic simulation of single electron transfer–living radical polymerization of methyl acrylate at 25° c. *Journal of Polymer Science Part A: Polymer Chemistry*, 45(10):1835–1847.

Ochoa, M. P., Estrada, V., Di Maggio, J., and Hoch, P. M. (2016). Dynamic global sensitivity analysis in bioreactor networks for bioethanol production. *Bioresource technology*, 200:666–679.

Odian, G. (2004). *Principles of polymerization*. John Wiley & Sons.

Pedersen, M. E. H. (2010). Good parameters for particle swarm optimization. *Hvass Lab., Copenhagen, Denmark, Tech. Rep. HL1001*, pages 1551–3203.

Percec, V., Guliashvili, T., Ladislaw, J. S., Wistrand, A., Stjerndahl, A., Sienkowska, M. J., Monteiro, M. J., and Sahoo, S. (2006). Ultrafast synthesis of ultrahigh molar mass polymers by metal-catalyzed living radical polymerization of acrylates, methacrylates, and vinyl chloride mediated by set at 25 c. *Journal of the American Chemical Society*, 128(43):14156–14165.

Qian, G. and Mahdi, A. (2020). Sensitivity analysis methods in the biomedical sciences. *Mathematical biosciences*, 323:108306.

Ribelli, T. G., Augustine, K. F., Fantin, M., Krys, P., Poli, R., and Matyjaszewski, K. (2017). Disproportionation or combination? the termination of acrylate radicals in ATRP. *Macromolecules*, 50(20):7920–7929.

Rios, L. M. and Sahinidis, N. V. (2013). Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization*, 56(3):1247–1293.

Salas, S. D., Brandão, A. L., Soares, J. B., and Romagnoli, J. A. (2019). Data-driven estimation of significant kinetic parameters applied to the synthesis of polyolefins. *Processes*, 7(5):309.

Saltelli, A. (2002a). Making best use of model evaluations to compute sensitivity indices. *Computer physics communications*, 145(2):280–297.

Saltelli, A. (2002b). Sensitivity analysis for importance assessment. *Risk analysis*, 22(3):579–590.

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., and Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental modelling & software*, 114:29–39.

Saltelli, A. and Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, 25(12):1508–1517.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global sensitivity analysis: the primer*. John Wiley & Sons.

Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models*, volume 1. Wiley Online Library.

Sansana, J., Joswiak, M. N., Castillo, I., Wang, Z., Rendall, R., Chiang, L. H., and Reis, M. S. (2021). Recent trends on hybrid modeling for industry 4.0. *Computers & Chemical Engineering*, page 107365.

Santos, M. R. E. (2020). *Atom transfer radical polymerization: study of ligands and preparation of block copolymers with controlled structures for application in high-performance coatings*. PhD thesis, 00500:: Universidade de Coimbra.

Schwaab, M., Biscaia Jr, E. C., Monteiro, J. L., and Pinto, J. C. (2008). Nonlinear parameter estimation through particle swarm optimization. *Chemical Engineering Science*, 63(6):1542–1552.

Shen, Y., Tang, H., and Ding, S. (2004). Catalyst separation in atom transfer radical polymerization. *Progress in polymer science*, 29(10):1053–1078.

Sobol, I. M. (1990). On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe modelirovanie*, 2(1):112–118.

Thiese, M. S., Ronna, B., and Ott, U. (2016). P value interpretations and considerations. *Journal of thoracic disease*, 8(9):E928.

Thomas, I. M. and Kiparissides, C. (1984). Sensitivity analysis of a batch polymerization reactor. *Journal of applied polymer science*, 29(6):2195–2204.

Vieira, R. P. et al. (2015). *Simulação e otimização da síntese de poliestireno funcionalizado via polimerização radicalar por transferência de átomo*. PhD thesis, Universidade Estadual de Campinas (UNICAMP).

Vieira, R. P. and Lona, L. M. F. (2016). Simulation of temperature effect on the structure control of polystyrene obtained by atom-transfer radical polymerization. *Polímeros*, 26:313–319.

Wang, A. and Solomatine, D. P. (2019). Practical experience of sensitivity analysis: Comparing six methods, on three hydrological models, with three performance criteria. *Water*, 11(5):1062.

Wang, J.-S. and Matyjaszewski, K. (1995). Controlled/" living" radical polymerization. atom transfer radical polymerization in the presence of transition-metal complexes. *Journal of the American Chemical Society*, 117(20):5614–5615.

Wu, Q.-L., Cournède, P.-H., and Mathieu, A. (2012). An efficient computational method for global sensitivity analysis and its application to tree growth modelling. *Reliability Engineering & System Safety*, 107:35–43.

Wu, Z., Yang, X., and Sun, X. (2017). Application of monte carlo filtering method in regional sensitivity analysis of aashtoware pavement me design. *Journal of traffic and transportation engineering (English edition)*, 4(2):185–197.

Xie, X., Schenkendorf, R., and Krewer, U. (2019). The effect of correlated kinetic parameters on (bio) chemical reaction networks. *Chemie Ingenieur Technik*, 91(5):632–636.

Yao, X. (2015). Sensitivity analysis for biorefineries. Master thesis report. Wageningen university.

Young, R. J. and Lovell, P. A. (2011). *Introduction to polymers*. CRC press.

