MODELAMIENTO ESTÓCASTICO DE BIORREACTORES

STOCHASTIC MODELING OF BIOREACTORS

ANDERSON VALENCIA ISAZA

Thesis of degree project presented as a partial requirement to quality for the title of:

Maestria en ingeniería - Ingeniería química

Director:

Ph.D. Javier Ignacio Carrero Mantilla

Co-Director:

Ph.D. Óscar Andrés Prado Rubio

Grupo de Investigación en Fisicoquímica Computacional

Grupo de Investigación en Aplicación de Nuevas Tecnologías, GIANT

Universidad Nacional de Colombia

Facultad de Ingeniería y Arquitectura, Departamento de Ingeriería Química

Manizales, Colombia

2021

ACKNOWLEDGMENTS

The path traced during these years as a master's student at the Universidad Nacional de Colombia - Sede Manizales, has been one of great personal, professional and academic growth. This has allowed me to know my limits, how to face and overcome them, he has to be resilient to face each new challenge with the best disposition, always giving my best and learning from my mistakes.

I have had the opportunity to be supervised by Prof. Javier Ignacio Carrero Mantilla, from whom I take his ability to overcome obstacles, find practical ways to solve problems, be specific when writing and always be willing to solve my problems. And by Prof. Óscar Andrés Prado Rubio, from whom I take his capacity for analysis, understanding, application, and presentation the issues, his high demand regarding the way of writing, arguing, and linking ideas, as well as his willingness to guide me and recognize my successes but also call my attention to my mistakes. For this and much, I am immensely grateful to both of them for allowing me to work under his direction, have trusted me, and has guided me during this enriching path in the art of research. Through the discussions, I had with both of them, allowed me to improve my capacity for critical thinking, question everything, and find a way to answer it.

None of this would have been possible without the help of God, who allowed me to have a job with which to pay for my studies, I can grace my co-advisers to support me on this path. Last but not least, I want to especially thank my parents Luz Bibiana Isaza Gómez, Hector Higinio Valencia Giraldo, and John Jairo Castrillón Castro because each of them has been fundamental support during this process that sometimes becomes a roller coaster of emotions, but thanks to your words, your listening, your direction, your councils continue forward and do not faint along the way. I also thank my family, friends, work colleagues, and study colleagues who in one way or another were important help at this time with their prayers, patience, advice, and support. I reiterate my sincere thanks.

Anderson Valencia Isaza

Manizales, Caldas

Abril, 2021

Abstract

Stochastic Modeling of Bioreactors

Modeling and simulation have become over time a key tool for the continuous improvement of designs, control systems, and understanding of industrial processes, including biotechnology. Processes that respond to the need to make use of alternative sources to those based on petroleum, for the production of consumer products of all kinds. This is why this work focuses on modeling biotechnological processes such as enzymatic transformation and fermentation. The modeling was investigated using stochastic and deterministic approaches for two case studies, enzymatic reactions with and without inhibition, as well as the xylitol bioproduction from glucose and xylose. The deterministic approach was based on the solution of the ordinary differential equations obtained from the mass balances for a batch reactor and the chemical kinetics of the mass action law and the Michaelis-Menten type. On the other hand, the stochastic approximation was based on the application of the Gillespie stochastic solution algorithm, from the stochastic chemical kinetics and the stochastic approximation of the Michaelis-Menten equations. Thus, the objective of this thesis is to investigate the application of stochastic models and evaluate the uncertainty of these systems, as well as to re-estimate parameters for the xylitol production model.

It was found in this work that stochastic methods have the same predictive power as deterministic ones for the most common cases of inhibition in the literature (Competitive, Non-competitive, Un-competitive) and the xylitol production, but with the advantage of being able to analyze the inherent uncertainty of biological systems (enzymatic reactions and fermentation). Besides, It was possible to evaluate the evolution of the uncertainty of the stochastic model throughout the simulation and its relationship with the relative uncertainty of the experimental measurements. The sensitivity analysis allowed making a conscious selection of the size of the system and the number of realizations necessary to obtain a practically constant uncertainty value. Moreover, these methods served to propose a translation of the xylitol bioproduction model into stochastic terms, despite the lack of information on the reactions that lead to the formation of biomass and the inclusion of the mass transport model. Besides, the deterministic re-adjustment of the parameters showed a better fit of the model to the experimental data of xylitol (compound of interest). The greatest contribution, in this case, was to illustrate how this method (ABC rejection sampler) offers greater robustness when evaluating the uncertainty of the model and being able to optimize under uncertainty. Therefore, this thesis includes contributions to the state of the art in Process System En-

gineering that will provide new perspectives for the modeling and analysis of systems under uncertainty.

Keywords: Stochastic Modeling, Gillespie Method, Bioprocess, Michaelis-Menten, Stochastic Chemical Kinetics.

Resumen

Modelamiento estocástico de biorreactores

El modelado y la simulación se han convertido con el tiempo en una herramienta clave para la mejora continua de los diseños, los sistemas de control y la comprensión de los procesos industriales, incluida la biotecnología. Procesos que responden a la necesidad de hacer uso de fuentes alternativas a las basadas en petróleo, para la elaboración de productos de consumo de todo tipo. Es por ello que este trabajo se centra en la modelización de procesos biotecnológicos como la transformación enzimática y las fermentaciones. El modelado se investigó utilizando enfoques estocásticos y deterministas para dos estudios de caso, reacciones enzimáticas con y sin inhibición, así como la bioproducción de xilitol a partir de glucosa y xilosa. El enfoque determinista se basó en la solución de las ecuaciones diferenciales ordinarias obtenidas de los balances de masa para un reactor discontinuo y la cinética química de la ley de acción de masas y el tipo de Michaelis-Menten. Por otro lado, la aproximación estocástica se basó en la aplicación del algoritmo de solución estocástica de Gillespie, a partir de la cinética química estocástica y la aproximación estocástica de las ecuaciones de Michaelis-Menten. Así, el objetivo de esta tesis es investigar la aplicación de modelos estocásticos y evaluar la incertidumbre de estos sistemas, así como reestimar parámetros para el modelo de producción de xilitol.

En este trabajo se encontró que los métodos estocásticos tienen el mismo poder predictivo que los deterministas para los casos más comunes de inhibición en la literatura (Competitivo, No competitivo, No competitivo) y la producción de xilitol, pero con la ventaja de poder analizar la incertidumbre inherente a los sistemas biológicos (reacciones enzimáticas y fermentación). Además, se pudo evaluar la evolución de la incertidumbre del modelo estocástico a lo largo de la simulación y su relación con la incertidumbre relativa de las medidas experimentales. El análisis de sensibilidad permitió hacer una selección consciente del tamaño del sistema y el número de realizaciones necesarias para obtener un valor de incertidumbre prácticamente constante. Además, estos métodos sirvieron para proponer una traducción del modelo de bioproducción de xilitol en términos estocásticos, a pesar de la falta de información sobre las reacciones que conducen a la formación de biomasa y la inclusión del modelo de transporte masivo. Además, el reajuste determinista de los parámetros mostró un mejor ajuste del modelo a los datos experimentales del xilitol (compuesto de interés). El mayor aporte, en este caso, fue ilustrar cómo este método (muestreador de rechazo ABC) ofrece una mayor robustez a la hora de evaluar la incertidumbre del modelo y poder optimizar bajo

incertidumbre. Por lo tanto, esta tesis incluye contribuciones al estado del arte en Ingeniería de Sistemas de Procesos que brindarán nuevas perspectivas para el modelado y análisis de sistemas bajo incertidumbre.

Palabras clave: Modelamiento Estocástico, Método Gillespie, Bioprocesos, Michaelis-Menten, Cinéticas Químicas Estocásticas.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Abstract

The study of the modeling of enzymatic reactions and fermentation processes from the beginning of the 20th century with the Michelis-Menten and Monod-based kinetics. Nowadays, important advances in process monitoring and control drive the need on proposing more robust models that allow predicting, evaluating, and understanding the different conditions that give rise to obtain a product, thus improve industrial processes. This chapter presents a brief review of the modeling and simulation evolution of these systems from two approaches, deterministic and stochastic. From those, there is a special focus on stochastic modeling of biotechnological processes since it allows evaluating the behavior of these systems from the inherent uncertainty of biological systems. Based on which the hypotheses and objectives of the thesis were raised. Also, the methodology followed for the two case studies analyzed is presented: enzymatic reactions with and without inhibition, and the xylitol bioproduction.

1.1 State of the art

According to their mode of operation, bioreactors can be classified as steady-state or continuous (with inlet and outlet flows), batch (without inlet or outlet flows), or fed-batch reactors (they have one or more inflows, but not outflows). batch or fed-batch is the most widely used because they are easy to sterilize and adapt to different processes, and also reduce the risk of genetic mutation of microorganisms, since they do not have very long residence times (Villadsen, Nielsen, & Lidén, 2011).

Table 1.1. Model classification for bioreactors, partially based on.

		Type of equations	
According to	Models	Stable Steady	Dynamic
Input information	Phenomenological Empirical		
Randomness	Deterministic	Nonlinear algebraic	ODE/PDE
	Probabilistic	Algebraic/difference	Stochastic ODE or Difference
	Correlation		
Time-space dependence	Lumped Distributed	Algebraic EPDE	ODE PPDE
Microorganism location	Spatial No spatial		
Applicability of superposition principle	Linear Non linear	Linear algebraic Nonlinear algebraic	Linear ODE Non linear ODE
Type of variable	Continuous Discrete Hybrids	Algebraic Difference	ODE Difference

ODE = Ordinary Differential Equiation, PDE = Partial Differential Equiation, EPDE = Eliptic Partial Differential Equation, PPDE = Parabolic Partial Differential Equiation. Cameron and Hangos (2001).

Classification of bioreactor models is difficult because several criteria can be applied simultaneously (see Table 1.1). In turn, bioreactor models can also be deterministic or stochastic according to how the parameters and values of the variables are defined (Dehling, Gottschalk, & Hoffmann, 2007). In deterministic models, the same input conditions will always produce the same output conditions. Stochastic models are based on the introduction of at least one parameter that determines probabilistic relationships between the variables, that is, it introduces the concept of uncertainty to the phenomenon studied, thus each run of the stochastic model will be different from the previous one. Some spatial stochastic models take into account the interactions of microorganisms with each other and with their environment, while others aim to study the dynamic behavior within the microorganism (Andrews, Dinh, & Arkin, 2009). An important part of bioreactor modeling is the study of chemical kinetics that describe microbial growth (Figure 1.1) and enzymatic kinetics. Within this field, the kinetics of microbial growth developed by Monod (1942) from an empirical base was the starting point for the rest of the kinetics of this type.

Figure 1.1 shows the 4 possible combinations of mathematical models that can be

built depending on the following approaches (González-Figueredo, Flores-Estrella, & Rojas-Rejón, 2018; Villadsen et al., 2011):

- Unstructured: manage the population as a mono-component system.
- Structured: model the population as a multi-component system since the microorganisms are suspended in the liquid phase.
- Non-segregated: consider the average behavior of the population of microorganisms.
- Segregated: describe the individual behavior of each organism of a heterogeneous population.

Therefore, the Monod kinetic is an un-structured and non-segregated model, since the model treats the population of the microorganisms as one component solute and its average behavior.

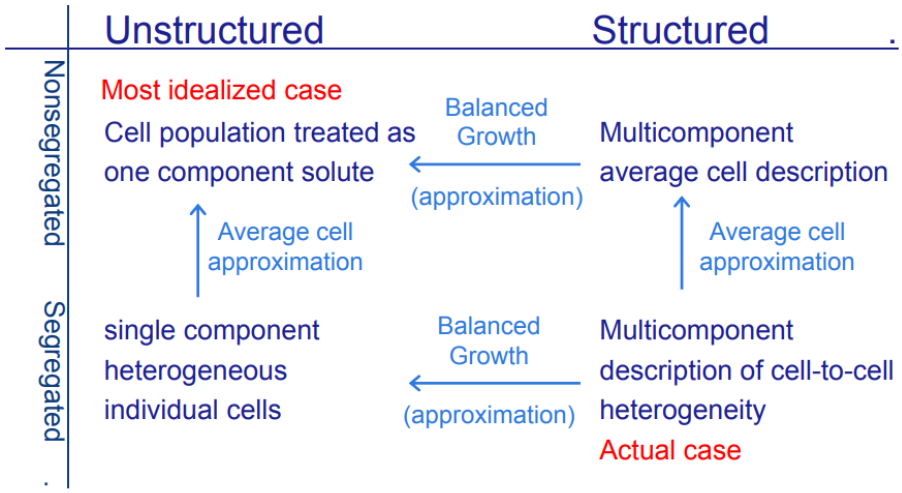

Figure 1.1. Classification of growth kinetics models.

On the other hand, in the field of enzymatic reactions, there are two main approaches to enzyme kinetics :

- Law mass action: describes the behavior of substrates, enzymes, complexes, and products from elemental reactions. The problem with this approach is that it is experimentally difficult to measure the concentrations of the enzyme complexes.

Thus, there is most likely insufficient information to estimate the parameters of the kinetic constants associated with these reactions (Enderle, 2012; Källén, 2018).

- Michaelis-Menten kinetics: describes the behavior of the substrates and the products from the Quasi-Steady-State Assumption (QSSA), this approach circumvents the need to know the concentration of complexes to represent the system behavior (Michaelis & Menten, 2007; Rogers & Gibon, 2009; Villadsen et al., 2011).

Both approaches have traditionally been modeled in a deterministic way solving Ordinary Differential Equations (ODEs).

Bioprocesses have been known empirically since ancient times, but the introduction of microorganism culture techniques permitted the manufacture of antibiotics, amino acids, organic acids, vaccines, etc. Later in the 1970s and 1980s, the ascent of genetic engineering allowed the manipulation of genes and metabolic pathways of living organisms (Ortega Quintana, Alvarez, & Botero Castro, 2017; Smith, 2002). This led to the development of equipment such as biochemical reactors or bioreactors to manipulate microorganisms by controlling the conditions of the environment (temperature, pH, aeration, substrate, etc.) to obtain the desired products.

Deterministic modeling of bioreactors began with the growth rate equation proposed by Blackman (1905). Then, Michaelis and Menten (1913) proposed a rate equation to describe the kinetic behavior of enzymes, which has been widely studied by Murugan (2018). Monod (1942) developed a kinetic growth equation that relates the substrate concentration with the growth rate of the microorganisms.

The Michaelis-Menten and Monod equations have served as a starting point for the development of a large number of kinetic models applicable to bioprocesses (Ortega Quintana et al., 2017). The first to propose phenomenological bases (based on physical and chemical principles) for bioprocesses proposed by Konak (1974) in 1974, with an empirical growth rate model. It was not until Button (1998) that a phenomenological approach was achieved by defining the concept of specific affinity, demonstrating the ability of the cell to choose a specific substrate.

Later Liu (2006) proposed a model of microbial growth based on thermodynamics, in which he managed to relate these processes to the change in Gibbs free energy. Even today it has not been possible to propose a completely phenomenological model, although semi-physical models are close to that ideal (Ortega Quintana et al., 2017).

The need to represent biological processes more accurately imposes a challenge to develop models that describe cellular metabolism and can be linked to bioprocess mod-

els. Most of the current efforts are based on the deterministic approach, but stochastic methods have become a suitable alternative to describe the intrinsic and extrinsic uncertainty of these systems. Instance: in their pioneering 1979 study, Stephanopoulos, Aris, and Fredrickson (1979) simulated the competition of two microorganisms for the same substrate source using stochastic models. But it took a little over 25 years for Imhof and Walcher (2005) to return to stochastic models to study the extinction or persistence of a population of microorganisms in a chemostat. Thus, they were able to demonstrate that random effects can lead to the extinction of the microorganism, while deterministic models predict their persistence.

In 2011 Campillo, Joannides, and Larramendy-Valverde (2011) proposed a set of stochastic models applicable for the operation of small, medium, and large-scale chemostats. The results of each model were compared with their respective deterministic solution. Subsequently, in 2015 Fritsch, Harmand, and Campillo (2015) proposed a model that serves as a bridge between deterministic and stochastic approach, since the system is interpreted as continuous populations of the same size (deterministic point view) and as a set of individual organisms with random and dynamic behaviors (stochastic point view). More recently Fontbona, Riquelme, and Silva (2017), proposed a stochastic model to describe the behavior of a sequential batch reactor used in wastewater treatment through the use of stochastic differential equation systems (Fontbona et al., 2017; Rincón, 2006).

On the other hand, the theoretical foundations of stochastic chemical kinetics and its simulation were given by McQuarrie (1967) and Gillespie (1977). Then, Rao and Arkin (2003) proposed a stochastic approximation of the QSSA from the Chemical Master Equation (CME) to be used in case studies that meet the conditions of Michaelis-Menten type kinetics. Subsequently, Wilkinson (2006) presents the applicability of stochastic chemical kinetics for the case of enzymatic reaction without inhibition, work that is taken up by Lecca (2013), Anderson and Kurtz (2015), and Marchetti, Priami, and Thanh (2017). Besides, authors such as Agarwal, Adams, Castellani, and Shouval (2012); Lawson, Petzold, and Hellander (2015); Sanft, Gillespie, and Petzold (2011); Warne, Baker, and Simpson (2019); Wu, Vidakovic, and Voit (2011) continue to expand the formulation of Rao and Arkin (2003); coming to Kang, KhudaBukhsh, Koeppl, and Rempała (2019)'s work, in which they make a referral for all types of QSSAs. The literature also includes the use of hybrid models or machine learning as a modelling and simulation strategy of biological processes. However, in this work, these topics are not developed, since it was focused on the advantages and applicability of pure stochastic methods, especially in the study of the uncertainty linked to these processes.

This review provides evidence of the relevance of stochastic models as another form of evaluation and description of biological processes but from a micro approach since the macro approach is well understood from deterministic models. Observing a possible point of union between both types of modeling, since each one can model a type of approach either macro or microscopic. Furthermore, after a review of the literature, a stochastic approach to diauxic growth with inhibition, enzymatic reactions with inhibition complemented with mass transport has not been reported in a single particular case. Therefore, there is a substantial contribution to the case of xylitol Bioproduction. In addition, the extension that was made for the cases of enzymatic reactions with inhibition from the Stochastic Chemical Kinetics approaches and the stochastic approach of the Quasi-Stable State Assumption.

1.2 Justification

Bioprocess modeling has become a relevant field of study to improve production processes in food, chemical and pharmaceutical industries, which base their processes on the use of microorganisms. Similarly, modeling has focused on understanding microbial growth, metabolite production, enzymatic processes, and more recently on the interaction between the medium and the cellular material of the microorganism (Ortega Quintana et al., 2017). Hence, the importance of modeling in the field of Process Systems Engineering (PSE). This could lead to optimal designs and operating conditions, which in the industry translate into greater efficiency, safety, and profitability of the processes. Thus, more and more robust models are required that are better adapted to the particularities of the phenomena, to obtain more realistic representations.

Traditionally, the study of enzymatic kinetics has been carried out from the deterministic approach. Nevertheless, there is a growing need to have models that allow representing both the uncertainty of biological processes. Specifically, this thesis studies the modeling of enzymatic processes since enzymes are excellent natural catalysts, which contributes to the development and obtaining of new products. The challenge in this field focuses on the understanding and modeling of the reaction mechanisms through which the design and optimization of industrial processes can be carried out (Lonsdale, Harvey, & Mulholland, 2012) The work focuses on the study of reactions through stochastic simulations since these models work at the molecular level and serve as a complement to experimental techniques. This is particularly relevant since biotechnological processes imply uncertainties regarding their behavior, dynamics, and interaction between enzymes/cells and their environment (intrinsic noise), in addition to the

uncertainty of experimental measurements (extrinsic noise) (Tsimrng, 2014; Uusitalo, Lehikoinen, Helle, & Myrberg, 2015). Therefore, studying this type of modeling will provide deeper process understanding. Besides, many of the current models do not take into account this degree of uncertainty in the face of the random behaviors of microorganisms. And this will bring new tools for better design, optimization, and control of biotechnological systems.