Zhang, Z., Gul, R., and Zeb, A. (2021). Global sensitivity analysis of covid-19 mathematical model. *Alexandria Engineering Journal*, 60(1):565–572.

Zhou, Y.-N. and Luo, Z.-H. (2014). Copper (0)-mediated reversible-deactivation radical polymerization: kinetics insight and experimental study. *Macromolecules*, 47(18):6218–6229.

Zi, Z., Zheng, Y., Rundell, A. E., and Klipp, E. (2008). Sbml-sat: a systems biology markup language (sbml) based sensitivity analysis tool. *BMC bioinformatics*, 9(1):1–14.

Zubov, A. and Sin, G. (2018). Multiscale modeling of poly (lactic acid) production: From reaction conditions to rheology of polymer melt. *Chemical Engineering Journal*, 336:361–375.

Appendices

A Sobol method demonstrations

A.1 Sobol demonstrations

Ilya Sobol proposed a way of calculating the sensitivity index using a Monte-Carlo based approach, evaluating the function $Y = f(X_1, X_2, ..., X_i, ..., X_K)$ multiple times with different input values. He proposed a decomposition of the response f, considering into a set of functions of increasing dimensions in which all the individual terms are square integrable in their domain and that the inputs are all mutually independent (Ochoa et al., 2016; Saltelli et al., 2008; Wu et al., 2012). In (1) the indices i,j and k represent different model inputs.

$$f = f_0 + \sum_i f_i(X_i) + \sum_i \sum_{j>i} f_{i,j}(X_i, X_j) + ... + f_{1,2,..,k}(X_i, X_j, ..., X_k) \tag{1}$$

Sobol proved that if all the individual terms of (1) have zero mean (Equation 2) then all the terms are orthogonal in pairs (Equation 3).

$$\int f(X_i) dX_i = 0 \tag{2}$$

$$\int f(X_i) f(X_j) dX_i dX_j = 0 \tag{3}$$

This orthogonal characteristic enables that these terms can be calculated using the conditional expectations of the model output (Y) as it is demonstrated in the Equations 4 and 5 (Saltelli, 2002a; Saltelli et al., 2008; Ochoa et al., 2016).

$$E(Y) = \int f(X) dX = f_0 \tag{4}$$

$$E(Y|X_i) = \int f(X) \prod_{k \neq i} dX_k = f_0 + f_i(X_i) \tag{5}$$

By square integrating each one of the terms of (1) except X_i, it is possible to obtain the decomposition of the variance, the so called the ANOVA-HMDR decomposition (Equation 6) (Ochoa et al., 2016; Cariboni et al., 2007; Saltelli et al., 2008; Wu et al., 2012).

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>i} V_{i,j} + ... + V_{1,2,...,k} \tag{6}$$

Here, $\sum_i V_i$ represents the main contribution to the total variance of only one input (X_i) and

$\sum_i \sum_{j>i} V_{i,j} + ... + V_{1,2,...,k}$ is all the decomposed parts that the input X_i appears, from two-elements ($V_{i,j}$) up to k-elements group ($V_{1,2,3,..,k}$) (Yao, 2015). As one can imagine, this last element indicates the contribution of inputs interaction to the total variance of the model output (Yao, 2015). Dividing 6 by $V(Y)$ and taking into account the definition of sensitivity index, (3.3), Equation 7 is obtained.

$$1 = \sum_i S_i + \sum_i \sum_{j>i} S_{i,j} + \sum_i \sum_{j>i} \sum_{l>j} S_{i,j,l} + ... + S_{1,2,3...k} \tag{7}$$

A.2 Equation 3.14 deduction

The expect value of the model output ($E(Y)$) and $V(Y)$ can be defined by the multidimensional integrals (8) and (9) (Saltelli, 2002a).

$$E(Y) = \int \int ... \int f(X_1, X_2, ..., X_K) \prod_{i=1}^{k} p_i X_i dX_i \tag{8}$$

$$V(Y) = \int \int ... \int f^2(X_1, X_2, ..., X_K) \prod_{i=1}^{k} p_i X_i dX_i - E^2(Y) \tag{9}$$

As demonstrated in Saltelli (2002a), the numerator (Equation 3.3) in the the first order sensitivity index equation ($V_{X_i}(E_{X_i}(Y|X_{\sim i}))$ can be calculated by solving the integral in (10).

$$V_{X_i}(E_{X_i}(Y|X_i)) = V(Y) - E_{X_i}(V(Y|X_i)) = \int E^2_{X_i}(Y|X_i = \tilde{X_i}) p_i X_i dX_i - E^2(Y) \tag{10}$$

In (10) the term $\int E^2_{X_i}(Y|X_i = \tilde{X_i}) p_i X_i dX_i$ can be substituted by U_i, as follows (Saltelli, 2002a):

$$U_i = \int E^2_{X_i}(Y|X_i = \tilde{X_i}) p_i X_i dX_i \tag{11}$$

Thus, from (3.3)is obtained:

$$S_i = \frac{Ui - E^2(Y)}{V(Y)} \tag{12}$$

B Moments differential equations development

For the $P_n^{\bullet}$ chains the differential terms corresponding to the balances $\frac{d[P_1^{\bullet}]}{dt}$ and $\frac{d\sum_{n=2}^{\infty} n^k [P_n^{\bullet}]}{dt}$ of (2.11), can be written as follows:

$$\frac{\mathrm{d}[\mathrm{P_1}^\bullet]}{\mathrm{d}t} = k_{\mathrm{a0p}}^{\mathrm{app}}[\mathrm{P_1X}] + k_{\mathrm{a1p}}[\mathrm{Cu^I/L}][\mathrm{P_1X}] - k_{\mathrm{d1p}}[\mathrm{Cu^{II}X/L}][\mathrm{P_1}^\bullet] + k_{\mathrm{add}}[\mathrm{R}^\bullet][\mathrm{M}] - k_{\mathrm{p}}[\mathrm{M}][\mathrm{P_1}^\bullet] - k_{\mathrm{tR}}[\mathrm{R}^\bullet][\mathrm{P_1}^\bullet] - (k_{\mathrm{tc}} + k_{\mathrm{td}})[\mathrm{P_1}^\bullet] \sum_{n=1}^{\infty}[\mathrm{P}_n{}^\bullet] \tag{13a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\sum_{n=2}^{\infty} n^k[\mathrm{P}_n{}^\bullet]) = k_{\mathrm{a0p}}^{\mathrm{app}} \sum_{n=2}^{\infty} n^k[\mathrm{P_nX}] + k_{a1p}[\mathrm{Cu^I/L}] \sum_{n=2}^{\infty} n^k[\mathrm{P}_n\mathrm{X}] - k_{\mathrm{d1p}}[\mathrm{Cu^{II}X/L}] \sum_{n=2}^{\infty} n^k[\mathrm{P}_n{}^\bullet] - k_{\mathrm{p}}[\mathrm{M}] \sum_{n=2}^{\infty} n^k([\mathrm{P}_n{}^\bullet] - [\mathrm{P}_{n-1}{}^\bullet]) - k_{\mathrm{tR}}[\mathrm{R}^\bullet] \sum_{n=2}^{\infty} n^k[\mathrm{P}_n{}^\bullet] - (k_{\mathrm{tc}} + k_{\mathrm{td}}) \sum_{n=2}^{\infty} n^k[\mathrm{P}_n{}^\bullet] \sum_{n=1}^{\infty}[\mathrm{P}_n{}^\bullet] \tag{13b}$$

From the sum of (13a) and (13b) is follows that:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\sum_{n=1}^{\infty} n^k[\mathrm{P}_n{}^\bullet]) = k_{\mathrm{a0p}}^{\mathrm{app}} \sum_{n=1}^{\infty} n^k[\mathrm{P_nX}] + k_{a1p}[\mathrm{Cu^I/L}] \sum_{n=1}^{\infty} n^k[\mathrm{P}_n\mathrm{X}] - k_{\mathrm{d1p}}[\mathrm{Cu^{II}X/L}] \sum_{n=2}^{\infty} n^k[\mathrm{P}_n{}^\bullet] - k_{\mathrm{p}}[\mathrm{M}] \sum_{n=1}^{\infty} n^k[\mathrm{P}_n{}^\bullet] + k_{\mathrm{p}}[\mathrm{M}] \sum_{n=1}^{\infty} n^k[\mathrm{P}_{n-1}{}^\bullet] - k_{\mathrm{tR}}[\mathrm{R}^\bullet] \sum_{n=1}^{\infty} n^k[\mathrm{P}_n{}^\bullet] - (k_{\mathrm{tc}} + k_{\mathrm{td}}) \sum_{n=1}^{\infty} n^k[\mathrm{P}_n{}^\bullet] \sum_{n=1}^{\infty}[\mathrm{P}_n{}^\bullet] \tag{14}$$

By Equation (14) is possible to obtain the 0th, 1st and 2nd differential moment equations. To obtain these equations, first it is necessary to remove the dependency on $[\mathrm{P}_{n-1}{}^\bullet]$. This is done by simplifying the term $k_{\mathrm{p}}[\mathrm{M}] \sum_{n=1}^{\infty} n^k[\mathrm{P}_n{}^\bullet] + k_{\mathrm{p}}[\mathrm{M}] \sum_{n=1}^{\infty} n^k[\mathrm{P}_{n-1}{}^\bullet])$ of (14). First of all, the summation index of the second sum is shifted as follows (Hungenberg and Wulkow, 2018):

$$-k_{\mathrm{p}}[\mathrm{M}] \sum_{n=1}^{\infty} n^k[\mathrm{P}_n{}^\bullet] + k_{\mathrm{p}}[\mathrm{M}] \sum_{n=1}^{\infty} n^k[\mathrm{P}_{n-1}{}^\bullet] = -k_{\mathrm{p}}[\mathrm{M}] \sum_{n=1}^{\infty} n^k[\mathrm{P}_n{}^\bullet] + k_{\mathrm{p}}[\mathrm{M}] \sum_{n=1}^{\infty} (n-1)^k[\mathrm{P}_n{}^\bullet] \tag{15}$$

Then, the final simplification of Equation 15 leads to changes according to the order of the moment considered. The final simplification result for the 0th, 1st and 2nd moments can be found in Table B.1 (Hungenberg and Wulkow, 2018).

Table B.1: Simplification results of Equation 15 for the 0th, 1st and 2nd moments.