Until now, the published stochastic models focus on the enzymatic reaction and microorganisms growth from purely mathematical aspects, lacking specific application cases (Fontbona et al., 2017; Kang et al., 2019; Wilkinson, 2018). On the other hand, stochastic chemical kinetics has been studied for generic cases, which are not necessarily centered on conditions at the industrial level. Therefore, this thesis seeks is focused in the following aspects of stochastic modeling:

- Definition of the system volume
- Realizations number
- Relationship between experimental and stochastic model uncertainty
- Real industrial application case

The literature review has shown that most of the advances in the field of stochastic bioprocess modelling have been carried out by researchers from the exact sciences (mainly mathematical, physical, and chemical). Therefore, this work seeks to give greater visibility to stochastic modelling in the field of chemical engineering, showing its advantages for equipment design under uncertainty at an industrial level.

This work aims to improve understanding of enzymatic biological processes with and without inhibition, as well as showing its applicability in the specific case of xylitol bioproduction and them parameters estimation. This is achieved, through the application of stochastic chemical kinetics as an alternative to the deterministic law of mass action and the use of the CME to translate the ODEs of the Michaelis-Menten to estimate the intrinsic uncertainty and model parameters. Thus, the industrial application of this model would allow improving the operation, design, and control under the uncertainty of the bioreactors.

1.3 Hypothesis and objectives

- **Hypothesis**

The key idea of this project is to investigate stochastic modeling to represent enzymatic catalysis and microbial catalysis. The project is focused on the uncertainty associated with these simulations and their relationship with the behavior of this type of system and experimental measurements. Therefore, the hypothesis of this thesis is:

Models based on the stochastic chemical equations or the approximation of the differential equations to propensity functions can reproduce the uncertainty associated with the experimental measurements and the behavior of the chemical species in biotechnological processes.

Note: Although studies on the application of stochastic modeling of bioprocesses are presented in the literature. The hypothesis of this thesis focuses on the analysis of uncertainty in more complex biological processes (enzymatic with inhibition and fermentative), as well as its applicability in specific cases.

- **General objective**

To investigate the application of a stochastic model for the enzymatic reaction and bioproduction of xylitol through stochastic chemical kinetics and the Stochastic Quasi-Stable State Assumption, to provide uncertainty information and predict the behavior of the system.

- **Specific objectives**
  - Develop a stochastic model that can describe the dynamics of enzymatic reactions for generic cases and xylitol bioproduction case.
  - Implement an algorithm that allows solving the proposed stochastic model for the planted cases.
  - Verify the proposed stochastic model by comparing its results versus experimental data and deterministic models.

1.4 Project Methodology and thesis content

This work focuses on two case studies: enzymatic reactions with and without inhibition and xylitol bioproduction modeling. The methodology is described in Figure 1.2, it starts with the phenomenological identification includes the elemental reactions that defend the enzymatic reactions studied, as well as microbial growth, enzymatic inhibition, and the transport of matter that participate in the bioproduction of xylitol. In the model construction, deterministic simulation becomes based on the application

of the law of mass action and the Michaelis-Menten kinetics, while stochastic simulation works with stochastic chemical kinetics, the stochastic approximation of the Quasi-Stable State Assumption (stochastic QSSA), and the propensity functions. The model implementation was done with numeric integration of differential equations and the Gillespie method for deterministic and stochastic simulation, respectively. Then, the structural verification was also performed for both cases through comparison with the literature for similar cases, deterministic results and for the specific case of xylitol with experimental data (Rao & Arkin, 2003; Sanft et al., 2011; Tochampa et al., 2005; Wilkinson, 2018). There is a stage of parameter identification and model verification. Notice that model validation is only performed in the case that enough experimental data are available for parameter identification and validation. This stage was only applied to the case of xylitol bioproduction since there were experimental data that allowed for the identification of parameters, but no validation was done since all the experimental data were used for the estimation of parameters.

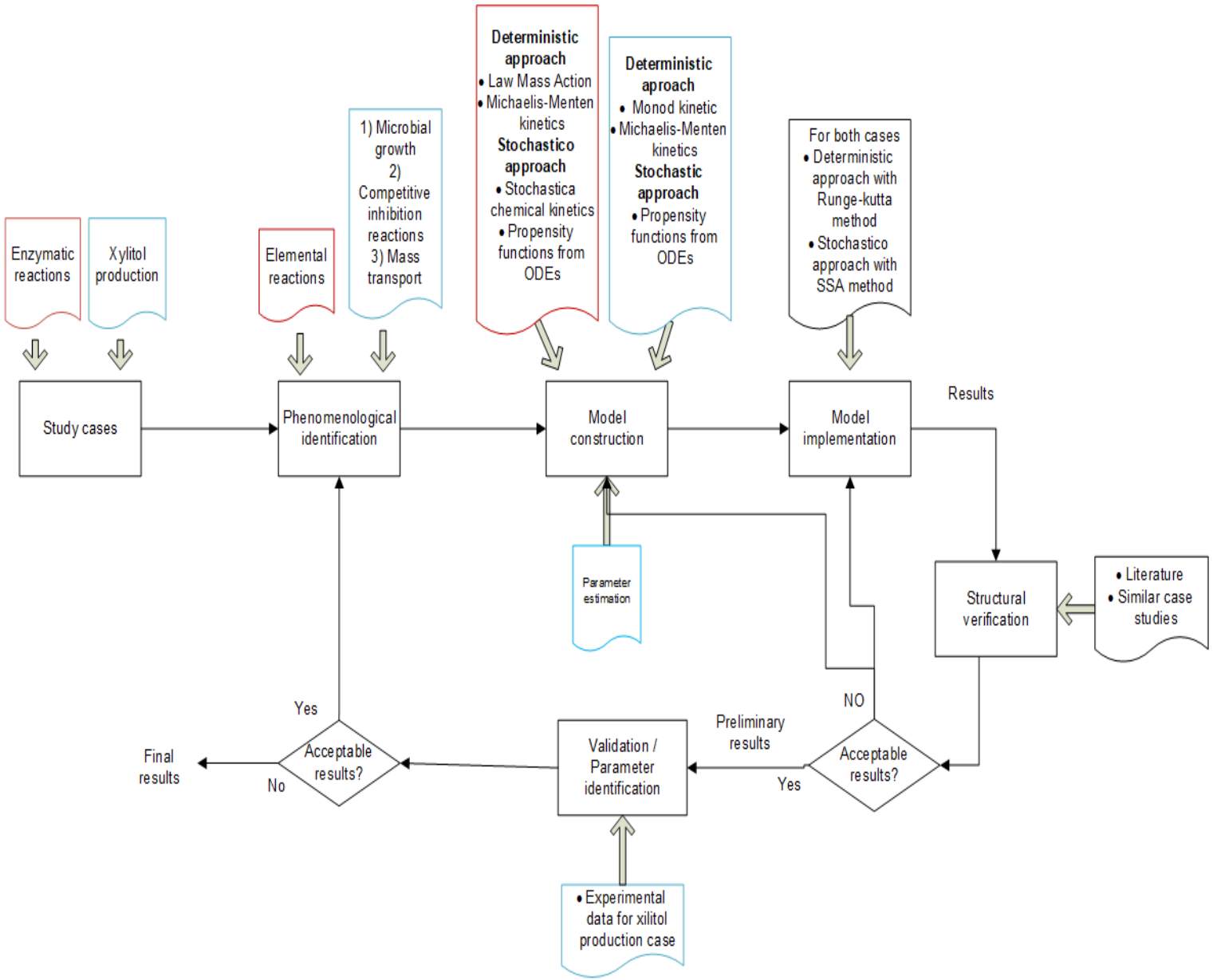

Figure 1.2. Flow diagram for the model development and validation.

In Chapter 2, four cases of enzymatic reactions (no inhibition, competitive inhibition, non-competitive and non-competitive inhibition) were studied. Each case was studied from two kinetic approaches: the law of mass action and kinetics of the Michaelis-Menten type. Both were simulated using deterministic and stochastic approaches. or the stochastic simulations, a sensitivity analysis was carried out to select the size of the system and the number of realizations based on the uncertainty produced by the method. The kinetic approaches are generically described below from deterministic and stochastic perspectives.

- Law of mass action: the deterministic approach starts from the approach of ordinary differential equations based on the kinetics obtained for each elemental reaction. From the derived stochastic chemical equations, the propensity functions are obtained (section 2.2.3).

- Michaelis-Menten kinetics: the deterministic approach started from the QSSA which leads to the derivation of the Michaelis-Menten kinetics and ODEs that describe the evolution of the substrate and the product. The stochastic approximation was carried out from the mass balances, which were translated into propensity functions through the application of the Master Chemical Equation (SCK section), this approximation is known as stochastic QSSA (Rao & Arkin, 2003; Toral & Colet, 2014).

This chapter corresponds to the AMIDIQ conference article and oral presentation, Besides, a paper submitted in the Journal of Mathematical Chemistry. On the other hand, Chapter 3 deals with the xylitol production model proposed by Tochampa et al. (2005), which is translated stochastic modeling using the stochastic QSSA. As a particular characteristic, a re-estimation of parameters of the xylitol model was performed with the deterministic method of the interior point and with the stochastic ABC rejection sampler method. Using this approach, it was possible to make a hand-to-hand comparison of both approaches for a real case from the literature.

The discussion of the case studies presented in chapters 2 and 3 focused more on the chemical kinetics and its relationship with the uncertainty obtained from the stochastic model, than the biological interpretation of the phenomena present in the bioreactor. Both case studies have been widely studied in the literature from the biological point of view, and it was wanted to give greater depth to the kinetic understanding of the processes.

1.5 Contributions

Papers

- Anderson Valencia Isaza, Oscar Andrés Prado-Rubio, Javier Ignacio Carrero Mantilla. (2020). Modeling and uncertainty assessment for enzymatic reactions through stochastic kinetic simulations. Status: submitted to ChemTexts (under review).

- Anderson Valencia Isaza, Oscar Andrés Prado-Rubio, Javier Ignacio Carrero Mantilla. (2020). Stochastic modeling approach and tuning for fermentation models. Status: for submitting.

Peer reviewed conference papers

- Anderson Valencia Isaza, Oscar Andrés Prado-Rubio, Javier Ignacio Carrero Mantilla. (2020). Modelamiento y Simulación Estocástica para Cinéticas Enzimáticas con Inhibición. In Proceedings of "XLI Encuentro Nacional de la AMIDIQ" (ISBN: en trámite). María del Rosario Enríquez Rosado (Editor). Pages: BIO-117-122. Academia Mexicana de Investigación y Docencia en Ingeniería Química (AMIDIQ).

- Anderson Valencia Isaza, Oscar Andrés Prado-Rubio, Javier Ignacio Carrero Mantilla. (2020). Modelamiento y Simulación Estocástica para Cinéticas Enzimáticas con Inhibición. Oral presentation at Academia Mexicana de Investigación y Docencia en Ingeniería Química (AMIDIQ), 22th to 24th of October, 2020. Virtual meeting.

Bibliography

Agarwal, A., Adams, R., Castellani, G. C., & Shouval, H. Z. (2012). On the precision of quasi steady state assumptions in stochastic dynamics. *The Journal of chemical physics*, *137*(4), 44105.

Anderson, D. F., & Kurtz, T. G. (2015). *Stochastic analysis of biochemical systems* (Vol. 1). Springer.

Andrews, S. S., Dinh, T., & Arkin, A. P. (2009). Stochastic Models of Biological Processes. *Encyclopedia of Complexity and Systems Science*, 8730–8749. Retrieved from `http://link.springer.com/10.1007/978-0-387-30440-3{_}524` doi: 10.1007/978-0-387-30440-3_524

Blackman, F. F. (1905). Optima and limiting factors. *Annals of botany*, *19*(74), 281–295.

Button, D. K. (1998). Nutrient uptake by microorganisms according to kinetic parameters from theory as related to cytoarchitecture. *Microbiology and molecular biology reviews*, *62*(3), 636–645.

Cameron, I. T., & Hangos, K. (2001). *Process modelling and model analysis* (Vol. 4). Elsevier.

Campillo, F., Joannides, M., & Larramendy-Valverde, I. (2011). Stochastic modeling of the chemostat. *Ecological Modelling*, *222*(15), 2676–2689.

Dehling, H. G., Gottschalk, T., & Hoffmann, A. C. (2007). *Stochastic modelling in process technology* (Vol. 211). Elsevier.

Enderle, J. D. (2012). Biochemical Reactions and Enzyme Kinetics. In *Introduction to biomedical engineering* (pp. 447–508). Elsevier.

Fontbona, J., Riquelme, V., & Silva, F. J. (2017). Stochastic modeling and control of bioreactors. *IFAC-PapersOnLine*, *50*(1), 12611–12616.

Fritsch, C., Harmand, J., & Campillo, F. (2015). A modeling approach of the chemostat. *Ecological Modelling*, *299*, 1–13. Retrieved from `http://dx.doi.org/10.1016/j.ecolmodel.2014.11.021` doi: 10.1016/j.ecolmodel.2014.11

.021

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

González-Figueredo, C., Flores-Estrella, R. A., & Rojas-Rejón, O. A. (2018). Fermentation: Metabolism, kinetic models, and bioprocessing. *Current topics in biochemical engineering*, 1–17. Retrieved from https://www.intechopen.com/books/current-topics-in-biochemical-engineering/fermentation-metabolism-kinetic-models-and-bioprocessing

Imhof, L., & Walcher, S. (2005). Exclusion and persistence in deterministic and stochastic chemostat models. *Journal of Differential Equations*, *217*(1), 26–53. doi: 10.1016/j.jde.2005.06.017

Källén, A. (2018). Lecture 7 Biomathematics (FMAN01). *Spring*. Retrieved from http://www.ctr.maths.lu.se/matematiklth/personal/andersk/kurser/biomath2018/index.html

Kang, H.-W., KhudaBukhsh, W. R., Koeppl, H., & Rempała, G. A. (2019). Quasi-steady-state approximations derived from the stochastic model of enzyme kinetics. *Bulletin of mathematical biology*, *81*(5), 1303–1336.

Konak, A. R. (1974). Derivation of a generalised Monod equation and its application. *Journal of Applied Chemistry and Biotechnology*, *24*(8), 453–455.

Lawson, M. J., Petzold, L., & Hellander, A. (2015). Accuracy of the Michaelis–Menten approximation when analysing effects of molecular noise. *Journal of The Royal Society Interface*, *12*(106), 20150054.

Lecca, P. (2013). Stochastic chemical kinetics. *Biophysical reviews*, *5*(4), 323–345.

Liu, Y. (2006). A simple thermodynamic approach for derivation of a general Monod equation for microbial growth. *Biochemical Engineering Journal*, *31*(1), 102–105. doi: 10.1016/j.bej.2006.05.022

Lonsdale, R., Harvey, J. N., & Mulholland, A. J. (2012, apr). A practical guide to modelling enzyme-catalysed reactions. *Chemical Society reviews*, *41*(8), 3025–3038. Retrieved from https://pubmed.ncbi.nlm.nih.gov/22278388https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3371381/ doi: 10.1039/c2cs15297e

Marchetti, L., Priami, C., & Thanh, V. H. (2017). Stochastic Simulation of Biochemical Reaction Systems. In *Simulation algorithms for computational systems biology* (pp. 7–28). Springer.

McQuarrie, D. A. (1967). Stochastic approach to chemical kinetics. *Journal of applied probability*, *4*(3), 413–478.

Michaelis, L., & Menten, M. L. (1913). Die kinetik der invertinwirkung. *Biochem. z*,

49(333-369), 352.

Michaelis, L., & Menten, M. L. (2007). *Die kinetik der invertinwirkung*. Universitätsbibliothek Johann Christian Senckenberg.

Monod, J. (1942). Recherches sur la croissance des cultures bacteriennes.

Murugan, R. (2018). *Theory on the rate equation of Michaelis–Menten type single-substrate enzyme catalyzed reactions* (Vol. 56) (No. 2). Springer International Publishing. doi: 10.1007/s10910-017-0791-3

Ortega Quintana, F. A., Alvarez, H., & Botero Castro, H. A. (2017). Enfrentando el modelado de bioprocesos: una revisión de las metodologías de modelado. *Revista ION*, *30*(1).

Rao, C. V., & Arkin, A. P. (2003). Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *The Journal of chemical physics*, *118*(11), 4999–5010.

Rincón, L. (2006). Introducción a las ecuaciones diferenciales estocásticas. *UNAM. México*.

Rogers, A., & Gibon, Y. (2009). Enzyme kinetics: theory and practice. In *Plant metabolic networks* (pp. 71–103). Springer.

Sanft, K. R., Gillespie, D. T., & Petzold, L. R. (2011). Legitimacy of the stochastic michaelis–menten approximation. *IET systems biology*, *5*(1), 58–69.

Smith, J. E. (2002). Biotechnology Third edition.

Stephanopoulos, G., Aris, R., & Fredrickson, A. G. (1979). A stochastic analysis of the growth of competing microbial populations in a continuous biochemical reactor. *Mathematical Biosciences*, *45*(1-2), 99–135.

Tochampa, W., Sirisansaneeyakul, S., Vanichsriratana, W., Srinophakun, P., Bakker, H. H. C., & Chisti, Y. (2005). A model of xylitol production by the yeast Candida mogii. *Bioprocess and Biosystems Engineering*, *28*(3), 175–183. doi: 10.1007/s00449-005-0025-0

Toral, R., & Colet, P. (2014). *Stochastic numerical methods: an introduction for students and scientists*. John Wiley & Sons.

Tsimrng, L. S. (2014). Noise in biology. *Reports on Progress in Physics*, *77*(2), 26601.

Uusitalo, L., Lehikoinen, A., Helle, I., & Myrberg, K. (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling Software*, *63*, 24–31.

Villadsen, J., Nielsen, J., & Lidén, G. (2011). *Bioreaction engineering principles*. Springer Science & Business Media.

Warne, D. J., Baker, R. E., & Simpson, M. J. (2019). Simulation and inference al-

gorithms for stochastic biochemical reaction networks: from basic concepts to state-of-the-art. *Journal of the Royal Society Interface*, *16*(151), 20180943.

Wilkinson, D. J. (2006). *Stochastic modelling for systems biology* (First Edit ed.). CRC press.

Wilkinson, D. J. (2018). *Stochastic modelling for systems biology* (Third Edit ed.). CRC press.

Wu, J., Vidakovic, B., & Voit, E. O. (2011). Constructing stochastic models from deterministic process equations by propensity adjustment. *BMC Systems Biology*, *5*(1), 187.

Chapter 2

Modeling and uncertainty assessment for enzymatic reactions through stochastic kinetic simulations

Abstract

The uncertain nature of biological systems is well known, however, it is interesting how pure deterministic modeling approaches have been majorly used for process system engineering. As an alternative, stochastic simulation of enzymatic kinetics can bring more insights compared to deterministic simulation. This work investigates the stochastic simulation for mass action and Michaelis-Menten kinetic models applied to generic examples of four enzymatic kinetic cases, expanding previous efforts to kinetics with and without inhibition. For each case, the deterministic kinetic model was translated to a stochastic formulation based on numbers of molecules. The simulations were performed applying the Gillespie algorithm, and the uncertainty of the results was estimated with the standard deviation, as a function of the number of realizations (repetitions) of the stochastic simulations. The uncertainty of the evolution of the state is related to the number of molecules used in the simulation settings. Through a sensitivity analysis, it was found a minimum number of realizations required to estimate the standard deviation favoring the computational requirements for the simulations. Par-

ticularly, it was found that the uncertainty of results can vary through the simulation due to the stochastic nature of the method. The stochastic kinetic modeling approach has shown to be a powerful tool to address biological models uncertainty alternative to correlation analysis of linear systems or Monte-Carlo approach and thus it could be further exploited for process system engineering design under uncertainty.

2.1 Introduction

The study of enzymes for specific applications using protein engineering has led to a growing interest in the development of realistic mathematical models for enzymatic reactions due to its application the in design, optimization, and control of industrial bioprocesses (Chapman, Ismail, & Dinu, 2018). Commonly, enzymatic reaction evolution has been simulated using mechanistic kinetics through the solution of ordinary differential equations. The deterministic approach is limited to a phenomenological understanding of kinetics, and the process uncertainty is addressed through a correlation analysis (linear case) or Monte-Carlo approach for non-linear models, accounting for the experimental data uncertainty or the expected molecular noise of the system (intrinsic noise). In the last two decades, there has been a growing interest in the application of stochastic approximations in biological systems. This has been motivated by the ability of these models to describe fluctuations, internally random processes (small chemical systems) without having any deterministic trend (Lente, 2013; Érdi & Lente, 2014; Wilkinson, 2018). Theoretical and mathematical aspects of stochastic enzymatic kinetics have been studied, including the appearance of fluctuations, the functions of the probability density, ergodicity, and approximation, inference methods, cellular noise associated with biological processes, and the noise patterns of experimental errors have been simulated (Kou, Cherayil, Min, English, & Xie, 2005; Qian & Elson, 2002; Schnoerr, Sanguinetti, & Grima, 2017; Tsimrng, 2014; Yang & Jiang, 2016). Previous results suggest that the uncertainty of experimental measurements can be replicated by applying stochastic methods based mainly on Gillespie's Stochastic Simulation Algorithm, or SSA, to chemical reactions (Gillespie, 1977, 2007; Lecca, 2013; Marchetti, Priami, & Thanh, 2017). An adaptation of the SSA to the occurrence of rare events was developed for the stiff enzyme-substrate reaction as a slow-scale SSA (Cao, Gillespie, & Petzold, 2005; Sanft, Gillespie, & Petzold, 2011). The validity of the adaptation of QSSA (a basis of the Michaelis-Menten model) to enzyme-catalyzed reactions has been covered (Barik, Paul, Baumann, Cao, & Tyson, 2008; Dhatt & Banerjee, 2019; Dóka, 2012; Kang, KhudaBukhsh, Koeppl, & Rempała, 2019; Rao & Arkin, 2003).

And stochastic enzyme kinetics has been included in reaction-diffusion simulations (Basu & Mohanty, 2009; Yi & Liu, 2010). However, the literature has only focused on the study of a single kinetic case, leaving aside very common industrial cases such as enzyme inhibitions. Hence, the stochastic approach is not a replacement for the traditional deterministic approach, but a new proposal to incorporate random aspects of this type of system that deterministic models cannot handle, for example, the main theme of this work, uncertainty propagation.

In this work, we propose to use stochastic kinetic models to include the uncertainty associated with the random behavior of the molecules in enzymatic processes allowing a prediction of the uncertainty propagation through the model. Previous efforts modeling enzymatic stochastic kinetics are extended to offer a detailed implementation of the SSA with both Law of Mass Action (LMA) or the Michaelis-Menten (MM) models, including for each one the case without inhibition (WI) and the three common types of enzymatic inhibition: competitive (CI), non-competitive (NCI), and un-competitive (UCI) (J. Donoso, 2006; Rogers & Gibon, 2009). The Figure 2.1 represents the different reaction routes from reactants, through unstable complexes until reaching products, depending on the case of the enzymatic reaction being studied (WI, CI, NCI, or ICU). Notice that the arrow (→) indicates a direct reaction, while the double arrow (↔) indicates reversible reactions.

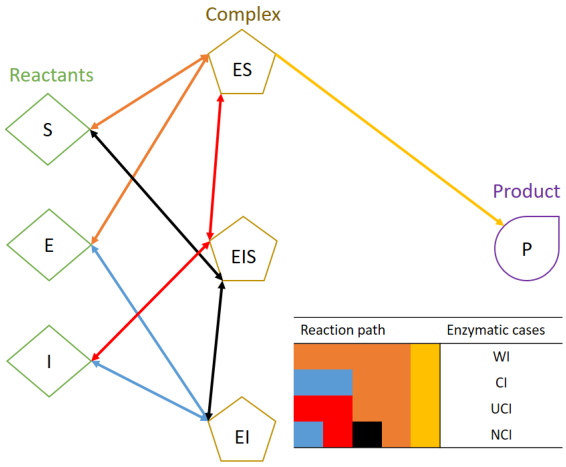

Figure 2.1. Chemical paths of the different cases of enzymatic reaction. Note: the color lines are referenced in the table to indicate the direct or reversible reaction path.

Particularly, practical aspects of setting a stochastic simulation are investigated

through a sensitivity analysis since this aspect has been overlooked in previous research. This provides useful insights to get better simulation times versus the number of runs (so-called realizations) and the selection of a system volume. Additionally, this analysis allows having an estimation of the system uncertainty from the standard deviation analysis. Notice, in literature the application of stochastic modeling is restricted to a very small-scale (namely cell size). In this research, the proposed approach allows using larges volumes aligned with conventional chemical systems.

The paper is structured as follows: theoretical and mathematical aspects of the deterministic and stochastic model are presented in section 2.2 and supplementary information is shown in the appendix 2.5. In the results section (section 2.3), the findings of the sensitivity analysis are addressed depicting simulation results (section 2.3.1), this is followed by the uncertainty assessment for the four enzymatic kinetics using both MM and LMA approaches (sections 2.3.2 and 2.3.3, respectively). Finally, the conclusions are drawn. This work represents the efforts to present an alternative of enzymatic systems modeling that could be used for process design, control, and optimization under uncertainty in the field of PSE, complementing the traditional vision of this type of systems.

Table 2.1. Nomenclature and symbols.

Greek letters	
ξ	Random number
κ	Reaction molecularity
τ	Stochastic step time
ν	Stoichiometric coefficient
Acronyms	
CI	Competitive Inhibition
E	Enzyme
EI	Enzyme-Inhibitor complex
EP	Enzyme-Product complex
ES	Enzyme-Substrate complex
EIS	Enzyme-Inhibitor-Substrate
I	Inhibitor
LMA	Law of Mass Action
MM	Michaelis-Menten
NCI	Non-Competitive Inhibition
P	Product
QSSA	Quasi-Steady State Approximation
S	Substrate
sdev	Standard deviation
SSA	Stochastic Simulation Algorithm
UCI	Un-Competitive Inhibition
WI	Without Inhibition
Variables	
a	Propensity
K	Equilibrium constant
t	Time
Subscripts	
C_i	Concentration of species i
c_j	Propensity constant
k_j	Kinetic constant, reaction j
N_i	Number of i molecules
r_j	Reaction rate

2.2 Mathematical models and methods

The deterministic simulations are based on the integration of differential equations obtained from the mass balance for a batch reactor. For the LMA cases, the ODEs correspond to the concentration of each participating chemical species, namely E, S, P, I, and complexes (Table 2.1); while in the MM cases, the state corresponding to the substrate concentration while product concentration is calculated stoichiometrically.

In the stochastic LMA simulations (SSA-LMA) chemical kinetics are decomposed into reaction channels that define the simulation events and their propensity functions. For Michaelis-Menten (SSA-MM), the only reaction channel is the transformation of the substrate into the product, and the right-hand side of the ODE defines the propensity function. The translation of mass-action expressions and ODEs into functions of numbers of molecules for SSA is covered in the supplementary material (appendix 2.5) and details of the implementation are depicted in the following sections. On the other hand, a sensitivity analysis of the number of realizations and the initial number of molecules was carried out, to determine the better conditions for the stochastic simulation and thus establish the uncertainty of the model by varying these parameters.

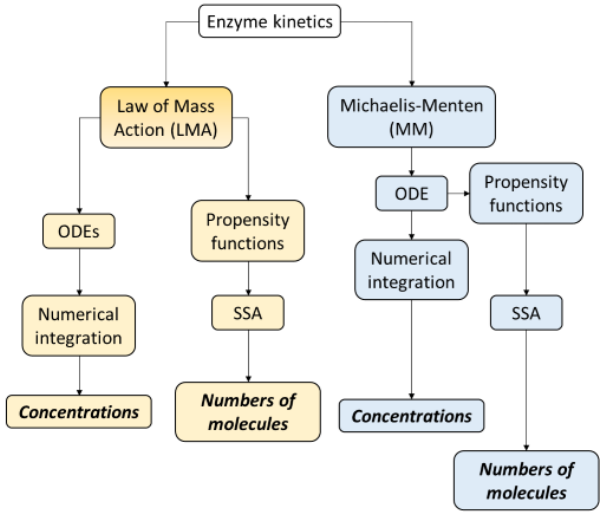

Figure 2.2. Models and simulation methods.

2.2.1 Deterministic simulation

In the mass-action law kinetics, the mass balances of a batch reactor have the form:

$$\frac{dC_i}{dt} = \sum_j \nu_{i,j} r_j \tag{2.1}$$

where r_j is the reaction rate j and $\nu_{i,j}$ the stoichiometric coefficient of species i in reaction j. Following the law of mass action rates are proportional to the product of

the reactant concentrations raised to their stoichiometric coefficients

$$r_j = k_j \prod_i C_i^{|\nu_{i,j}|} \tag{2.2}$$

where i represents reactants. For the sake of brevity the resulting ODEs for all chemical species (S, E, ES, I, P) are summarized in section 2.5.1.

On the other hand, the MM model applies the standard QSSA, considering chemical species and enzyme complexes transient and highly reactive in such a way that their net reaction rates are set equal to zero (Briggs & Haldane, 1925; Rao & Arkin, 2003). The ODEs for the four cases (WI, CI, NCI, and UCI) are summarized in appendix 2.5.2. The system of differential equations was integrated using the ode15s function of the software Matlab 2019b due to the stiff nature of the equations.

2.2.2 Stochastic Simulation Algorithm

Originally, stochastic simulations were developed to solve the chemical master equation, an exceedingly complex procedure, since it generates an ODE for each possible combination of reactant molecules (Gillespie, 2007). For instance, for 200 molecules, there are a million different molecular combinations, that is, a million ODEs (Rao & Arkin, 2003). Alternatively, the SSA is not based on ODEs. Instead, it is a stochastic method that defines the evolution of the system one reaction event at a time in intervals that depends on the propensity of the reaction, or reactions (the probability that a reaction occurs in a determined time) (Gillespie, 2007). For example, for the reaction

$$\mathrm{E} + \mathrm{S} \xrightarrow{k_1} \mathrm{ES} \tag{2.3}$$

the numbers of molecules are upgraded at each step of the SSA according to

$$N_\mathrm{E} \rightarrow N_\mathrm{E} - 1 \tag{2.4}$$

$$N_\mathrm{S} \rightarrow N_\mathrm{S} - 1 \tag{2.5}$$

$$N_\mathrm{ES} \rightarrow N_\mathrm{ES} + 1 \tag{2.6}$$

(see Eq. 2.20 and Table 2.4), with propensity $a_1 = c_1 N_\mathrm{E} N_\mathrm{S}$. In general, the SSA takes the following steps (Gillespie, 1977, 2007; Lecca, 2013):

1. Calculate the propensity a of the event from the numbers of molecules.

2. Use the propensity to define the time τ to the next event

$$\tau = -\frac{\ln(\xi)}{a} \tag{2.7}$$

where ξ is a random number generated in the interval $[0, 1]$.

3. If there are multiple reactions (i.e. *reaction channels*) a becomes the sum of all propensities

$$a = \sum_j a_j, \tag{2.8}$$

where the subscript j stands for each reaction. The next reaction event is chosen according to the relative weight of the propensities, that is a_i/a. A second random number ξ_2 is generated, and the index of the next reaction event corresponds to the smallest J such that

$$\xi_2 < \sum_{j=1}^{J} \frac{a_j}{a}. \tag{2.9}$$

4. Update the time, i.e. $t := t + \tau$, and the molecule numbers.

5. Repeat until reaching the simulation final time, or running out of reactant molecules.

2.2.3 Stochastic Chemical Kinetics (SCK)

Only direct reactions ($\rightarrow$) are allowed in the stochastic formulation of chemical kinetics, but reversible reactions ($\rightleftarrows$) are described with separate reaction channels, one for each direction (Martínez Urreaga, Mira, & González Fernández, 2003). For example, enzyme kinetics without inhibition,

$$\mathrm{E} + \mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \mathrm{ES} \xrightarrow{k_3} \mathrm{E} + \mathrm{P}, \tag{2.10}$$

becomes a combination of three reaction channels: $\xrightarrow{k_1}$, $\xleftarrow{k_2}$, and $\xrightarrow{k_3}$.

In the SSA-LMA, the occurrence probability of a reaction event depends on the number of possible combinations of reacting molecules, therefore the propensity function (a) to implement the SSA is described as (Lecca, 2013; Wilkinson, 2018):

$$a_j = c_j \prod_i \binom{N_i}{\nu_{i,j}} \tag{2.11}$$

where N is the number of molecules and

$$\binom{N_i}{\nu_{i,j}} = \frac{N_i!}{\nu_{i,j}! \cdot (N_i - \nu_{i,j})!}. \tag{2.12}$$

The propensity constant c_j can be obtained from the kinetic constant k_j in Eq. (2.2) (Lecca, 2013).

$$c_j = \frac{k_j}{(N_{\mathrm{Av}} V)^{\kappa - 1}} \tag{2.13}$$

where κ is the molecularity of the reaction, N_{Av} the Avogadro's number, and V the volume of the system (Lecca, 2013). It is illustrated with Eq. (2.10), for an unimolecular reaction channel, such as $\mathrm{ES} \xrightarrow{k_3} \mathrm{E} + \mathrm{P}$,

$$a_3 = c_3 N_{\mathrm{ES}}$$
$$c_3 = k_3,$$

for a bimolecular reaction channel, such as $\mathrm{E} + \mathrm{S} \xrightarrow{k_1} \mathrm{ES}$,

$$a_1 = c_1 N_{\mathrm{E}} N_S$$
$$c_1 = \frac{k_1}{(N_{\mathrm{Av}} V)}.$$

Table 2.4 in Appendix 2.5 resumes the propensity, molecularity, and molecule number changes for the nine reaction channels that appear in the reactions of the LMA model cases.

On the other hand, the stochastic version of the MM kinetics includes only the reaction channel $\mathrm{S} \rightarrow \mathrm{P}$, implying that each step of the SSA algorithm updates the numbers of product and substrate molecules as $N_{\mathrm{P}} \rightarrow N_{\mathrm{P}} + 1$ and $N_{\mathrm{S}} \rightarrow N_{\mathrm{S}} - 1$. Unfortunately, the propensity cannot be obtained from Eq. (2.11) as it does not follow the law of mass action. This issue was addressed by Rao and Arkin applying the CME to translate the ODE of the WI case

$$\frac{dC_{\mathrm{S}}}{dt} = -\frac{k_3 C_{\mathrm{E},0} C_{\mathrm{S}}}{C_{\mathrm{S}} + K_{\mathrm{ES}}} \tag{2.14}$$

into an *equivalent* propensity (Rao & Arkin, 2003). They proposed the same translation for the competitive inhibition (CI) case, and the translation of kinetic ODEs into propensities has become common in subsequent studies of stochastic Michaelis-Menten-QSSA (Agarwal, Adams, Castellani, & Shouval, 2012; Kang et al., 2019; Lawson, Petzold, & Hellander, 2015; Rao & Arkin, 2003; Sanft et al., 2011; Warne, Baker, & Simpson, 2019; Wu, Vidakovic, & Voit, 2011). Herein, this approach was applied

using the fact that the system volume is constant to replace the concentrations with the numbers of molecules by doing

$$C_i = N_i \left(\frac{C_{\mathrm{S},0}}{N_{\mathrm{S},0}} \right) \tag{2.15}$$

where $C_{\mathrm{S},0}$ and $N_{\mathrm{S},0}$ are substrate's initial concentration and number of molecules, respectively. In this way, Eq. (2.24) becomes

$$\frac{dN_{\mathrm{S}}}{dt} = -k_3 \frac{N_{\mathrm{E},0} N_{\mathrm{S}}}{N_{\mathrm{S}} + K_{\mathrm{ES}} \left(\frac{N_{\mathrm{S},0}}{C_{\mathrm{S},0}} \right)}$$

which leads to the propensity for Michaelis-Menten kinetics without inhibition

$$a_{\mathrm{WI}} = k_3 \frac{N_{\mathrm{E},0} N_{\mathrm{S}}}{N_{\mathrm{S}} + K_{\mathrm{ES}}^N} \tag{2.16}$$

where

$$K_{\mathrm{ES}}^N = K_{\mathrm{ES}} \left(\frac{N_{\mathrm{S},0}}{C_{\mathrm{S},0}} \right).$$

The same transformation leads to the propensities for all cases, WI, CI, NCI, and UCI; see Sec. 2.5.2.

Table 2.2. Parameters and initial conditions for LMA examples .

Parameter	Value	Parameter	Value
k_1	$10^6(\mathrm{M}\cdot\mathrm{h})^{-1}$	k_9	$5\cdot10^{-5}\mathrm{h}^{-1}$
k_2	$10^{-4}\mathrm{h}^{-1}$	$C_{\mathrm{S},0}$	$5\cdot10^{-7}\mathrm{M}$
k_3	$0.1\mathrm{h}^{-1}$	$C_{\mathrm{E},0}$	$2\cdot10^{-7}\mathrm{M}$
k_4	$10^5(\mathrm{M}\cdot\mathrm{h})^{-1}$	$C_{\mathrm{I},0}$	$2\cdot10^{-7}\mathrm{M}$
k_5	$10^{-5}\mathrm{h}^{-1}$	$N_{\mathrm{S},0}$	300
k_6	$5\cdot10^5(\mathrm{M}\cdot\mathrm{h})^{-1}$	$N_{\mathrm{E},0}$	120
k_7	$5\cdot10^{-5}\mathrm{h}^{-1}$	$N_{\mathrm{I},0}$	120
k_8	$5\cdot10^5(\mathrm{M}\cdot\mathrm{h})^{-1}$	V	$10^{-15}\mathrm{L}$

1 M=1 mol/L. Wilkinson (2018)

Table 2.3. Parameters and initial conditions for Michaelis-Menten simulations.

Parameter	Value	Parameter	Value	Parameter	Value
k_3	$1\mathrm{h}^{-1}$	$C_{\mathrm{S},0}$	10M	$N_{\mathrm{S},0}$	1000
K_{ES}	11M	$C_{\mathrm{E},0}$	1M	$N_{\mathrm{E},0}$	100
K_{EI}	1M	$C_{\mathrm{I},0}$	1M	$N_{\mathrm{I},0}$	100

1 M=1 mol/L. Sanft et al. (2011).

It is worth mentioning that this conversion has been published before in such a way that the K_{ES} value becomes the same in the kinetic ODE and the propensity (Rao & Arkin, 2003; Sanft et al., 2011; Wu et al., 2011). It is achieved by setting $(N_{\mathrm{Av}}V) = 1\mathrm{L/mol}$, however, this fixed unitary value of $N_{\mathrm{Av}}V$ is not a requisite of the procedure.