Momentum	Result
0th	0
1st	$k_p[\mathrm{M}]\sum_{n=1}^{\infty}[\mathrm{P}_n^{\bullet}]$
2nd	$k_p[\mathrm{M}](\sum_{n=1}^{\infty}[\mathrm{P}_n^{\bullet}] + 2\sum_{n=1}^{\infty} n[\mathrm{P}_n^{\bullet}])$

Now, it is possible to obtain the differential equation for the 3 moments of the propagating chains. These are given by (16), using the notation described in Chapter 2.

$$\begin{aligned}\frac{\mathrm{d}\lambda_0}{\mathrm{d}t} &= k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_0 + k_{\mathrm{a1p}}\,[\mathrm{Cu^I/L}]\mu_0 - k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\lambda_0 \\ &\quad + k_{\mathrm{add}}\,[M][\mathrm{R}] - k_{\mathrm{tR}}\,[\mathrm{R}^{\bullet}]\,\lambda_0 - (k_{\mathrm{tc}} + k_{\mathrm{td}})\lambda_0\,\lambda_0\,,\end{aligned} \tag{16a}$$

$$\begin{aligned}\frac{\mathrm{d}\lambda_1}{\mathrm{d}t} &= k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_1 + k_{\mathrm{a1p}}\,[\mathrm{Cu^I/L}]\mu_1 - k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\lambda_1 + k_{\mathrm{p}}\,[\mathrm{M}]\lambda_0 \\ &\quad + k_{\mathrm{add}}\,[M][\mathrm{R}^{\bullet}] - k_{\mathrm{tR}}[\mathrm{R}]\,\lambda_1 - (k_{\mathrm{tc}} + k_{\mathrm{td}})\lambda_1\,\lambda_0\,,\end{aligned} \tag{16b}$$

$$\begin{aligned}\frac{\mathrm{d}\lambda_2}{\mathrm{d}t} &= k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_2 + k_{\mathrm{a1p}}\,[\mathrm{Cu^I/L}]\mu_2 - k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\,\lambda_2 + k_{\mathrm{p}}\,[\mathrm{M}]\,(\lambda_0 + 2\,\lambda_1) \\ &\quad + k_{\mathrm{add}}\,[M][\mathrm{R}] - k_{\mathrm{tR}}\,[\mathrm{R}^{\bullet}]\,\lambda_2 - (k_{\mathrm{tc}} + k_{\mathrm{td}})\lambda_2\,\lambda_0\,.\end{aligned} \tag{16c}$$

Dormant chains $\mathrm{P}_n\mathrm{X}$

For the $\mathrm{P}_n\mathrm{X}$ chains the derivation of Equation (2.11) can be made directly, because the chains with one monomer unit have the same exact molar balance that the ones with two or more repeating units. The balance to the $\mathrm{P}_n\mathrm{X}$ is then given by (17).

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\,t}\Big(\sum_{n=1}^{\infty} n^k\,[\mathrm{P}_n\mathrm{X}]\Big) &= -k_{a0p}^{\mathrm{app}}\sum_{n=1}^{\infty} n^k\,[\mathrm{P}_n\mathrm{X}] - k_{\mathrm{a1p}}\,[\mathrm{Cu^I/L}]\sum_{n=1}^{\infty} n^k\,[\mathrm{P}_n\mathrm{X}] \\ &\quad + k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\sum_{n=1}^{\infty} n^k\,[\mathrm{P}_n^{\bullet}]\end{aligned} \tag{17}$$

Then, the 0th, 1st and 2nd momentum equations are obtained from (17), and are given by (18)

$$\frac{\mathrm{d}\mu_0}{\mathrm{d}t} = -k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_0 - k_{\mathrm{a1p}}\,[\mathrm{Cu^{I}/L}]\mu_0 + k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\lambda_0 \tag{18a}$$

$$\frac{\mathrm{d}\mu_1}{\mathrm{d}t} = -k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_1 - k_{\mathrm{a1p}}\,[\mathrm{Cu^{I}/L}]\mu_1 + k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\lambda_1 \tag{18b}$$

$$\frac{\mathrm{d}\mu_2}{\mathrm{d}t} = -k_{\mathrm{a0p}}^{\mathrm{app}}\,\mu_2 - k_{\mathrm{a1p}}\,[\mathrm{Cu^{I}/L}]\mu_2 + k_{\mathrm{d1p}}\,[\mathrm{Cu^{II}X/L}]\,\lambda_2 \tag{18c}$$

Dead chains D_n

For the D_n chains (for simplicity D_n denotes the chains terminated by combination and disproportionation) the two terms of Equation (2.11) can be written as follows:

$$\frac{\mathrm{d}[\mathrm{D_1}]}{\mathrm{d}t} = k_{\mathrm{tR}}\,[\mathrm{R^{\bullet}}]\,[\mathrm{P_1^{\bullet}}] + k_{\mathrm{td}}\,[\mathrm{P_1^{\bullet}}]\sum_{n=1}^{\infty}[\mathrm{P}_n{}^{\bullet}] \tag{19a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{n=2}^{\infty}[\mathrm{D}_n] = k_{\mathrm{tR}}\,[\mathrm{R^{\bullet}}]\sum_{n}[\mathrm{D_n}] + k_{\mathrm{td}}\sum_{n=2}^{\infty}[\mathrm{P}_n]\sum_{n=1}^{\infty}[\mathrm{P}_n{}^{\bullet}] + \frac{k_{\mathrm{tc}}}{2}\sum_{n=1}^{\infty}n^k\sum_{r=1}^{n-1}[\mathrm{P}_n{}^{\bullet}]\,[\mathrm{P}_{n-r}{}^{\bullet}] \tag{19b}$$