2.2.4 Simulation settings

Initial conditions for deterministic simulations are conventionally known from the experimental setup. However, stochastic simulations are based on numbers of molecules and given that

$$N_{\mathrm{Av}}V = \frac{N_{i,0}}{C_{i,0}} \tag{2.17}$$

(being V the system volume and N_{Av} Avogadro's number). It is possible to either define the simulation volume and an initial concentration (V, $C_{i,0}$) to get the numbers of molecules, or an initial number of molecules and concentration ($N_{i,0}$, $C_{i,0}$) to set the value of $(N_{\mathrm{Av}}V)$.

The settings for the LMA stochastic simulations were adapted from available examples calculating the initial numbers of molecules from the initial concentration values with a fixed volume of $V = 10^{-15}\mathrm{L}$, see Table 2.2, Eq. (2.17) (Wilkinson, 2018). These settings aim to represent a microorganism's volume, a diluted system with very low initial concentrations of the species. However, our purpose is to illustrate the numerical

simulation methods rather than simulate a living organism, for this reason, the volume of a specific microorganism was not used. Readers interested in molecular and cell biology simulations can use the BioNumbers database, which includes parameters for specific microorganisms (Milo, Jorgensen, Moran, Weber, & Springer, 2010).

In the Michaelis-Menten stochastic simulations, the initial state of the system was defined with the numbers of molecules and initial concentrations listed in Table 2.3. This indirectly sets the volume of the system through Eq. (2.17), although the value of V is not necessary for the stochastic simulations. Also, parameters were chosen considering that the QSSA is only suitable when the initial concentration of the enzyme is almost negligible compared to the sum of the initial concentration of substrate plus the equilibrium constant, i.e. $C_{\mathrm{E},0} << C_{\mathrm{S},0} + K_{\mathrm{ES}}$, or

$$N_{\mathrm{E},0} << N_{\mathrm{S},0} + K_{\mathrm{ES}}^{N} \tag{2.18}$$

because the stochastic MM becomes accurate when the same deterministic validity conditions are fulfilled (Kang et al., 2019; Lawson et al., 2015; Sanft et al., 2011; Segel, 1988; Segel & Slemrod, 1989).

The simulations were run in the software Matlab 2019b on a laptop with an Intel Core i5 processor and 4Gb of RAM, with which an execution time of less than one hour was obtained for the most demanding case with 50 000 realizations of the MM-WI model and 100 000 initial molecules.

2.3 Results

2.3.1 Sensitivity analysis

Notice that for each stochastic simulation result, a so-called *"realization"*, is different, although the path dictated by the kinetics through the reaction propensity is the same in the stochastic simulation algorithm. Due to this, and in the same way that concentration measurements have a relative uncertainty $\Delta C/C$, the SSA carries an inherent uncertainty due to the change of the number of molecules (N_i) at each step of the Gillespie algorithm. For example, if the number of i molecules is reduced by one the relative change is $1/N_i$. Thus, given that the simulation volume is constant, Eq. (2.17) leads to

$$\frac{\Delta C_i}{C_i} = \frac{1}{N_i}, \tag{2.19}$$

Note: The ability of stochastic models to simulate extrinsic and intrinsic uncer-

tainty has already been validated for biological processes (Székely Jr & Burrage, 2014; Tsimrng, 2014). Therefore, it is proposed to make use of the relationship between the relative experimental uncertainty to calculate the number of initial molecules to define the size of the system (Eq. 2.19).

implying that the SSA can emulate the order of magnitude of concentration uncertainty by setting the (initial) numbers of molecules. This is illustrated in Figure 2.3 when the integration of the kinetic differential equation for the enzymatic reaction with non-competitive inhibition is superimposed within several realizations of the SSA.

The analogy between experimentation and stochastic simulation indicates that statistic measurements, mean and standard deviation can be obtained from SSA realizations. However, literature has not addressed how to define the numbers of molecules (system volume) and realizations without taking advantage of the uncertainty given by the stochastic model Romero-Severson, Ribeiro, and Castro (2018). In previous research, the number of realizations has been arbitrarily selected between 100 and 10 000 (Agarwal et al., 2012; Rao & Arkin, 2003; Sanft et al., 2011; Wu et al., 2011), to assess the effects of system size on stochastic Michaelis-Menten simulations and stochastic solutions of population balances (Lawson et al., 2015; Zhou, Jiang, & Chan, 2020). But even so, it is not clear how to systematically define those two parameters. To answer this question a sensitivity analysis is proposed to study the influence of the initial number of molecules and realizations.

Results shown in Figure 2.4, indicate that the logarithm of the standard deviation of the concentration decreases linearly with the initial number of substrate molecules ($N_{S,0}$). It corresponds to the fact that the stochastic simulation becomes a better approximation of the real system with higher values of N and converges to the deterministic solution when $N \to \infty$ (Hahl & Kremling, 2016; Lawson et al., 2015; Marchetti et al., 2017; Menz, 2013). Despite that, the emulation of large-scale chemical systems with a huge N value would lead to an unrealistic small uncertainty in Eq. (2.19). The selection of the volume of the system can be made based on: replicating the experimental uncertainty of the process (Eq. 2.19), the uncertainty around the experimental data, or even, a simulation volume equal to the volume of the microorganism in the case fermentation.

Regarding the number of realizations (R), the results in Figure 2.5 show that standard deviation converges to a stable value after 500 realizations for all $N_{S,0}$ values investigated. It implies that the probability distribution function becomes constant, in agreement with the full probability distribution corresponds to the limit $R \to \infty$ (Székely Jr & Burrage, 2014).

The results for the law of mass action model were analogous, so for both mod-

els, MM and LMA the statistics were calculated for $R = 1\,000$ realizations. Also, it is confirmed that the initial numbers of molecules proposed in Tables 2.2-2.3 are true, acceptable. It is worth mentioning a technical detail, given that Gillespie's algorithm produces results at random times generated by the propensities in Eq. (2.7), the averages and standard deviations were calculated from SSA results interpolated to 1 h intervals.

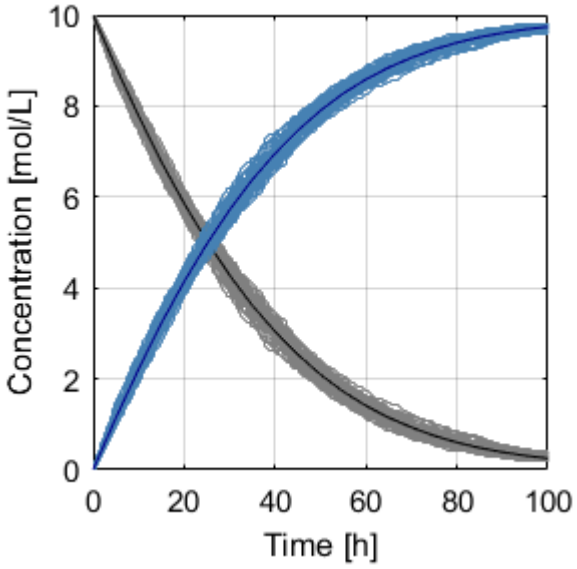

Figure 2.3. Deterministic (ODE based) and stochastic (SSA) simulations of Michaelis-Menten kinetics with non-competitive inhibition (MM-NCI, Eq. 2.28).
Black solid lines: substrate (S) and product (P) concentrations obtained from ode integration. Staircase lines: stochastic results, 100 realizations of Gillespie algorithm. Concentrations are calculated from the numbers of molecules using Eq. 2.15.

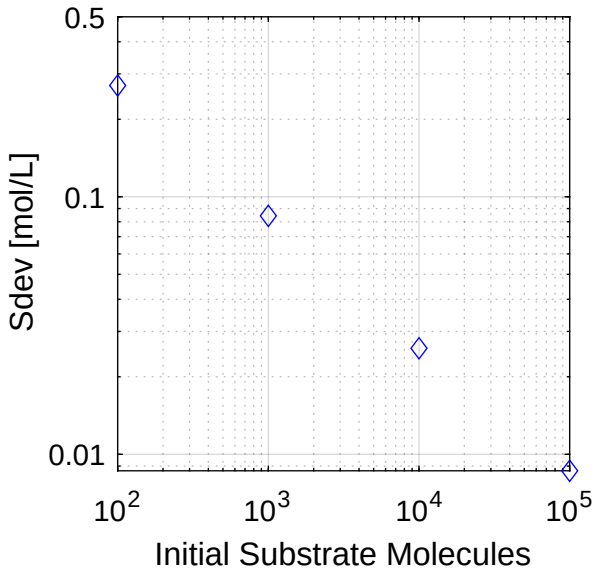

Figure 2.4. Standard deviation of concentration for different $N_{S,0}$ values (initial numbers of substrate molecules).
MM-WI kinetics with 1 000 realizations.Values are shown for t = 40h, results for other times and number of realizations are similar.

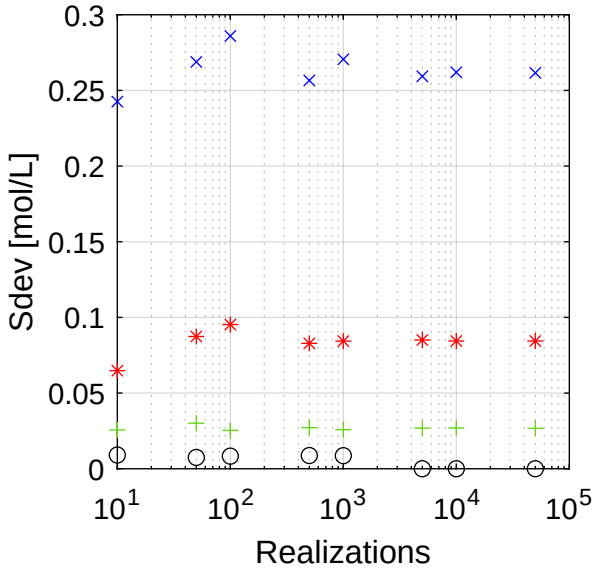

Figure 2.5. Standard deviation of concentration for different numbers of realizations. Initial numbers of substrate molecules for MM-WI kinetics (each marker is a different $N_{S,0}$, i.e. '×' =100, '∗'=1000, '+'=10000, '◯'=100000). Values are shown for $t = 40$h, results for other times are similar.

2.3.2 Michaelis-Menten stochastic and deterministic simulations

A comparison of stochastic and deterministic simulations for the four enzymatic kinetics are depicted in Figure 2.6 with the mean value and standard deviations calculated from the SSA, shown at 5 h intervals. The small standard deviation values suggest that the system is big enough to be statistically homogeneous (see Figure 2.7).

Despite not evident to the naked eye, uncertainty (sdev) is not constant through the simulation. It is null in the initial time because all realizations start from the same state, and tend to be zero in the final point where the molecules are depleted because each realization tends to this state at a similar t value. Figure 2.7, illustrates how sdev reaches a maximum at times in which the realizations present a greater dispersion of the value of N. A more detailed analysis is presented in the description of the Figure 2.7.

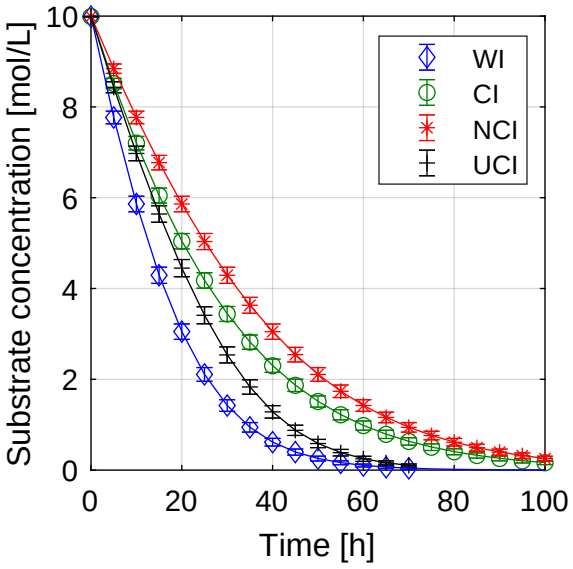

Figure 2.6. Stochastic and deterministic trajectories of substrate concentration with Michaelis-Menten kinetics.
Solid lines: deterministic model (ODE) results. Markers: average and standard deviation from stochastic simulation (SSA). Without inhibition the substrate depletes faster and the slope of $C(t)$ is higher for the WI simulation.

Figure 2.7 shows that in the beginning of the simulations, before the maximum, the results of the inhibited cases produce standard deviations lower than sdev for the WI case. It comes from the fact that inhibited reactions proceed at a slower rates (slopes in Figure 2.3) than the WI case. The small propensities associated to inhibition increase the times τ to the next reaction in the SSA, therefore few different states of the system, i.e. N values, correspond to each sampled time which leads to a lower variance (see τ, Eq. (2.7), and the term $1 + N_{\mathrm{I}}/K_{\mathrm{EI}}^{N}$ in the equations for a in Section 2.5.2). On the contrary, the order of sdev values becomes reversed in the long term because the WI propensity becomes smaller due to the substrate depletion. The transition between the two propensity regimes induces the maximum that appears in the figure

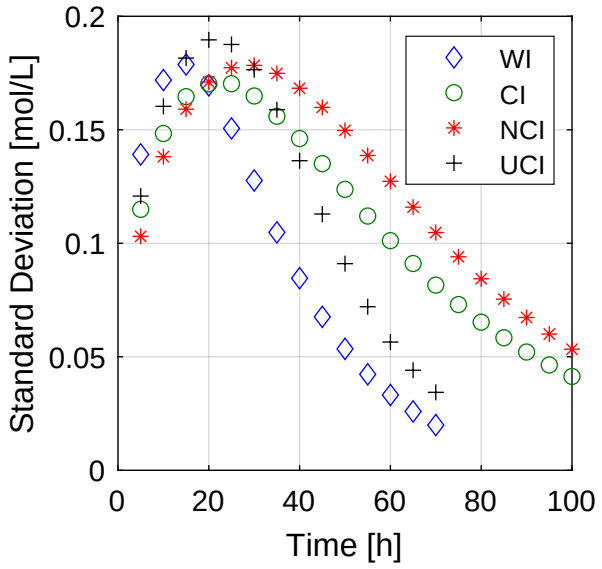

Figure 2.7. Concentration standard deviation from SSA simulations of Michaelis-Menten kinetics for S $\rightarrow$ P.

2.3.3 Law of mass action stochastic and deterministic simulations

Simulations of the four cases WI, CI, NI, and UCI, with both deterministic and stochastic methods (ODE integration and SSA), are shown in Figure 2.8-2.11. Mean and the standard deviation were calculated at 1 h intervals from the results of 1000 realizations. Although the plots of the stochastic simulation algorithm results are illustrated with a single realization for clarity purposes, also the plots include the ± one-standard-deviation intervals, which encloses the deterministic solution and gives an estimation of the uncertainty of the simulation provided by the stochastic model. For all cases, the simulation results were homogeneous, in the sense that the stochastic (Gillespie's predictions) results follow the deterministic concentration profiles. In the description of Fig. 2.8 - 2.11, the analysis are presented in the figures caption.

Figure 2.8 shows that substrate and product concentrations follow the same pattern observed in MM kinetics. Concentrations of enzyme and enzyme-substrate complex (E, EI) behave in opposite ways due to the initial consumption of enzyme to produce enzyme-substrate complex, which later decomposes to generate the product, recovering the enzyme concentration. The SSA reproduces correctly the behavior predicted by integration of the ODEs, including substrate's uptake, product's output, and E and EI concentrations.

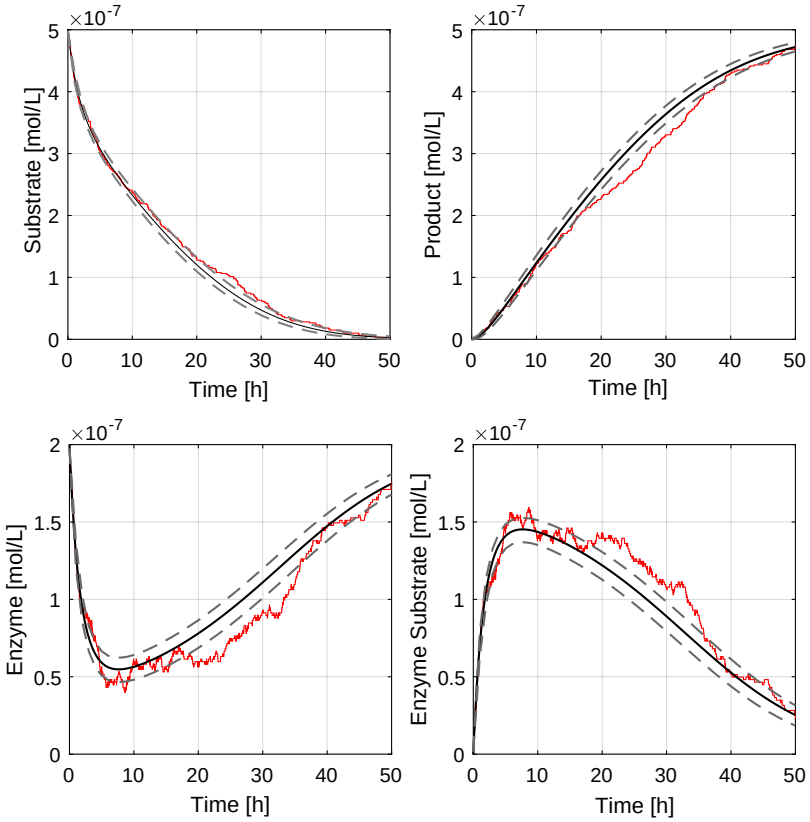

Figure 2.8. Law of mass action kinetics without inhibition (LMA-WI). Black solid line: deterministic model results. Staircase red line: results from a single realization of the SSA. Gray dashed lines: one-standard-deviation envelope, $x \pm \mathrm{sdev}\,(x)$.

In Figure 2.9, enzyme concentration initially decreases due to the formation of enzyme-substrate and enzyme-inhibitor complexes, however it later increases and peaks close to $t = 60$ due to the decomposition of the ES complex. Also it can be observed how after 60 hours the behavior of the system is governed by the inhibitor through enzyme consumption to form EI complex until the end of the simulation. On the other hand the inhibitor concentration decreases with time for the entire simulation due to the formation of the enzyme-inhibtor complex (EI), whose concentration increases in consequence. Standard deviations of substrate and product were smaller than the ones of the other species, perhaps due to the inhibition reducing the effect of the enzyme.

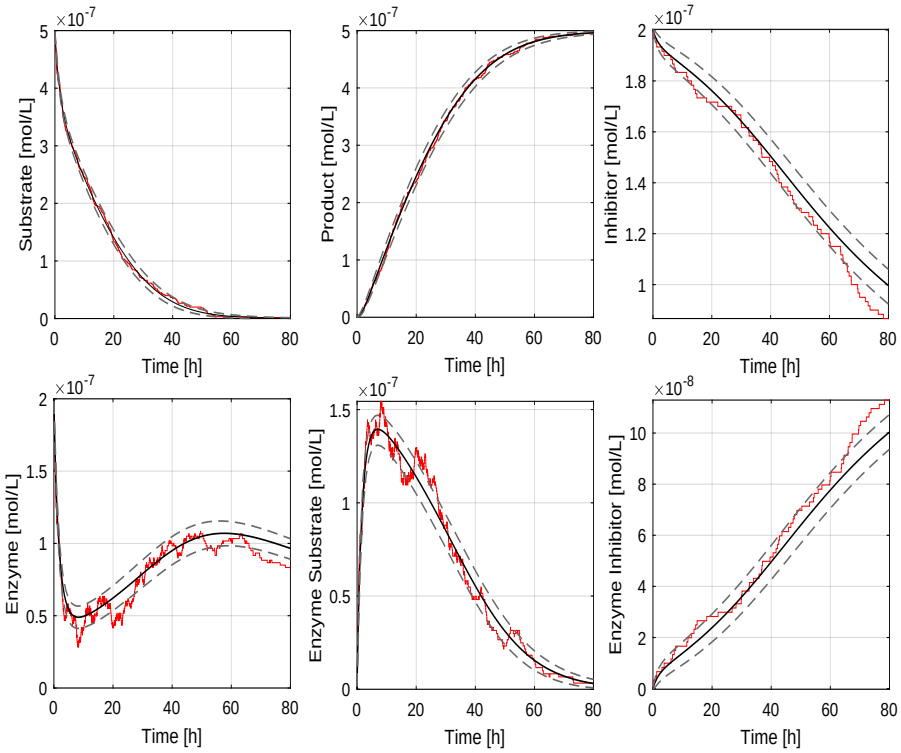

Figure 2.9. Law of mass action kinetics with competitive inhibition (LMA-CI). Solid line: deterministic model results. Staircase red line: results from a single realization of the SSA. Gray dashed lines are the one-standard-deviation envelope, $x \pm \text{sdev}(x)$.

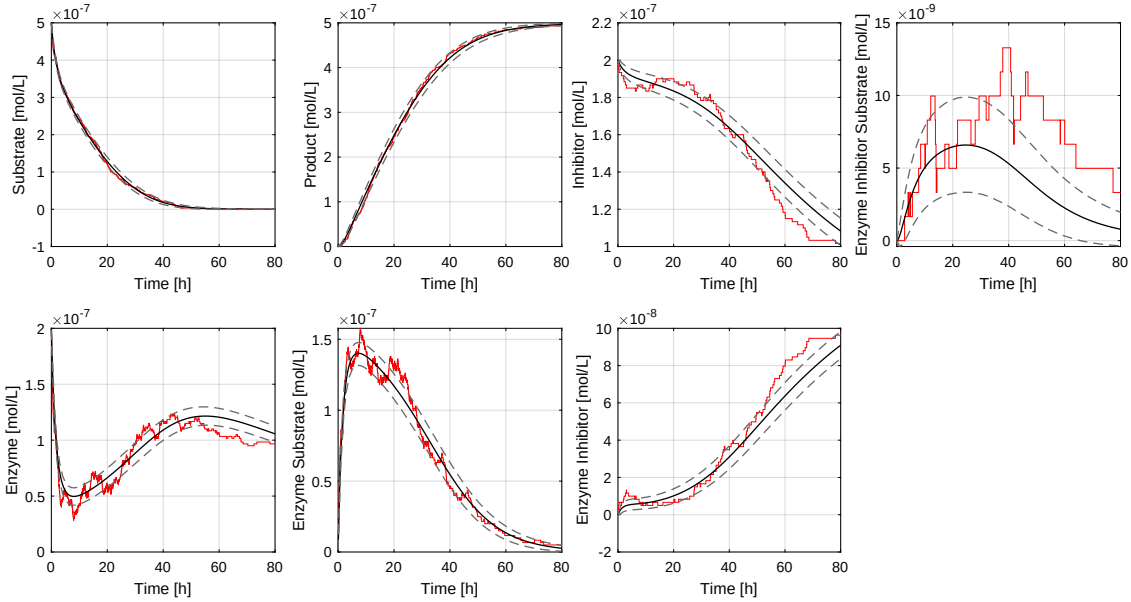

Figure 2.10. Law of mass action, non-competitive inhibition enzyme kinetics (LMA-NCI). Black solid line: deterministic model results. Staircase red line: results from a single realization of the SSA. Gray dashed lines are the one-standard-deviation envelope, $x \pm \mathrm{sdev}(x)$.

.

The most remarkable feature in the results presented in Figure 2.11 is that the standard deviation of the inhibitor species (I, EIS) increases with time to maximum values, contrary to the results in Figure 2.7. However, it can be explained by the fact that SSA realizations were terminated just before the N_{I} and N_{EIS} became invariant, causing each realization to end with a different value which increases sdev.

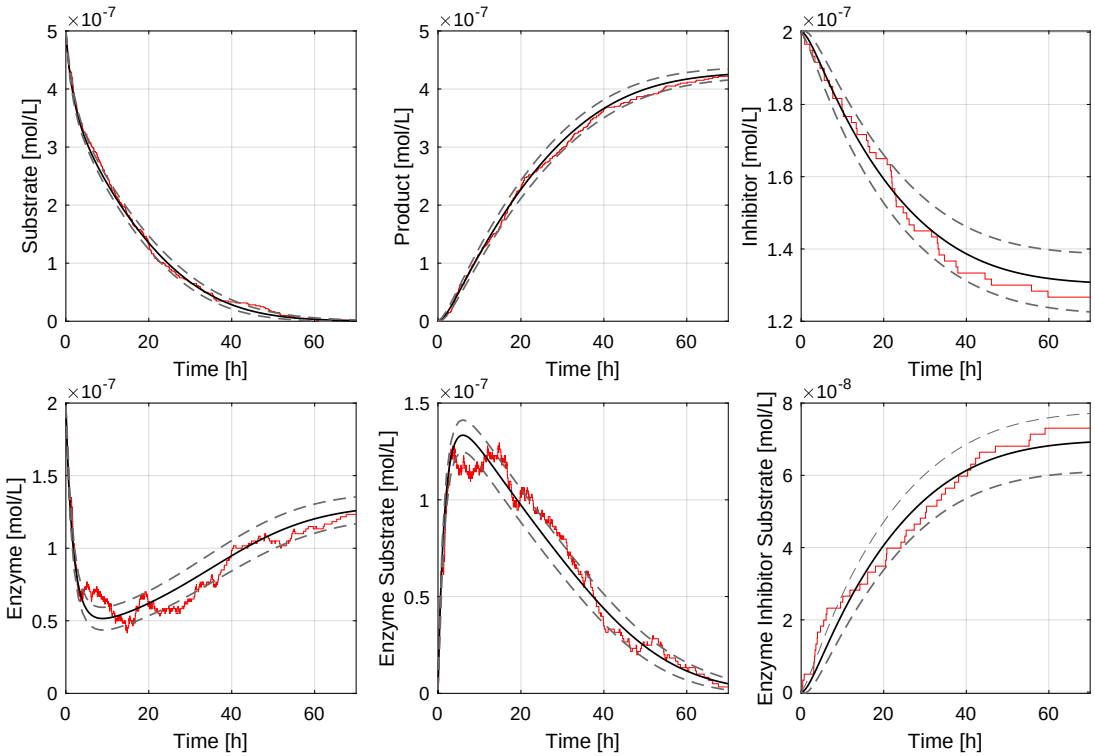

Figure 2.11. Law of mass action un-competitive inhibition enzyme kinetics (LMA-UCI). Black solid line: deterministic model results. Staircase red line: results from a single realization of the SSA. Gray dashed lines are the one-standard-deviation envelope, $x \pm \mathrm{sdev}(x)$.

In summary, the sensitivity analysis allowed it to make an optimal selection of the initial number of molecules, as well as of the realizations to obtain a more precise value of the uncertainty found from the stochastic simulations. Besides, the results of the stochastic approaches in comparison with the deterministic ones for both cases (LMA and MM), not only describe in the same way the behavior of the species involved but also provide information on the level of uncertainty of the systems and the variation of this throughout the simulation.

2.4 Conclusion

This work shows that the uncertainty of concentrations enzyme kinetics can be estimated with repeated realizations of the SSA stochastic simulation, in models based either on the law of mass action or the Michaelis-Menten kinetics and with and without inhibition. It is also possible to emulate the relative uncertainty of experimental measurements by setting the number of molecules. In this sense the SSA can be a useful complement of ODE-based simulations of enzyme kinetics, however, it should be considered that this simulated uncertainty varies with time, it tends to be higher when the variable is changing quickly, making it advisable to calculate it for the whole extension of the simulation. Furthermore, uncertainty estimation can be executed in reasonable times in a common computer, as the standard deviation results converge after a small number of realizations. We envision that in future works stochastic simulation could be applied to design and control processes under conditions of uncertainty.

2.5 Appendix: detailed models

2.5.1 Law of mass action

Differential equations for each case are shown first, while reactions channels for stochastic simulation are summarized in Table 2.4. Some reactions are repeated for clarity. Sources: (J. Donoso, 2006; Rogers & Gibon, 2009).

- *Without inhibition* (WI)

$$\mathrm{E}+\mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \mathrm{ES} \xrightarrow{k_3} \mathrm{E}+\mathrm{P} \tag{2.20}$$

Species (i)	$dC_i/dt = \cdots$
E	$-k_1 C_\mathrm{E} C_\mathrm{S} + (k_2 + k_3)\, C_\mathrm{ES}$
S	$-k_1 C_\mathrm{E} C_\mathrm{S} + k_2 C_\mathrm{ES}$
ES	$k_1 C_\mathrm{E} C_\mathrm{S} - (k_2 + k_3)\, C_\mathrm{ES}$
P	$k_3 C_\mathrm{ES}$

- *Competitive inhibition* (CI)

As in the previous case

$$\mathrm{E}+\mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \mathrm{ES} \xrightarrow{k_3} \mathrm{E}+\mathrm{P}$$

but an inhibitor is included

$$\mathrm{E}+\mathrm{I} \underset{k_5}{\overset{k_4}{\rightleftarrows}} \mathrm{EI}. \tag{2.21}$$

the ODEs for substrate and product remain the same as in the kinetics without inhibition, but the inhibitor reaction changes dC_E/dt and requires two new differential equations for I and the complex enzyme-inhibitor (EI)

Species (i)	$dC_i/dt = \cdots$
E	$-k_1 C_\mathrm{E} C_\mathrm{S} + (k_2 + k_3)\, C_\mathrm{ES} - k_4 C_\mathrm{E} C_\mathrm{I} + k_5 C_\mathrm{EI}$
S	$-k_1 C_\mathrm{E} C_\mathrm{S} + k_2 C_\mathrm{ES}$
ES	$k_1 C_\mathrm{E} C_\mathrm{S} - (k_2 + k_3)\, C_\mathrm{ES}$
P	$k_3 C_\mathrm{ES}$
I	$-k_4 C_\mathrm{E} C_\mathrm{I} + k_5 C_\mathrm{EI}$
EI	$k_4 C_\mathrm{E} C_\mathrm{I} - k_5 C_\mathrm{EI}$

- *Non-competitive inhibition* (NCI)

This case includes the product, inhibitor reactions (Eqs. 2.20-2.21)

$$\mathrm{E}+\mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \mathrm{ES} \overset{k_3}{\rightarrow} \mathrm{E}+\mathrm{P}$$

$$\mathrm{E}+\mathrm{I} \underset{k_5}{\overset{k_4}{\rightleftarrows}} \mathrm{EI}$$

and EIS (enzyme-inhibitor-substrate) complex reactions

$$\mathrm{EI}+\mathrm{S} \underset{k_7}{\overset{k_6}{\rightleftarrows}} \mathrm{EIS} \tag{2.22}$$

$$\mathrm{ES}+\mathrm{I} \underset{k_9}{\overset{k_8}{\rightleftarrows}} \mathrm{EIS}. \tag{2.23}$$

The additional EIS reactions lead to

Species (i)	$dC_i/dt = \cdots$
E	$-k_1 C_\mathrm{E} C_\mathrm{S} + (k_2 + k_3)\, C_\mathrm{ES} - k_4 C_\mathrm{E} C_\mathrm{I} + k_5 C_\mathrm{EI}$
S	$-k_1 C_\mathrm{E} C_\mathrm{S} + k_2 C_\mathrm{ES} - k_6 C_\mathrm{EI} C_\mathrm{S} + k_7 C_\mathrm{EIS}$
ES	$k_1 C_\mathrm{E} C_\mathrm{S} - (k_2 + k_3)\, C_\mathrm{ES} - k_8 C_\mathrm{ES} C_\mathrm{I} + k_9 C_\mathrm{EIS}$
P	$k_3 C_\mathrm{ES}$
I	$-k_4 C_\mathrm{E} C_\mathrm{I} + k_5 C_\mathrm{EI} + k_8 C_\mathrm{ES} C_\mathrm{I} - k_9 C_\mathrm{EIS}$
EI	$k_4 C_\mathrm{E} C_\mathrm{I} - k_5 C_\mathrm{EI} - k_6 C_\mathrm{EI} C_\mathrm{S} + k_7 C_\mathrm{EIS}$
EIS	$k_6 C_\mathrm{EI} C_\mathrm{S} - k_7 C_\mathrm{EIS} + k_8 C_\mathrm{ES} C_\mathrm{I} - k_9 C_\mathrm{EIS}$

- *Un-competitive inhibition* (UCI)

Product comes only from Eq. (2.20), while the reaction between complex ES and I is shown in Eq. (2.23).

$$\mathrm{E}+\mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \mathrm{ES} \overset{k_3}{\rightarrow} \mathrm{E}+\mathrm{P}$$

$$\mathrm{ES}+\mathrm{I} \underset{k_9}{\overset{k_8}{\rightleftarrows}} \mathrm{EIS}$$

Species (i)	$dC_i/dt = \cdots$
E	$-k_1 C_\mathrm{E} C_\mathrm{S} + (k_2 + k_3)\, C_\mathrm{ES}$
S	$-k_1 C_\mathrm{E} C_\mathrm{S} + k_2 C_\mathrm{ES}$
ES	$k_1 C_\mathrm{E} C_\mathrm{S} - (k_2 + k_3)\, C_\mathrm{ES} - k_8 C_\mathrm{ES} C_\mathrm{I} + k_9 C_\mathrm{EIS}$
P	$k_3 C_\mathrm{ES}$
I	$-k_8 C_\mathrm{ES} C_\mathrm{I} + k_9 C_\mathrm{EIS}$
EIS	$k_8 C_\mathrm{ES} C_\mathrm{I} - k_9 C_\mathrm{EIS}$

Stochastic kinetics for the law of mass action cases is reduced to the reaction channels in Table (2.4). It includes the propensities and molecule number changes.

Table 2.4. Propensities, reaction molecularities (K), and molecule number changes of reaction channels in the law of mass action examples. Channels necessary for each case: WI 1-3, CI 1-5, NCI 1-9, UCI 1-3 and 8-9.

Channel j	Reaction channel	a_j	K	E	I	P	S	EI	EP	ES	EIS
1	$\mathrm{E}+\mathrm{S} \xrightarrow{k_1} \mathrm{ES}$	$c_1 N_\mathrm{E} N_\mathrm{S}$	2	-1			-1			+1	
2	$\mathrm{ES} \xrightarrow{k_2} \mathrm{E}+\mathrm{S}$	$c_2 N_\mathrm{ES}$	1	+1			+1			-1	
3	$\mathrm{ES} \xrightarrow{k_3} \mathrm{E}+\mathrm{P}$	$c_3 N_\mathrm{ES}$	1	+1		+1				-1	
4	$\mathrm{E}+\mathrm{I} \xrightarrow{k_4} \mathrm{EI}$	$c_4 N_\mathrm{E} N_\mathrm{I}$	2	-1	-1			+1			
5	$\mathrm{EI} \xrightarrow{k_5} \mathrm{E}+\mathrm{I}$	$c_5 N_\mathrm{EI}$	1	+1	+1			-1			
6	$\mathrm{EI}+\mathrm{S} \xrightarrow{k_6} \mathrm{EIS}$	$c_6 N_\mathrm{EI} N_\mathrm{S}$	2				-1	-1			+1
7	$\mathrm{EIS} \xrightarrow{k_7} \mathrm{EI}+\mathrm{S}$	$c_7 N_\mathrm{EIS}$	1				+1	+1			-1
8	$\mathrm{ES}+\mathrm{I} \xrightarrow{k_8} \mathrm{EIS}$	$c_8 N_\mathrm{ES} N_\mathrm{I}$	2		-1					-1	+1
9	$\mathrm{EIS} \xrightarrow{k_9} \mathrm{ES}+\mathrm{I}$	$c_9 N_\mathrm{EIS}$	1		+1					+1	-1

2.5.2 Michaelis-Menten

Differential equations for substrate concentration (C_S), from refs. (Rogers & Gibon, 2009; Villadsen, Nielsen, & Lidén, 2011). Propensities come from the procedure described in Sec. 2.2.3. Some equations are repeated here for clarity:

- *Without Inhibition (WI)*

$$\frac{\mathrm{d}C_\mathrm{S}}{\mathrm{d}t} = -\frac{k_3 C_{\mathrm{E},0} C_\mathrm{S}}{C_\mathrm{S} + K_\mathrm{ES}} \tag{2.24}$$

$$a_\mathrm{WI} = k_3 \frac{N_{\mathrm{E},0} N_\mathrm{S}}{N_\mathrm{S} + K_\mathrm{ES}^N} \tag{2.25}$$

where

$$K_\mathrm{ES}^N = K_\mathrm{ES} \left(\frac{N_{\mathrm{S},0}}{C_{\mathrm{S},0}} \right)$$

- *Competitive Inhibition (CI)*

$$\frac{\mathrm{d}C_\mathrm{S}}{\mathrm{d}t} = -\frac{k_3 C_{\mathrm{E},0} C_\mathrm{S}}{C_\mathrm{S} + K_\mathrm{ES} \left(1 + \frac{C_\mathrm{I}}{K_\mathrm{EI}} \right)} \tag{2.26}$$

$$a_\mathrm{CI} = \frac{k_3 N_{\mathrm{E},0} N_\mathrm{S}}{N_\mathrm{S} + K_\mathrm{ES}^N \left(1 + \frac{N_\mathrm{I}}{K_\mathrm{EI}^N} \right)} \tag{2.27}$$

where

$$K_\mathrm{EI}^N = K_\mathrm{EI} \left(\frac{N_{\mathrm{S},0}}{C_{\mathrm{S},0}} \right)$$

- *Non-Competitive Inhibition*

$$\frac{\mathrm{d}C_\mathrm{S}}{\mathrm{d}t} = -\left(\frac{k_3 C_{\mathrm{E},0} C_\mathrm{S}}{C_\mathrm{S} + K_\mathrm{ES}} \right) \times \left(1 + \frac{C_\mathrm{I}}{K_\mathrm{EI}} \right)^{-1} \tag{2.28}$$

$$a_\mathrm{NCI} = \frac{k_3 N_{\mathrm{E},0} N_\mathrm{S}}{N_\mathrm{S} + K_\mathrm{ES}^N} \left(1 + \frac{N_\mathrm{I}}{K_\mathrm{EI}^N} \right)^{-1} \tag{2.29}$$

- *Un-Competitive Inhibition*

$$\frac{\mathrm{d}C_\mathrm{S}}{\mathrm{d}t} = -\frac{k_3 C_{\mathrm{E},0} C_\mathrm{S}}{K_\mathrm{ES} + \left(1 + \frac{C_\mathrm{I}}{K_\mathrm{EI}} \right) C_\mathrm{S}} \tag{2.30}$$

$$a_\mathrm{UCI} = \frac{k_3 N_{\mathrm{E},0} N_\mathrm{S}}{K_\mathrm{ES}^N + \left(1 + \frac{N_\mathrm{I}}{K_\mathrm{EI}^N} \right) N_\mathrm{S}} \tag{2.31}$$

Bibliography

Agarwal, A., Adams, R., Castellani, G. C., & Shouval, H. Z. (2012). On the precision of quasi steady state assumptions in stochastic dynamics. *The Journal of chemical physics*, *137*(4), 44105.