The difference between Equation (19a) and (19b) is due to the fact that is not possible to have a D_n chain, which ended by combination, with only one repeating unit. The minimum number of units in this type of chain is two. The sum $\sum_{r=1}^{n-1}[\mathrm{P}_n{}^{\bullet}]\,[\mathrm{P}_{n-r}{}^{\bullet}]$ is the result of a modification of the original summation derived in Hungenberg and Wulkow (2018). This is because in the case of termination by combination there is a different summation term for polymers with odd or pair repeating units. This two different sum terms, derived in Hungenberg and Wulkow (2018), would increase the problem complexity and the difficulty of the numerical treatment, so a simplification is necessary (Hungenberg and Wulkow, 2018) (for more information about this subject and the mathematical demonstrations the reader is refereed to Hungenberg and Wulkow (2018)).

To obtain the differential moments equations it is necessary to simplify the double sum in (19b) as follows (Hungenberg and Wulkow, 2018; Mastan and Zhu, 2015):

$$\frac{k_{\mathrm{tc}}}{2}\sum_{n=1}^{\infty}n^k\sum_{r=1}^{n-1}[\mathrm{P}_n{}^{\bullet}]\,[\mathrm{P}_{n-r}{}^{\bullet}] = \frac{k_{\mathrm{tc}}}{2}\sum_{n=1}^{\infty}[\mathrm{P}_r{}^{\bullet}]\sum_{r=1}^{\infty}(n+r)^k\,[\mathrm{P}_n{}^{\bullet}] \tag{20}$$

Table B.2 show the expressions for the 0th, 1st and 2nd obtained by Mastan and Zhu (2015) by simplifying (20).

Table B.2: Simplification of (20) for the 0th, 1st and 2nd moments.

Momentum	Result
0th	$\frac{k_{\mathrm{tc}}}{2} \sum_{n=1}^{\infty} [\mathrm{P}_n^{\bullet}] \sum_{n=1}^{\infty} [\mathrm{P}_n^{\bullet}]$
1st	$k_{\mathrm{tc}} \sum_{n=1}^{\infty} n\,[\mathrm{P}_n^{\bullet}] \sum_{n=1}^{\infty} [\mathrm{P}_n^{\bullet}]$
2nd	$k_{\mathrm{tc}} \left(\sum_{n=1}^{\infty} n^2\,[\mathrm{P}_n^{\bullet}] \sum_{n=1}^{\infty} [\mathrm{P}_n^{\bullet}] + \left(\sum_{n}^{\infty} n\,[\mathrm{P}_n^{\bullet}]\right)^2 \right)$

Through the sum of (19a) and (19b), and using the expression in Table B.2, the moments equation are obtained:

$$\frac{\mathrm{d}\delta_0}{\mathrm{d}t} = k_{\mathrm{tR}}\,[\mathrm{R}^{\bullet}]\,\lambda_0 + k_{\mathrm{td}}\,\lambda_0\,\lambda_0 + \frac{k_{\mathrm{tc}}}{2}\,\lambda_0\,\lambda_0 \tag{21a}$$

$$\frac{\mathrm{d}\delta_1}{\mathrm{d}t} = k_{\mathrm{tR}}\,[\mathrm{R}^{\bullet}]\,\lambda_1 + (\,k_{\mathrm{td}} + k_{\mathrm{tc}}\,)\,\lambda_1\,\lambda_0 \tag{21b}$$

$$\frac{\mathrm{d}\delta_2}{\mathrm{d}t} = k_{\mathrm{tR}}\,[\mathrm{R}^{\bullet}]\,\lambda_2 + (\,k_{\mathrm{td}} + k_{\mathrm{tc}}\,)\,\lambda_2\,\lambda_0 + k_{tc}\,\lambda_1\lambda_1 \tag{21c}$$

C Experimental points

In Table C.3 are presented the experimental points used for the model validation in Chapter 4.

Table C.3: Experimental points for the model validation in Chapter 4.

time (min)	monomer conversion	PDI	M_n (g/mol)
5	0.0967	1.030	2590
17	0.2847	1.031	6308
38	0.5757	1.024	13070
60	0.7479	1.016	14599
82	0.8675	1.030	16604
105	0.9065	1.039	18644
130	0.9524	1.040	19669
160	0.9698	1.039	20038
195	0.9857	1.039	20230
240	0.9916	1.036	20895

The experimental points of the 2nd and 3rd experiment used in Chapter 6 are reported in the Tables C.4 and C.5.

Table C.4: Experimental points of the 2nd experience.

time (min)	monomer conversion	PDI	M_n (g/mol)
12	0.1713	1.052	14560
24	0.2411	1.033	21530
36	0.3400	1.028	29470
48	0.4524	1.034	40143
60	0.5306	1.040	47606
80	0.6201	1.036	56149
100	0.7353	1.039	68701
140	0.8536	1.052	78711
200	0.9287	1.074	86051

Table C.5: Experimental points of the 3rd experience.

time (min)	monomer conversion	PDI	M_n (g/mol)
5	0.0912	1.480	3590
11	0.2074	1.310	6308
20	0.3587	1.230	9070
30	0.5263	1.160	11599
45	0.6411	1.130	13604
70	0.8054	1.090	16644
95	0.8632	1.080	18669
125	0.9041	1.090	20038

D Approximated graphical representation of the empirical correlation matrix

For the monomer conversion the graphical representation of the empirical correlation matrix for the 1st, 2nd and 3rd interval are represented in the Figures D.1, D.2 and D.3, respectively.