Barik, D., Paul, M. R., Baumann, W. T., Cao, Y., & Tyson, J. J. (2008). Stochastic simulation of enzyme-catalyzed reactions with disparate timescales. *Biophysical Journal*, *95*(8), 3563-3574. doi: 10.1529/biophysj.108.129155

Basu, M., & Mohanty, P. (2009). Stochastic modeling of single molecule Michaelis Menten kinetics. *arXiv preprint arXiv:0901.2844*.

Briggs, G. E., & Haldane, J. B. S. (1925). A note on the kinetics of enzyme action. *Biochemical journal*, *19*(2), 338–339.

Cao, Y., Gillespie, D. T., & Petzold, L. R. (2005). Accelerated stochastic simulation of the stiff enzyme-substrate reaction. *Journal of Chemical Physics*, *123*(14), 144917. doi: 10.1063/1.2052596

Chapman, J., Ismail, A. E., & Dinu, C. Z. (2018). Industrial applications of enzymes: recent advances, techniques, and outlooks. *Catalysts*, *8*(6), 238.

Dhatt, S., & Banerjee, K. (2019). Efficacy of quasi-steady-state approximation in Michaelis-Menten kinetics: a stochastic signature. *Journal of Mathematical Chemistry*, *57*(7), 1797-1805. doi: 10.1007/s10910-019-01038-9

Dóka, G., Éva Lente. (2012). Stochastic mapping of the Michaelis-Menten mechanism. *The Journal of Chemical Physics*, *136*(5), 054111. doi: 10.1063/1.3681942

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

Gillespie, D. T. (2007, apr). Stochastic Simulation of Chemical Kinetics. *Annual Review of Physical Chemistry*, *58*(1), 35–55. Retrieved from `https://doi.org/10.1146/annurev.physchem.58.032806.104637` doi: 10.1146/annurev.physchem.58.032806.104637

Hahl, S. K., & Kremling, A. (2016). A comparison of deterministic and stochastic modeling approaches for biochemical reaction systems: On fixed points, means, and modes. *Frontiers in genetics*, *7*, 157.

J. Donoso. (2006). *Biopolímeros.* Retrieved from `http://facultatciencies.uib.cat/prof/josefa.donoso/campus/modulos/modulo4/modulo4{_}14.htm`

Kang, H.-W., KhudaBukhsh, W. R., Koeppl, H., & Rempała, G. A. (2019). Quasi-steady-state approximations derived from the stochastic model of enzyme kinetics. *Bulletin of mathematical biology*, *81*(5), 1303–1336.

Kou, S. C., Cherayil, B. J., Min, W., English, B. P., & Xie, X. S. (2005). Single-molecule michaelis-menten equations. *The Journal of Physical Chemistry B*, *109*(41), 19068-19081. doi: 10.1021/jp051490q

Lawson, M. J., Petzold, L., & Hellander, A. (2015). Accuracy of the Michaelis–Menten approximation when analysing effects of molecular noise. *Journal of The Royal Society Interface*, *12*(106), 20150054.

Lecca, P. (2013). Stochastic chemical kinetics. *Biophysical reviews*, *5*(4), 323–345.

Lente, G. (2013). A binomial stochastic kinetic approach to the Michaelis-Menten mechanism. *Chemical Physics Letters*, *568-569*, 167-169. doi: 10.1016/j.cplett.2013.03.011

Marchetti, L., Priami, C., & Thanh, V. H. (2017). Stochastic Simulation of Biochemical Reaction Systems. In *Simulation algorithms for computational systems biology* (pp. 7–28). Springer.

Martínez Urreaga, J., Mira, J., & González Fernández, C. (2003). Introducing the stochastic simulation of chemical reactions: Using the Gillespie algorithm and MATLAB,. *Chem. Engrg. Educ*, *37*, 14–19.

Menz, S. (2013). *Hybrid stochastic-deterministic approaches for simulation and analysis of biochemical reaction networks.*

Milo, R., Jorgensen, P., Moran, U., Weber, G., & Springer, M. (2010). BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic acids research*, *38*(suppl_1), D750–D753.

Qian, H., & Elson, E. L. (2002). Single-molecule enzymology: stochastic Michaelis-Menten kinetics. *Biophysical Chemistry*, *101-102*, 565-576. doi: https://doi.org/10.1016/S0301-4622(02)00145-X

Rao, C. V., & Arkin, A. P. (2003). Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *The Journal of chemical physics*, *118*(11), 4999–5010.

Érdi, P., & Lente, G. (2014). *Stochastic Chemical Kinetics*. Springer New York. doi:

10.1007/978-1-4939-0387-0

Rogers, A., & Gibon, Y. (2009). Enzyme kinetics: theory and practice. In *Plant metabolic networks* (pp. 71–103). Springer.

Romero-Severson, E. O., Ribeiro, R. M., & Castro, M. (2018). Noise is not error: detecting parametric heterogeneity between epidemiologic time series. *Frontiers in microbiology*, *9*, 1529.

Sanft, K. R., Gillespie, D. T., & Petzold, L. R. (2011). Legitimacy of the stochastic michaelis–menten approximation. *IET systems biology*, *5*(1), 58–69.

Schnoerr, D., Sanguinetti, G., & Grima, R. (2017). Approximation and inference methods for stochastic biochemical kinetics—a tutorial review. *Journal of Physics A: Mathematical and Theoretical*, *50*(9), 093001. doi: 10.1088/1751-8121/aa54d9

Segel, L. A. (1988). On the validity of the steady state assumption of enzyme kinetics. *Bulletin of mathematical biology*, *50*(6), 579–593.

Segel, L. A., & Slemrod, M. (1989, September). The quasi-steady-state assumption: A case study in perturbation. *SIAM Review*, *31*(3), 446-477. doi: 10.1137/1031091

Székely Jr, T., & Burrage, K. (2014). Stochastic simulation in systems biology. *Computational and structural biotechnology journal*, *12*(20-21), 14–25.

Tsimrng, L. S. (2014). Noise in biology. *Reports on Progress in Physics*, *77*(2), 26601.

Villadsen, J., Nielsen, J., & Lidén, G. (2011). *Bioreaction engineering principles*. Springer Science & Business Media.

Warne, D. J., Baker, R. E., & Simpson, M. J. (2019). Simulation and inference algorithms for stochastic biochemical reaction networks: from basic concepts to state-of-the-art. *Journal of the Royal Society Interface*, *16*(151), 20180943.

Wilkinson, D. J. (2018). *Stochastic modelling for systems biology* (Third Edit ed.). CRC press.

Wu, J., Vidakovic, B., & Voit, E. O. (2011). Constructing stochastic models from deterministic process equations by propensity adjustment. *BMC Systems Biology*, *5*(1), 187.

Yang, Y., & Jiang, D. (2016). Long-time behavior of a perturbed enzymatic reaction model under negative feedback process by white noise. *Journal of Mathematical Chemistry*, *54*(4), 854-865. doi: 10.1007/s10910-016-0590-2

Yi, M., & Liu, Q. (2010). Michaelis-Menten mechanism for single-enzyme and multi-enzyme system under stochastic noise and spatial diffusion. *Physica A: Statistical Mechanics and its Applications*, *389*(18), 3791-3803. doi: 10.1016/j.physa.2010.05.041

Zhou, K., Jiang, X., & Chan, T. L. (2020). Error analysis in stochastic solutions of population balance equations. *Applied Mathematical Modelling*, *80*, 531–552.

Chapter 3

Stochastic modeling approach and tuning for fermentation models

Abstract

The study and application of bioprocesses at an industrial level, such as the production of xylitol, makes it necessary to develop mathematical models that fit, increasingly predict the behavior of systems, as well as the possibility of evaluating their inherent uncertainty. It is precisely in the study of uncertainty and its application to biological systems where stochastic modeling has its greatest contribution and relevance. In this work, a significant contribution is made, since it is the first case in the literature in which a stochastic model is proposed in which enzymatic kinetic phenomena with inhibition, microbial growth, and mass transport are included in the same case. Herein, it is presented the modeling, simulation, and re-estimation of both deterministic and stochastic parameters of the xylitol production model from two substrate sources. Model calibration for the deterministic model was performed with the interior-point algorithm and, in turn, for the stochastic model was performed through the Approximate Bayesian Computation rejection sampler method. The re-estimation of parameters of the deterministic model showed a substantial improvement in experimental data reproduction, especially for xylitol. The stochastic simulation was carried out with Gillespie's Stochastic Simulation Algorithm. The stochastic model results were able to

reproduce the behavior of the different chemical species in the system, evaluate the uncertainty associated with the model through the standard deviation and apply a method of parameter estimation to a stochastic model of this type.

3.1 Introduction

Xylitol is a five-carbon sugar alcohol, currently used as a sweetener in the pharmaceutical, nutraceutical, food, beverage industries among others (Edelstein et al., 2007; Maguire & Rugg-Gunn, 2003). Xylitol has become a healthier and more ecological alternative to sucrose, as it has an equivalent relative sweetness, with lower "energy content". Besides, it can be obtained from natural sources such as agro-industrial waste (Dasgupta, Bandhu, Adhikari, & Ghosh, 2017). In recent decades, it has become a feasible replacement of production by chemical synthesis, given the need for more environmentally friendly processes (Dasgupta et al., 2017; Park, Sang, Park, & Park, 2005). Despite interesting, this bioproduction faces technical and economic limitations which have delayed its implementation at a large scale. Therefore, there is a need for process optimization which translates into the need for mathematical models that facilitate process design, control, optimization, and scaling up bioproduction processes.

Tochampa et al. (2005) were the first to introduce a mathematical model of fermentation with *Candida mogii*, which took into account the inhibition generated by the presence of glucose in the system (a very common sugar in agro-industrial waste). Since then, other authors have adapted this model for fed-batch processes, immobilized cells, and different microorganisms, among many improvements (Dorantes-Landa, Cocotle-Ronzón, Morales-Cabrera, & Hernández-Martínez, 2020; Nguyen, Le, & Boontawan, 2020; Prado-Rubio, Hernández-Escoto, Rodriguez-Gomez, Sirisansaneeyakul, & Morales-Rodriguez, 2015; Sirisansaneeyakul, Wannawilai, & Chisti, 2013). Also, the xylitol model proposed by Tochampa et al. (2015) for a fed-batch reactor has been applied in process stochastic optimization, improving the productivity of xylitol from the ratio of the feed flow to the bioreactor (Koop, Corazza, Voll, & Bonilla-Petriciolet, 2017).

On the other hand, (Rao & Arkin, 2003) were the first to apply the stochastic approach of QSSA for the case of enzymatic reaction without inhibition. Barik, Paul, Baumann, Cao, and Tyson (2008) proposed how to improve the time and efficiency of the stochastic simulation when there are different time scales between the reactions. (Dóka, 2012) determine the regions in the parameter space for which stochastic kinetic

approach is unavoidable and (Kang, KhudaBukhsh, Koeppl, & Rempała, 2019) makes a complete stochastic derivation for the different types of QSSA, indicating the applicability criteria of the same in comparison with the deterministic version. Based on the state of the art, there are still no reported examples on modeling microbial kinetics with enzymatic reactions from the stochastic approach.Therefore, this work aims to present the complete translation of a complex fermentation model (including microbial growth, competitive inhibition, and mass transport) to propensities and to propose a simulation methodology of the SSA from the ODEs, when the elemental reactions are not known, nor the stoichiometry associated with them. On the other hand, a parameter re-estimation was made from the deterministic approach with the interior-point algorithm and with the Approximate Bayesian Computation rejection Sampler (ABC rejection sampler) method to have a better fit of the model, as well as having more limited and realistic uncertainty of the system (Liu & Gunawan, 2017; Soize, 2013).

Table 3.1. Nomenclature symbols.

Greek letters		
γ_{cell}	Reductance degree of biomass	-
γ_{xit}	Reductance degree of xylitol	-
η	Energy yield coefficient for biomass production	-
μ_{xit}	Specific growth rate on xylitol	h^{-1}
μ_{glc}^{max}	Maximum specific growth rate on glucose	h^{-1}
μ_{xit}^{max}	Maximum specific growth rate on xylose	h^{-1}
ρ_{x}	yeast cell density	g DCW/L
σ_{cell}	Weight fraction of carbon in biomass	g atom-C g DCW^{-1}
σ_{xit}	Weight fraction of carbon in xylitol	g atom-C g DCW^{-1}
Variables		
a_{cell}	Specific surface area of the cell,	m^2/gDCW
C_{x}	Biomass concentration	g Biomass/L
C_{glc}	Glucose concentration	g Glucose/L
C_{xyl}	Xylose concentration	g Xylose/L
$C_{\text{xit}}^{\text{in}}$	Intracellular xylitol concentration	g Xyilitol/L
$C_{\text{xit}}^{\text{ex}}$	Extracellular xylitol concentration	g Xyilitol/L
$k_{i,glc}$	Glucose inhibition constant	gglucoseL^{-1}
$k_{i,xyl}$	Xylose inhibition constant	gxyloseL^{-1}
K_r	Repression constant by glucose	gglucoseL^{-1}
$K_{s,glc}$	Saturation constant based on glucose	gglucoseL^{-1}
$K_{s,xyl}$	Saturation constant based on xylose	gxyloseL^{-1}
$K_{s,xit}$	Saturation constant based on xylitol	gxylitolL^{-1}
N_{Av}	Avogadro's number	mol^{-1}
P_{xit}	Permeability coefficient of xylitol	ms^{-1}
q_{glc}^{max}	Specific rate of maximum glucose consumption	$\text{g glucose} \cdot \text{gDCW}^{-1}\text{h}^{-1}$
q_{xyl}^{max}	Specific rate of maximum xylose consumption	$\text{g xylose} \cdot \text{gDCW}^{-1}\text{h}^{-1}$
$r_{\text{f,xit}}$	Specific formation rate of xylitol,	$\text{g xylitol g DCW}^{-1}\,\text{h}^{-1}$
$r_{\text{t,xit}}$	Mass flow of xylitol on a cell dry basis	$\text{g xylitol g DCW}^{-1}\,\text{h}^{-1}$
$r_{\text{u,xit}}$	Intracellular xylitol consumption rate	$\text{g xylitol g DCW}^{-1}\,\text{h}^{-1}$
V	Volume	L
$Y_{\text{x/xit}}$	Biomass-xylitol yield	g DCW/g xylitol

3.2 Methods

Figure 3.1 shows the two methodologies followed to simulate and re-estimate the xylitol production model proposed by Tochampa et al. (2005):

- Using a deterministic approach of the model by Tochampa et al. (2005) through the interior-point algorithm coupled with the function *fmincon* of Matlab 2020a. The right branch of Figure 3.1 shows the methodology (Section 3.2.3) traveled from the deterministic approach of the model formulated by Tochampa et al. (2005) (the system of differential equations that describe the concentration profiles over time for biomass, glucose, xylose, and xylitol) until the new parameter values are obtained.

- Applying the Stochastic Simulation Algorithm in the ABC rejection sampler. This approach was simulated, taking the ODEs to propensity functions (Sections 2.2.3 and 2.5.2), with which the SSA is applied (Section 2.2.2) (Gillespie, 2007; Rao & Arkin, 2003). This second option of parameter estimation was carried out to find finding a substantial improvement in the fit of the model to the experimental data, especially for xylitol, which is the compound of interest.

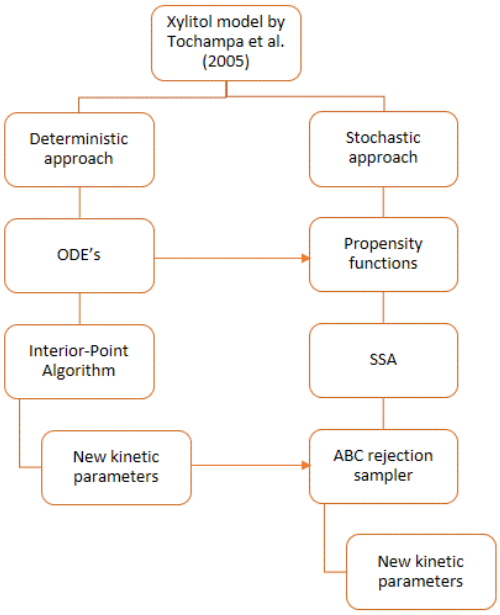

Figure 3.1. Metodological flowsheet.

3.2.1 Deterministic approach

The mathematical model proposed by Tochampa et al. (2005) yields the mass concentration of biomass (x, *Candida moggi*), glucose (glc), xylose (xyl), intra and extracellular xylitol (xit); in a batch reactor as a function of time. It takes into account the transport of xylitol outside the cell and competitive inhibition between glucose and xylose. The nomenclature and units for each parameter are presented in Table 3.1. The deterministic reactor mass balances and kinetic models are described below:

- The microbial growth rate was modeled with an auto-catalytic growth equation, which takes into account the growth of biomass in two substrates (xylose and glucose) with a term of inhibition of growth in xylose by glucose.

$$\frac{\mathrm{d}C_{\mathrm{x}}}{\mathrm{d}t} = \mu \cdot C_{\mathrm{x}} \tag{3.1}$$

$$\mu = \mu_{\mathrm{glc}}^{\max} \frac{C_{\mathrm{glc}}}{K_{\mathrm{s,glc}} + C_{\mathrm{glc}}} + \mu_{\mathrm{xit}}^{\max} \frac{C_{\mathrm{xit}}^{\mathrm{in}}}{K_{\mathrm{s,xi}t} + C_{\mathrm{xit}}^{in}} \cdot \frac{K\mathrm{r}}{K_{\mathrm{r}} + C_{\mathrm{glc}}} \tag{3.2}$$

- The glucose consumption rate was modeled as a function of the specific glucose consumption rate (q_{glc}), where the q_{glc} corresponds an enzymatic kinetic with competitive inhibition, that is, that the same active site of the enzyme can be attacked by both the glucose and xylose. In this case, the xylose is taken as the inhibitor.

$$\frac{\mathrm{d}C_{\mathrm{glc}}}{\mathrm{d}t} = -q_{\mathrm{glc}} \cdot C_{\mathrm{x}} \tag{3.3}$$

$$q_{\mathrm{glc}} = q_{\mathrm{glc}}^{\mathrm{max}} \frac{C_{\mathrm{glc}}}{C_{\mathrm{glc}} + K_{\mathrm{s,glc}} \left(1 + \frac{C_{\mathrm{xyl}}}{k_{\mathrm{i,xyl}}}\right)} \tag{3.4}$$

- The xylose consumption rate was modeled as a function of the specific xylose consumption rate (q_{xyl}). Also, the q_{xyl} represents the competitive inhibition generated by the glucose.

$$\frac{\mathrm{d}C_{\mathrm{xyl}}}{\mathrm{d}t} = -q_{\mathrm{xyl}} \cdot C_{\mathrm{x}} \tag{3.5}$$

$$q_{\mathrm{xyl}} = q_{\mathrm{xyl}}^{\mathrm{max}} \frac{C_{\mathrm{xyl}}}{C_{\mathrm{xyl}} + K_{\mathrm{s,xyl}} \left(1 + \frac{C_{\mathrm{glc}}}{k_{\mathrm{i,glc}}}\right)} \tag{3.6}$$

- The rate of change in intracellular xylitol concentration. This involves the generation of xylitol within the cell and its transport across the cell membrane.

$$\frac{\mathrm{d}C_{\mathrm{xit}}^{\mathrm{in}}}{\mathrm{d}t} = \rho_{\mathrm{x}} \left(r_{\mathrm{f,xit}} - r_{\mathrm{u,xit}} - r_{\mathrm{t,xit}}\right) - \mu \cdot C_{\mathrm{xit}}^{in} \tag{3.7}$$

$$r_{\mathrm{f,xi}t} = \frac{(\mathrm{mw})_{\mathrm{xit}}}{(\mathrm{mw})_{\mathrm{xyl}}} \cdot q_{\mathrm{xyl}} \tag{3.8}$$

Where $(\mathrm{mw})_{\mathrm{xit}}$and $(\mathrm{mw})_{\mathrm{xyl}}$ are the molecular masses of xylitol and xylose, respectively.

$$r_{\mathrm{u,xit}} = \frac{\mu_{\mathrm{xit}}}{Y_{\mathrm{x/xit}}} \tag{3.9}$$

$$Y_{\mathrm{x/xit}} = \eta \frac{\sigma_{\mathrm{xit}} \gamma_{\mathrm{xit}}}{\sigma_{\mathrm{cell}} \gamma_{\mathrm{cell}}} \tag{3.10}$$

$$r_{\mathrm{t,xit}} = 3.6 \times 10^{6} P \cdot a_{\mathrm{cell}} \left(C_{\mathrm{xit}}^{\mathrm{in}} - C_{\mathrm{xit}}^{\mathrm{ex}}\right) \tag{3.11}$$

- The rate of change in the concentration of extracellular xylitol, which is a function of $r_{\mathrm{t,xit}}$and C_{x}.

$$\frac{\mathrm{d}C_{\mathrm{xit}}^{\mathrm{ex}}}{\mathrm{d}t} = r_{\mathrm{t,xit}} \cdot C_{\mathrm{x}} \tag{3.12}$$

Table 3.2 shows the constant parameters and initial conditions for the deterministic and

stochastic model.

Table 3.2. Constant parameters for xylitol model.

Parameter	Value	Parameter	Value
η	0.58	a_{cell}	7.6
σ_{xit}	0.39	ρ_x	120
σ_{cell}	0.49	$C_{\text{x},0}$	5.0655
γ_{cell}	4.19	$C_{\text{glc},0}$	4.2795
γ_{xit}	4.4	$C_{\text{xyl},0}$	31.0044
$(\text{mw})_{\text{xyl}}$	150.13	$C^{\text{in}}_{\text{xit},0}$	0
$(\text{mw})_{\text{xit}}$	152.15	$C^{\text{ex}}_{\text{xit},0}$	0

DCW = Dry Cell Weight, the units of each parameter are presented in Table 3.1. (Tochampa et al., 2005).

3.2.2 Stochastic approach

The propensity functions for the stochastic model were derived from the deterministic model through the application of the CME to translate the ODEs (see Section 2.2.3 and Eqs. 3.1, 3.3, 3.5, 3.7 and 3.12) into their *equivalent* propensities (Rao & Arkin, 2003).

The following is the procedure performed to convert each differential equation of the xylitol model to its propensity function. The easiest case is when the same concentrations appears in both sides, e.g.

$$\frac{\mathrm{d}C_{\text{x}}}{\mathrm{d}t} = \mu C_{\text{x}} \tag{3.13}$$

$$C_i = \frac{N_i(\text{mw})_i}{N_{Av}V} \tag{3.14}$$

Replacing Eq. 3.14 into Eq. 3.13, it is obtained

$$\frac{\mathrm{d}N_{\text{x}}}{\mathrm{d}t} = \mu N_{\text{x}} \tag{3.15}$$

The propensity function is equal of the right side of the Eq. 3.15.

$$a_{\text{x}} = \mu N_{\text{x}} \tag{3.16}$$

In the same way, the propensities for Eqs. 3.3 and 3.5 are

$$a_{\text{glc}} = \left[\frac{(\text{mw})_x}{(\text{mw})_{\text{glc}}}\right] q_{\text{glc}} N_{\text{x}} \tag{3.17}$$

$$a_{\mathrm{xyl}} = \left[\frac{(\mathrm{mw})_x}{(\mathrm{mw})_{\mathrm{xyl}}}\right] q_{\mathrm{xyl}} N_{\mathrm{x}} \tag{3.18}$$

Then, the conversion from concentration to number of molecules of μ, q_{glc} and q_{xyl} is made from Eqs. (3.2), (3.4) and (3.6) respectively, as presented below.

$$\mu = \mu_{\mathrm{glc}} + \mu_{\mathrm{xit}} \tag{3.19}$$

where

$$\mu_{\mathrm{glc}} = \mu_{\mathrm{glc}}^{\max} \frac{C_{\mathrm{glc}}}{K_{\mathrm{s,glc}} + C_{\mathrm{glc}}} \tag{3.20}$$

When replacing Eq. 3.48 into Eq. 3.20, it is obtained:

$$\mu_{\mathrm{glc}} = \mu_{\mathrm{glc}}^{\max} \frac{\frac{N_{\mathrm{glc}}(\mathrm{mw})_{\mathrm{glc}}}{(N_{\mathrm{Av}}V)}}{K_{\mathrm{s,glc}} + \frac{N_{\mathrm{glc}}(\mathrm{mw})_{\mathrm{glc}}}{(N_{\mathrm{Av}}V)}} \tag{3.21}$$

Then, it is taken a common factor $(\mathrm{mw})_{\mathrm{glc}}/(N_{Av}V)$ in the denominator of the fraction and it is canceled in the numerator and denominator to obtain:

$$\mu_{\mathrm{glc}} = \mu_{\mathrm{glc}}^{\max} \frac{N_{\mathrm{glc}}}{K_{\mathrm{s,glc}} \frac{(N_{\mathrm{Av}}V)}{(\mathrm{mw})_{\mathrm{glc}}} + N_{\mathrm{glc}}} \tag{3.22}$$

$$K_{\mathrm{s,glc}}^{N} = K_{\mathrm{s,glc}} \frac{(N_{\mathrm{Av}}V)}{(\mathrm{mw})_{\mathrm{glc}}} \tag{3.23}$$

And finally, Eq. 3.20 is rewritten as

$$\mu_{\mathrm{glc}} = \mu_{\mathrm{glc}}^{\max} \frac{N_{\mathrm{glc}}}{K_{\mathrm{s,glc}}^{N} + N_{\mathrm{glc}}} \tag{3.24}$$

An analogous procedure is now followed for μ_{xit}

$$\mu_{\mathrm{xit}} = \mu_{\mathrm{xit}}^{\max} \frac{C_{\mathrm{xit}}}{K_{\mathrm{s,xit}} + C_{\mathrm{xit}}} \cdot \frac{K_{\mathrm{r}}}{K_{\mathrm{r}} + C_{\mathrm{glc}}} \tag{3.25}$$

After the translation process:

$$\mu_{\mathrm{xit}} = \mu_{\mathrm{xit}}^{\max} \frac{N_{\mathrm{xit}}}{K_{\mathrm{s,xit}}^{N} + N_{\mathrm{xit}}} \cdot \frac{K_{\mathrm{r}}^{N}}{K_{\mathrm{r}}^{N} + N_{\mathrm{glc}}} \tag{3.26}$$

$$K_{\mathrm{s,xit}}^{N} = K_{\mathrm{s,xit}} \frac{(N_{\mathrm{Av}}V)}{(\mathrm{mw})_{\mathrm{xit}}} \tag{3.27}$$

$$K_{\mathrm{r}}^{N} = K_{\mathrm{r}} \frac{(N_{\mathrm{Av}} V)}{(\mathrm{mw})_{\mathrm{glc}}} \tag{3.28}$$

For the specific glucose and xylose consumption rates (see Eqs. 3.43.6), an analogous procedure is applied.

$$q_{glc} = q_{\mathrm{glc}}^{\max} \frac{N_{\mathrm{glc}}}{N_{\mathrm{glc}} + K_{\mathrm{S,glc}}^{N} \left(1 + \frac{N_{\mathrm{xyl}}}{K_{\mathrm{i,xyl}}^{N}}\right)} \tag{3.29}$$

$$q_{\mathrm{xyl}} = q_{\mathrm{xyl}}^{\max} \frac{N_{\mathrm{xyl}}}{N_{\mathrm{xyl}} + K_{\mathrm{S,xyl}}^{N} \left(1 + \frac{N_{\mathrm{glc}}}{K_{\mathrm{i,glc}}^{N}}\right)} \tag{3.30}$$

$$K_{\mathrm{i,xyl}}^{N} = K_{\mathrm{i,xyl}} \frac{(N_{\mathrm{Av}} V)}{(\mathrm{mw})_{\mathrm{xyl}}} \tag{3.31}$$

$$K_{\mathrm{i,glc}}^{N} = K_{\mathrm{i,glc}} \frac{(N_{\mathrm{Av}} V)}{(\mathrm{mw})_{\mathrm{glc}}} \tag{3.32}$$

$$K_{\mathrm{s,xyl}}^{N} = K_{\mathrm{s,xyl}} \frac{(N_{\mathrm{Av}} V)}{(\mathrm{mw})_{\mathrm{xyl}}} \tag{3.33}$$

Finally, for Eqs. 3.7 and 3.12 that correspond to the intra and extracellular xylitol, the following expressions are obtained:

$$a_{\mathrm{xit}}^{\mathrm{in}} = \frac{(N_{\mathrm{Av}} V)}{(\mathrm{mw})_{\mathrm{xit}}} \cdot \rho_{\mathrm{x}} \left(r_{\mathrm{f,xit}} - r_{\mathrm{u,xit}} - r_{\mathrm{t,xit}}\right) - \mu \cdot N_{\mathrm{xit}}^{\mathrm{in}} \tag{3.34}$$

$$a_{\mathrm{xit}}^{\mathrm{ex}} = r_{\mathrm{t,xit}} \frac{(\mathrm{mw})_{\mathrm{x}}}{(\mathrm{mw})_{\mathrm{xit}}} \cdot N_{\mathrm{x}} \tag{3.35}$$

where $r_{\mathrm{f,xit}}$, $r_{\mathrm{u,xi}t}$ and $r_{\mathrm{t,xi}t}$ are defined as

$$r_{\mathrm{f,xit}} = \frac{(\mathrm{mw})_{\mathrm{xit}}}{(\mathrm{mw})_{\mathrm{xyl}}} q_{\mathrm{xyl}} \tag{3.36}$$

$$r_{\mathrm{u,xit}} = \frac{\mu_{\mathrm{xit}}}{Y_{x/\mathrm{xit}}} \tag{3.37}$$

$$r_{\mathrm{t,xit}} = 3.6 \times 10^{6} \frac{P_{\mathrm{xit}} a_{\mathrm{cell}} (\mathrm{mw})_{\mathrm{xit}}}{(N_{\mathrm{Av}} V)} \left(N_{\mathrm{xit}}^{\mathrm{in}} - N_{\mathrm{xit}}^{\mathrm{ex}}\right) \tag{3.38}$$

A stochastic simulation of this model was complex since Gillespie's algorithm requires

- The set of reactions that takes place.

- The stoichiometry is associated with each reaction to define the consumption or production of each species.

For example, in the SSA for two reactions.

$$\mathrm{A} \rightarrow \mathrm{B} \tag{3.39}$$

$$2\mathrm{B} \rightarrow \mathrm{C} \tag{3.40}$$

a reaction channel must be selected at each time step, and after that, the numbers of molecules are updated. If the selected channel was $2\mathrm{B} \rightarrow \mathrm{C}$, then one molecule of C is produced and two of B are consumed. However, a stochastic simulation of the xylitol model is difficult because only the mass balances (ODE's) are known, but he reactions that give rise to microbial growth are unknown, which leads to a lack of knowledge of their stoichiometry. To overcome this obstacle, the SSA formulation of the $\mathrm{S} \rightarrow \mathrm{P}$ Michelis-Menten enzymatic kinetics in Section 2.2.3 was followed, assuming that each propensity function obtained from each ODE (Eqs. 3.16 - 3.35), corresponds to an independent reaction channel with a 1 to 1 stoichiometry, even for biomass (J. Donoso, 2006; Rogers & Gibon, 2009; Villadsen, Nielsen, & Lidén, 2011). Hence, 5 reaction channels are obtained, 1 channel for each species (i.e. biomass, glucose, xylose, intra, and extracellular xylitol). Therefore, the right side of a_{x} is interpreted as increase of biomass molecules

$$N_{\mathrm{x}} \rightarrow N_{\mathrm{x}} + 1, \tag{3.41}$$

given the positive sing of the Eq. 3.1, which represents species production. On the other hand, the a_{glc} and a_{xyl} are interpreted as a decrease of glucose and xylose molecules, respectively. Notice in Eqs. 3.3 and 3.5, the negative sign which represent species consumption.

$$N_{\mathrm{glc}} \rightarrow N_{\mathrm{glc}} - 1 \tag{3.42}$$

$$N_{\mathrm{xyl}} \rightarrow N_{\mathrm{xyl}} - 1 \tag{3.43}$$

The situation for the intra and extra xylitol propensities ($a^{\mathrm{in}}_{\mathrm{xit}}$ and a^{ex}_{xit}) is particular, since they should always be positive. However in this case, the propensities can take negative and positive values given the interaction between the terms of the ODEs defining xylitol consumption, production and transport. Therefore, the questions to be solved are:

1. How to decide if xylitol is produced or consumed?

2. How to calculate the probability intervals associated with each reaction channel and the step time? Since neither times nor negative probabilities can be obtained (see Section 2.2.2)

The first question was solved to assign the channel based on the sign of the propensity value, following the convention for chemical reactions that the negative sign indicates consumption and the positive sign production. Thus, if $a^{\text{in}}_{\text{xit}}$ is positive

$$N^{\text{in}}_{\text{xit}} \rightarrow N^{\text{in}}_{\text{xit}} + 1 \tag{3.44}$$

and if it is negative

$$N^{\text{in}}_{\text{xit}} \rightarrow N^{\text{in}}_{\text{xit}} - 1 \tag{3.45}$$

On the other hand, if $a^{\text{ex}}_{\text{xit}}$ is positive

$$N^{\text{ex}}_{\text{xit}} \rightarrow N^{\text{ex}}_{\text{xit}} + 1 \tag{3.46}$$

and if it is negative

$$N^{\text{ex}}_{\text{xit}} \rightarrow N^{\text{ex}}_{\text{xit}} - 1 \tag{3.47}$$

And the second question was solved taking the absolute value of $|a^{\text{in}}_{\text{xit}}|$ and $|a^{\text{ex}}_{\text{xit}}|$ since to define the probability intervals and the time step, only the value of each propensity is required but not the sign. The mean and standard deviations were calculated from 1000 simulation runs based on an analogous sensitivity analysis presented in Section 2.2.4. The simulation results were interpolated at the same times read from the experimental results presented by Tochampa et al. (2005).

The simulation volume (see Table 3.3) was chosen with a preliminary sensitivity analysis, for each tested V the initial numbers of molecules were determined as

$$N_i = \frac{C_i N_{Av} V}{(\text{mw})_i}, \tag{3.48}$$

and the $N_i(t)$ profiles were generated with the SSA. This process was repeated until the uncertainty generated by the stochastic method concerning the mean value covered all the experimental data for xylitol since it is the substance of interest (see Figure 3.4).

Note: the molecular weight of the biomass is estimated as

$$(\mathrm{mw})_x = (\mathrm{mw})_\mathrm{C} + 1.79(\mathrm{mw})_\mathrm{H} + 0.5(\mathrm{mw})_\mathrm{O} + 0.2(\mathrm{mw})_\mathrm{N} \tag{3.49}$$

where C is carbon, H is hydrogen, O is oxygen and N is Nitrogen (Tochampa et al., 2005).

Table 3.3. Initial conditions for stochastic model.

Parameter	Value	Parameter	Value
$N_{\mathrm{x},0}$	62	$N^{in}_{xit,0}$	0
$N_{\mathrm{glc},0}$	7	$N^{ex}_{xit,0}$	0
$N_{\mathrm{xyl},0}$	62	V	5×10^{-22} L

3.2.3 Parameter estimation

The estimated parameters were: $\mu^{\max}_{\mathrm{glc}}$, $\mu^{\max}_{\mathrm{xit}}$, $q^{\max}_{\mathrm{xyl}}$, $q^{\max}_{\mathrm{glc}}$, $K_{\mathrm{s,xyl}}$, $K_{\mathrm{s,glc}}$, $K_{\mathrm{s,xit}}$, $K_{\mathrm{i,xyl}}$, $K_{\mathrm{i,glc}}$, K_r, P_{xit}. These were estimated both with the deterministic model and the interior point method, as well as with the stochastic model and the ABC rejection Sampler method. The experimental data were obtained from Tochampa et al. (2005), in which to report the concentrations of biomass, glucose, xylose and extracellular xylitol for a batch reactor. The parameter estimation was made for a set of 28 experimental data for each chemical species. The objective function was

$$F_{\mathrm{obj}} = \sum_{i=1}^{M} W_i \sum_{j=1}^{O} \left(\hat{y}_{i,j} - y_{i,j}\right)^2 \tag{3.50}$$

$$W_i = \frac{1}{\max(y_{i,j})} \tag{3.51}$$

where M is the number of experimentally observable species variables (i.e. concentrations of biomass, glucose, xylose, and xylitol concentration; subscript i), O is the total number of sampling points of concentration for each the species (subscript j), W_i is the weighting coefficient (a constant value for each species i), $y_{i,j}$ and $\hat{y}_{i,j}$ are experimental and model-predicted concentrations, respectively. The minimization of the objective function was performed with the interior-point algorithm as implemented in the function *fmincon* of Matlab2020a. This gradient-based numerical method calculates the minimum of a nonlinear model where the function and the constraints are continuous, creating a succession of points on a curve within the feasible region (Ruiz

Paredes, 2019).

Also, the confidence intervals for the parameters were estimated with a 95% confidence level using an asymptotic normal distribution for the parameter estimate through the *nlparci* function of Matlab 2020a. The confidence intervals of the predictor (concentration) were analyzed using the Jacobian matrix ($\boldsymbol{J}$) for the deterministic model (Englezos & Kalogerakis, 2000; Myers, Montgomery, Vining, & Robinson, 2012),

$$\hat{y}_m \pm t_{st} \sqrt{\sigma^2 \left(1 + \mathbf{J_m} (\mathbf{J'J})^{-1} \mathbf{J'_m}\right)} \tag{3.52}$$

where t_{st} corresponds to *Student's t* for a 95% confidence interval, the subscript m refers to the m-th row (i.e. m-th experimental concentration value) of the Jacobian matrix, The variance (σ^2) is calculated as

$$\sigma^2 = \frac{F_{val}}{(O \cdot M) - n_p} \tag{3.53}$$

where F_{val} is the value of the objective function evaluated with the definitive set of parameters reached in the estimation and n_p is the number of estimated parameters.

The parameter estimation of the stochastic model was done to compare with the deterministic approach. It was carried out through the ABC method since it is a parameter estimation method applicable to stochastic models (Warne, Baker, & Simpson, 2019). This method estimates the parameters through the generation of random values for each parameter within a previously defined interval, then it is measured the "distance" between the experimental data and the stochastic model results (a metric discrepancy, $\rho(y, \hat{y})$)

$$\rho(y, \hat{y}) = \left[\sum_{j=1}^{O} \left(y_j - \hat{y_j}\right)^2 \right]^{1/2}, \tag{3.54}$$

for a candidate parameter set and then it is accepted or rejected based on a tolerance (discrepancy threshold, ϵ) (Gupta, 2013).