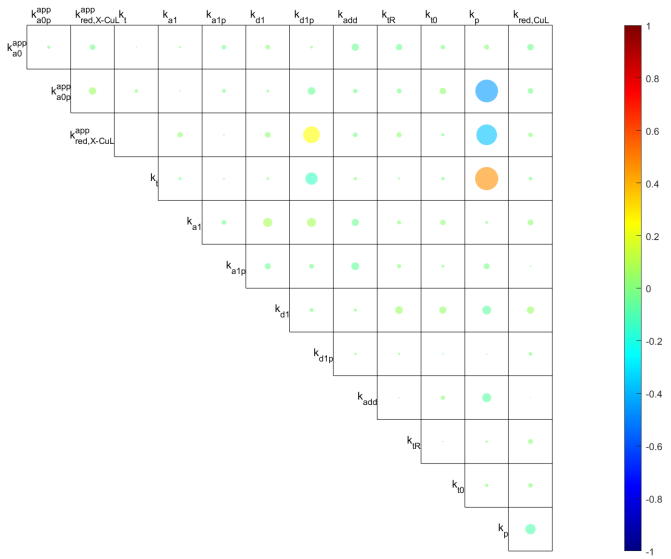


Figure D.1: 1st interval for the monomer conversion.

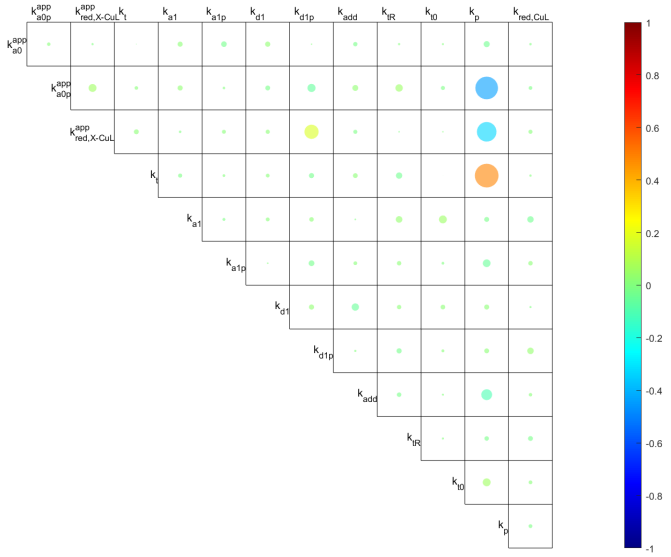


Figure D.2: 2nd interval for the monomer conversion.

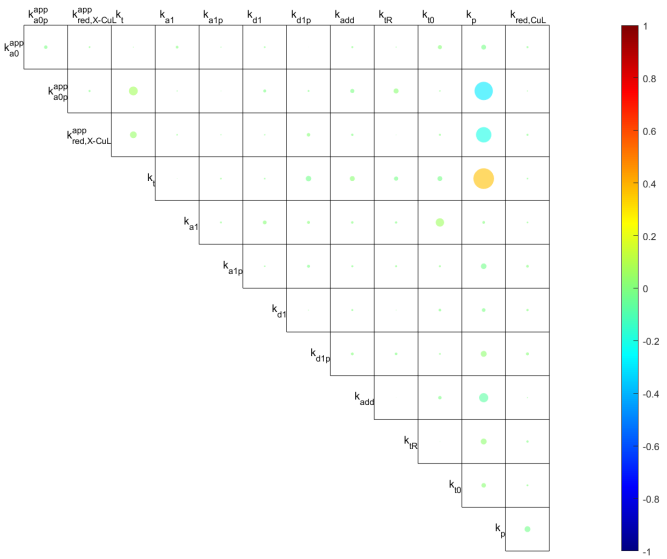


Figure D.3: 3nd interval for the monomer conversion.

For the PDI the graphical representation of the empirical correlation matrix for the 1st, 2nd and 3rd interval are represented in the Figures D.4, D.5 and D.6, respectively.

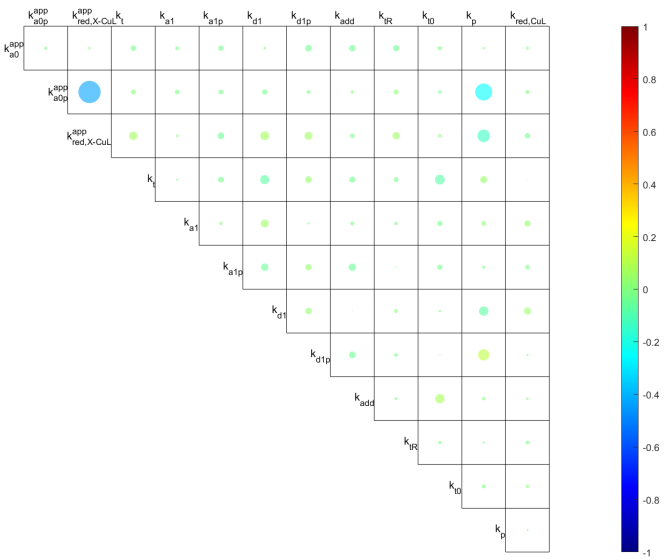


Figure D.4: 1st interval for the PDI.

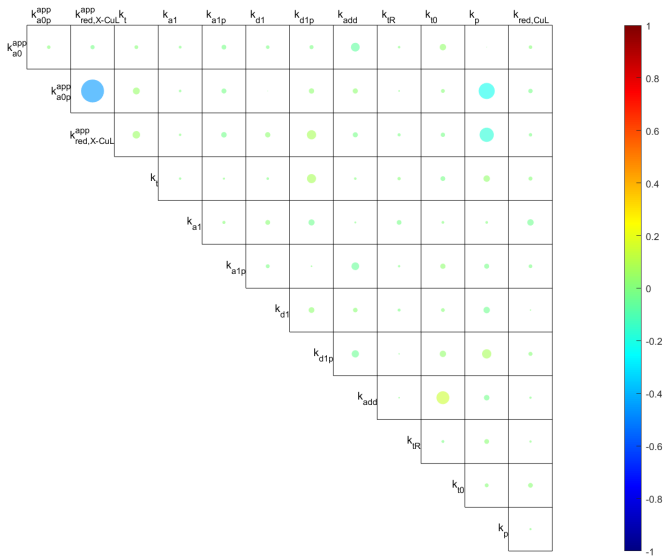


Figure D.5: 2nd interval for the PDI.

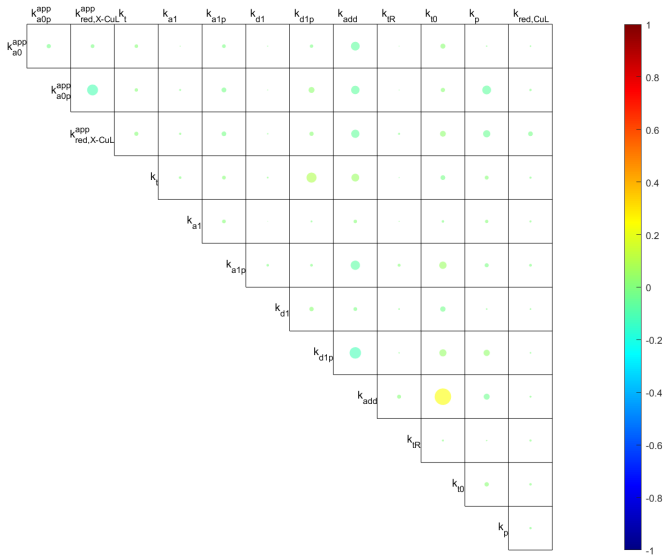


Figure D.6: 3rd interval for the PDI.

For the M_n the graphical representation of the empirical correlation matrix for the 1st, 2nd and 3rd interval are represented in the Figures D.7, D.8 and D.9, respectively.

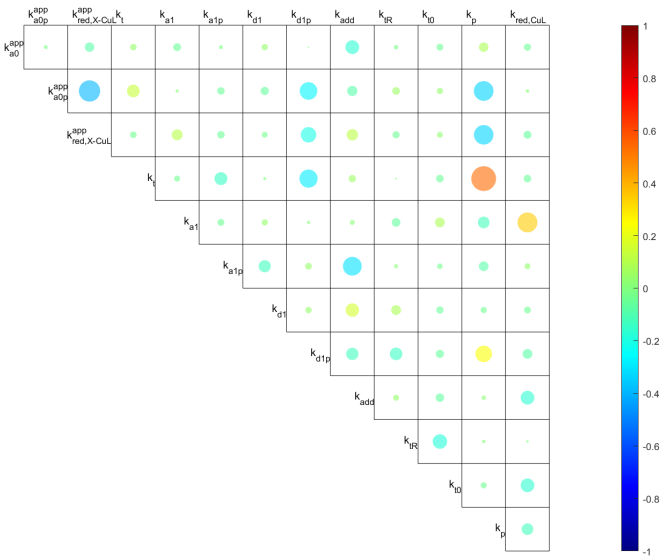


Figure D.7: 1st interval for the M_n.

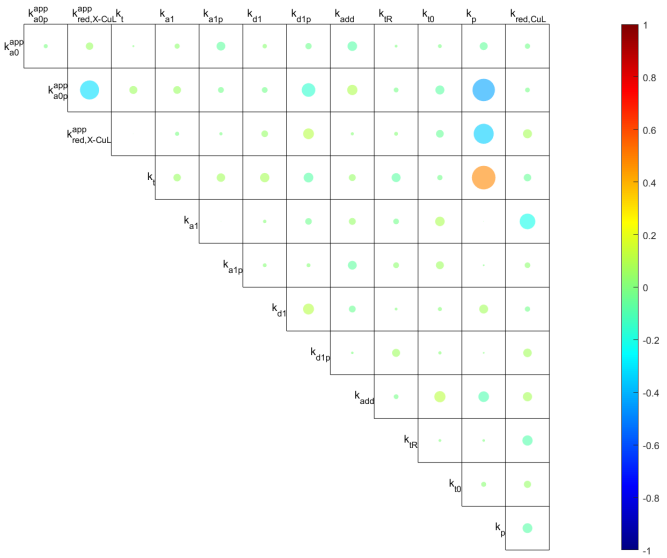


Figure D.8: 2nd interval for the M_n.

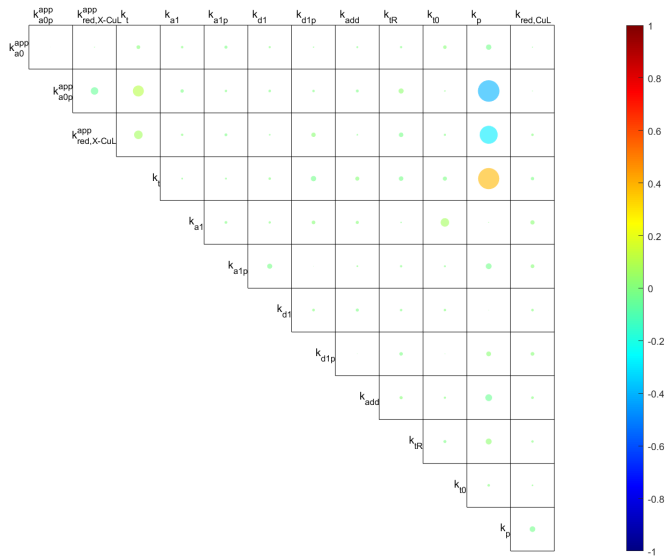


Figure D.9: 3rd interval for the M_n.

E Implementation test

An example is considered to test whether the method was properly implemented in MATLAB®, for the calculation of the first order s indices and total sensitivity indices as well. The code was tested with a simple function (22) where the parameters X_i are uniformly distributed between 0 and 1. This function was selected because the first order and total order sensitivity indices were calculated using a Monte-Carlo method based on Sobol approximations in Qian and Mahdi (2020), which facilitates the comparison to the results of the code in this work.

$$y = X_1^2 + X_2X_3 + X_4 \tag{22}$$

The author does not mention the exact method they used to specify the number of samples generated for this study. Because the computational cost to study this function is low, it will be used a N= 10000, since it is the number of N all the sensitivity indices approximated the value calculated analitically, in the previous example. The result for sensitivity indices obtained with the code implemented in this work and Qian and Mahdi (2020) are presented in Table E.6. The obtained sensitivity indices are consistent with those obtained by Qian and Mahdi (2020) using a similar method. It can be concluded that the method is well implemented in MATLAB®.

Table E.6: Global sensitivity analysis results of Equation 22.

	Qian and Mahdi (2020)		This work	
Parameter	S_i	S_{T_i}	S_i	S_{T_i}
X_1	0.403	0.403	0.402	0.402
X_2	0.094	0.126	0.094	0.126
X_3	0.094	0.126	0.094	0.126
X_4	0.377	0.377	0.377	0.378

F Algorithm for the weights determination

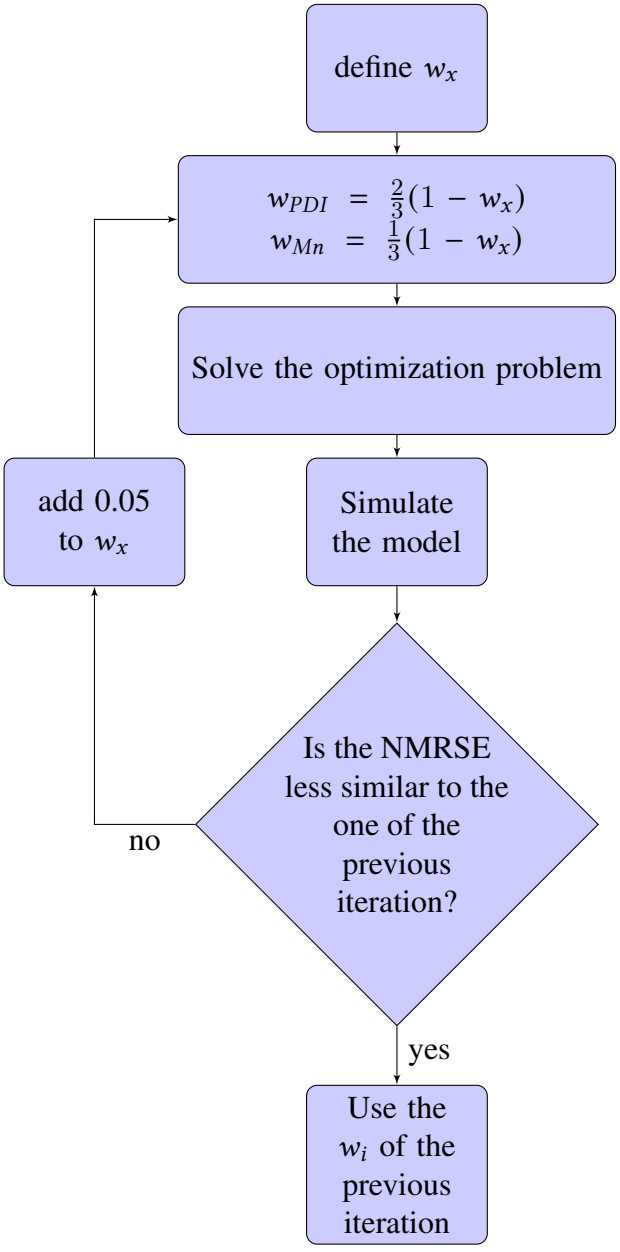


Figure F.10: Steps to calculte w_i.