As a feasible region for the search for parameters, ± 50% of the values found with the new parameters settings (Table 3.5)**.** Therefore, it was decided that the total number of accepted parameter sets (m) should be 100 since which generate a standard deviation lower than one. This indicates a homogeneous distribution between the different sets of parameters found. In addition, it was taken a value of ϵ equal to 2.5, based on Warne et al. (2019)'s work and that when the value ϵ decreases, the estimate becomes longer and longer. Also, the volume was set for a value of 1×10^{-20} L, which yields initial $N_i s$ values in the range of thousands, which narrows the interval of the SSA

results improving their precision. It comes from the fact that the stochastic simulation becomes a better approximation of the real system with higher values of N ($N \to \infty$), at which point it reaches the thermodynamic limit (Hahl & Kremling, 2016; Lawson, Petzold, & Hellander, 2015; Marchetti, Priami, & Thanh, 2017; Menz, 2013).

On the other hand, the confidence intervals for the parameters were calculated from

$$IP = P \pm t_s \frac{\sigma}{\sqrt{Z}} \tag{3.55}$$

where σ is the standard deviation of 100 accepted parameter set, IP is the confidence interval value by each parameter and P symbolizes the value of each parameter. The *Student's t* (t_s) was calculated with the Matlab 2020a *tinv* function for a probability of 95% and degrees of freedom equal to $(O \cdot M) - n_p$.

The methods for estimating parameters were compared using three statistical tools Montgomery, Peck, and Vining (2021):

- Mean square error (MSE): measures the average of the squared errors between the model results and the experimental data.

$$MSE = \frac{1}{O \cdot M} \sum_{i=1}^{M} \sum_{j=1}^{O} \left(\hat{y}_{i,j} - y_{i,j}\right)^2 \tag{3.56}$$

- Determination coefficient (r^2): quantify the percentage of the experimental results that are explained by the model output.

- Parity diagram: plot model output versus experimental data. Then, the model quality is evidenced in how close the point is closer to the 45° line.

3.3 Results

Table 3.4 presents the comparison between the original model proposed by Tochampa et al. (2005) (ORI), the re-calibrated deterministic model (IP) and an the stochastic model calibration. It is also worth noting that the value of total r^2 (i.e. all models) is very close to one, which indicates that the model can be linearized. The results of both the MSE and r^2 show a significant improvement in the model predictions for biomass and xylose. However the most relevant improvement is for xylitol increasing from a r^2 of 0.78 for ORI to 0.98 with IP. As xylitol is the product of interest, this new set of parameters allows optimization of the production process. On the other hand, when comparing the results between the IP and ABC, basically approaches have the same data reproduction power. Given the high computational demand of the stochastic method (22 h) compared with the 4 min of the deterministic one, the latter is the best estimation method for this particular case. Therefore, the results of both the stochastic and deterministic models will be performed with the data set obtained from the interior-point algorithm.

Table 3.4. Fit of experimental data with three models.

	ORI		IP		ABC	
	MSE	r^2	MSE	r^2	MSE	r^2
Biomass	2.5257	0.8819	1.4293	0.9252	1.3532	0.9274
Glucose	0.0627	0.9473	0.0671	0.9381	0.0677	0.9375
Xylose	2.7740	0.9798	1.9048	0.9860	1.9306	0.9857
Xylitol	2.2640	0.7898	0.1601	0.9814	0.1893	0.9790
Total	1.9066	0.9794	0.8903	0.9900	0.8852	0.9900

ORI: original parameters from Tochampa et al. (2005). IP: deterministic model re-fiting with the interior point algorithm. ABC: parameters from this work, fitted using the stochastic results with the ABC method. The variable in each case is the concentration ($V = 1 \times 10^{-20}$ L). The units of each parameter are presented in Table 3.1.

Figure 3.2 shows the parity results of the model calibration process. The results indicate that the new values estimated by both the interior point method and the ABC, improved the model predictive power, having most of the predictions an error below 15%. Also, it should be noted that the xylitol concentration adjustment was significantly improved compared to the original (ORI).

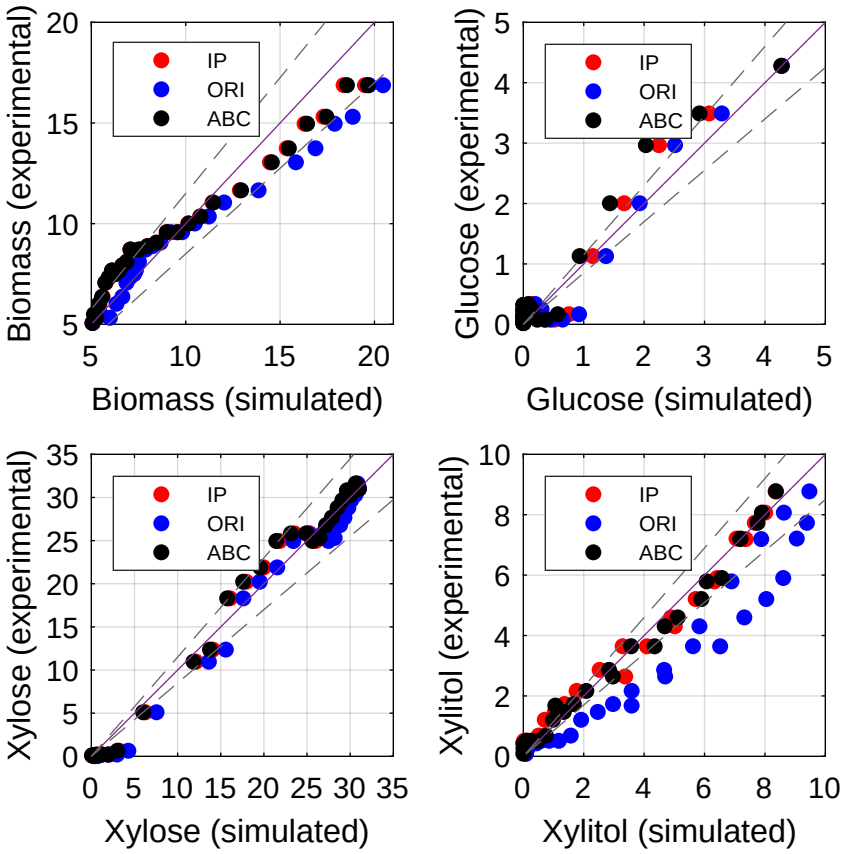

Figure 3.2. Parity diagrams for the concentrations of Biomass, Glucose, Xylose, and Xylitol. Gray dashed lines represent an error of 15% concerning the 45 ° line.

Table 3.5 shows a comparison between the values of the parameters estimated by (Tochampa et al., 2005) and the estimated in this using the deterministic method (interior-point) with 95% confidence interval, where the deviation percentage corresponds to the relative difference of the re-calibrated parameter and its confidence interval. Among the most notable differences is that the re-estimated $K_{s,xit}$ is practically zero, which indicates that cell growth is governed by glucose concentration, maximum specific growth rate (μ_{glc}^{max}) and the repression constant (K_r). Unlike that reported by authors such as Tochampa et al. (2005), Tochampa et al. (2015), and Hernández-Escoto, Prado-Rubio, and Morales-Rodriguez (2016), where their results show that $K_{s,xit}$ has an important incidence within microbial growth. However, this situation can be generated by the multiple combinations of parameters that could be obtained due to the

non-convex nature of the problem.

On the other hand, the confidence intervals for the parameters show that for 7 of them ($\mu_{\text{glc}}^{\max}$, $q_{\text{glc}}^{\max}$, $K_{\text{s,glc}}$, $K_{\text{s,xit}}$, $k_{\text{i,xyl}}$, K_{r}, P_{xit}) they have CI/IP percentages above 100%, which implies that the parameters could take negative values. This means that the degree of correlation between them is very large and its effect could be offset by changes in the other variables. Therefore, these cannot be reliably identified from the experimental data. This same problem had already been identified by Hernández-Escoto et al. (2016). Besides, there are parameters with very large confidence intervals ($K_{\text{s,xit}}$, $k_{\text{i,glc}}$, among others), which show an over-parameterization of the model.

Table 3.5. Model parameters, original (ORI), re-estimated using a interior-point algorithm (IP), 95% confidence interval (±95% CI) and relative deviation.

Parameter	ORI	IP	±95% CI	CI/IP (%)
$\mu_{\text{glc}}^{\max}$	0.662	0.044	2.3932	5465.76
$\mu_{\text{xit}}^{\max}$	0.189	0.059	0.0035	5.94
$q_{\text{xyl}}^{\max}$	0.342	0.315	0.0854	27.07
$q_{\text{glc}}^{\max}$	3.276	6.380	328.1336	5142.77
$K_{\text{s,xyl}}$	11.761	8.219	5.7057	69.42
$K_{\text{s,glc}}$	9.998	15.992	953.4967	5962.33
$K_{\text{s,xit}}$	16.068	$2.443 \cdot 10^{-5}$	$6.8550 \cdot 10^{-5}$	280.58
$k_{i,xyl}$	14.780	15.167	$2.3061 \cdot 10^{3}$	15204.11
$k_{i,\text{glc}}$	0.1	0.219	0.2124	96.69
K_r	0.1	20	80.7648	403.82
P_{xit}	$7.591 \cdot 10^{-9}$	$1.262 \cdot 10^{-6}$	$5.1968 \cdot 10^{-5}$	4118.20

ORI: original model parameters obtained by Tochampa et al. (2005). IP: Parameter values obtained from the interior-pint algorithm for deterministic model. The units of each parameter are presented in Table 3.1.

Figure 3.3 shows the predictions of the xylitol model readjusted with the confidence intervals of the predictor, predictions of the original model, and the experimental data. It can be seen that there is a good fit between the deterministic model and the experimental data, which are within the confidence interval of the model predictions for glucose and xylitol. However, the biomass and xylitol prediction leaves the experimental data at the limit or even outside the limit. This shows that the numerical method loses precision when it comes to representing the behavior of xylose and biomass, but it fits very well with the compound of interest. The results are in agreement with the analysis of the confidence intervals that show an over-parameterization of the model, implying that several parameters can be compensated by others and therefore improve the adjustment of some substances while sacrificing that of others.

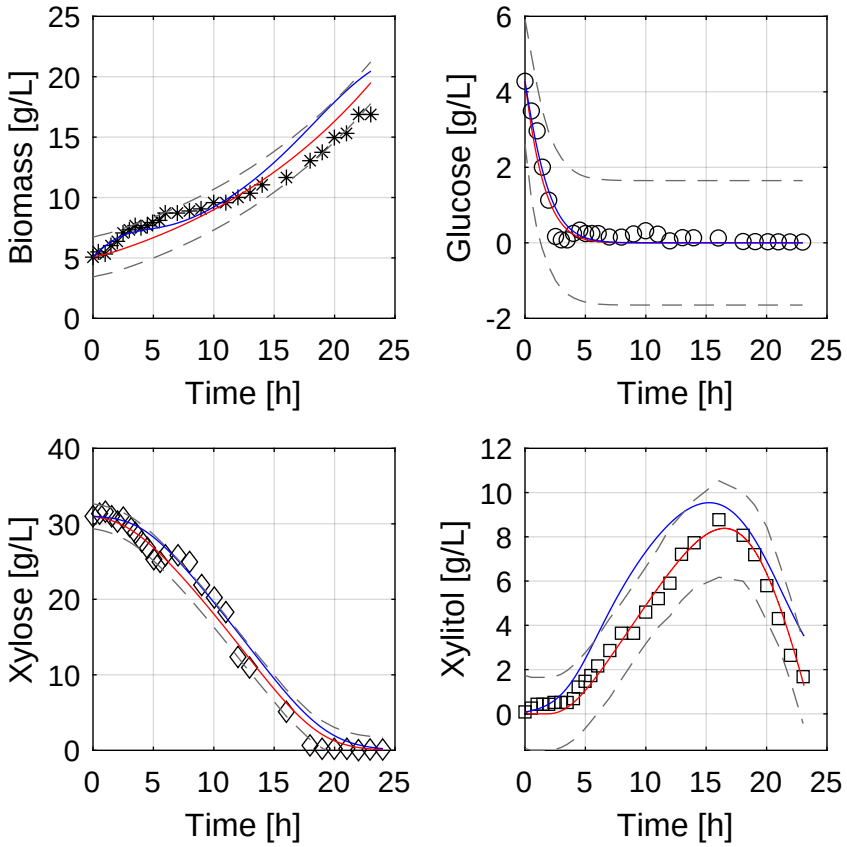

Figure 3.3. Concentration profiles of Biomass (∗), Glucose (◯), Xylose (◊), and Xylitol (□). The solid red line represents deterministic model results with the re-estimated parameters. The solid blue line represents the deterministic model results by Tochampa et al. (2005). Gray dashed lines: predictor confidence intervals and markers represent the experimental data for each species.

Table 3.6 presents the values found for the kinetic parameters from the ABC method. The values are very similar to those found with the deterministic method. Being the $k_{i,xyl}$, the parameter with the highest percentage of difference (10%) between the values estimated by both methods. Besides, the confidence intervals (uncertainty) are small and with practically homogeneous deviation percentages between the different parameters, compared to those calculated for the deterministic method. This shows that the stochastic method offers greater robustness and reliability when estimating the uncertainty of the model, compared to the deterministic method that presents very large

and therefore unrealistic uncertainties (Liu & Gunawan, 2017; Soize, 2013).

Table 3.6. Estimated parameters calculated from ABC rejection sampler method and stochastic model.

Parameter	SP	±95% CI	CI/IP (%)
$\mu_{\text{glc}}^{\text{max}}$	0.044	0.0020	4.56
$\mu_{\text{xit}}^{\text{max}}$	0.060	$2.8293 \cdot 10^{-4}$	0.47
$q_{\text{xyl}}^{\text{max}}$	0.309	0.0054	1.74
$q_{\text{glc}}^{\text{max}}$	6.788	0.2777	4.09
$K_{\text{s,xyl}}$	7.785	0.3821	4.91
$K_{\text{s,glc}}$	15.463	0.6407	4.14
$K_{\text{s,xit}}$	$2.466 \cdot 10^{-5}$	$1.1831 \cdot 10^{-6}$	4.80
$k_{i,xyl}$	16.696	0.6893	4.13
$k_{i,\text{glc}}$	0.227	0.0106	4.65
K_r	21.008	0.9645	4.59
P_{xit}	$1.294 \cdot 10^{-6}$	$5.8511 \cdot 10^{-8}$	4.52

SP: parameter values obtained from the ABC rejection sampler method ($V = 1 \times 10^{-20}$ L). The units of each parameter are presented in Table 3.1.

Figure 3.4 presents the mean behavior as well as the uncertainty associated with each of the states of the system from the stochastic simulation results. The parameter values correspond to those presented in column DP of Table 3.5 and the initial conditions presented in Table 3.3. When comparing the deterministic model predictions with the re-estimated parameters versus the stochastic model, it is observed that the greatest difference is for xylitol concentration. This is because the simulation was run with a system volume of $5 \cdot 10^{-22}$ L to increase the uncertainty of the method until containing the experimental data. In this way, the model allows evaluating the different scenarios on which the concentrations of the species could move, as well as determining the most probable behavior of the system from the average.

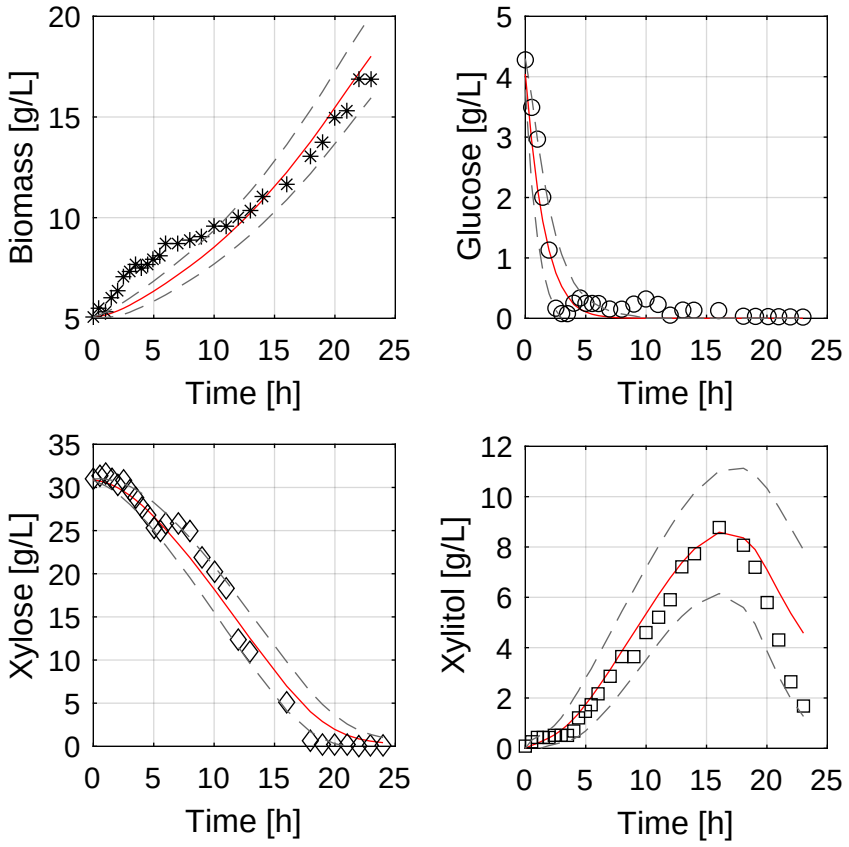

Figure 3.4. Concentration profiles of Biomass (∗), Glucose (◯), Xylose (◊), and Xylitol (□). The solid red line represents deterministic model results. Gray dashed lines: one-standard deviation envelope, $x \pm sdev(x)$, and markers represent the experimental data for each species.

3.4 Conclusions

Through this research, it was possible to Translate a semi-empirical model lacking elementary chemical reactions for microbial growth, to a stochastic simulation. This was made through the selection and calculation of the probability intervals of the reaction channels for intra and extracellular xylitol based on the sign and the evaluation of the absolute value of the propensity, respectively. Simulation results show the predictive power of the stochastic approach to correctly describe the evolution of the species based on the experimental data as well as the range of uncertainty over which the model

moves. This is a test of the validity of the assumptions made in this work to apply the stochastic QSSA and although the unique result is lost, since each simulation is different, the uncertainty estimate is gained.

On the other hand, it was possible to improve the fit of the deterministic model compared to the results previously reported in the literature and even the parameters can be re-estimated from stochastic simulation. Although the parameter estimation ABC method is slower than the deterministic method, the gain of its execution lies in the possibility of analyzing the uncertainty of the model more realistically. Therefore, this work leaves the doors open to simulate and optimize processes under uncertainty from purely stochastic models.

Bibliography

Barik, D., Paul, M. R., Baumann, W. T., Cao, Y., & Tyson, J. J. (2008). Stochastic simulation of enzyme-catalyzed reactions with disparate timescales. *Biophysical Journal*, *95*(8), 3563-3574. doi: 10.1529/biophysj.108.129155

Dasgupta, D., Bandhu, S., Adhikari, D. K., & Ghosh, D. (2017). Challenges and prospects of xylitol production with whole cell bio-catalysis: A review. *Microbiological research*, *197*, 9–21.

Dóka, G., Éva Lente. (2012). Stochastic mapping of the Michaelis-Menten mechanism. *The Journal of Chemical Physics*, *136*(5), 054111. doi: 10.1063/1.3681942

Dorantes-Landa, D. N., Cocotle-Ronzón, Y., Morales-Cabrera, M. A., & Hernández-Martínez, E. (2020). Modeling of the xylitol production from sugarcane bagasse by immobilized cells. *Journal of Chemical Technology & Biotechnology*, *95*(7), 1936–1945.

Edelstein, S., Smith, K., Worthington, A., Gillis, N., Bruen, D., Kang, S. H., … Guiducci, G. (2007). Comparisons of six new artificial sweetener gradation ratios with sucrose in conventional-method cupcakes resulting in best percentage substitution ratios. *Journal of Culinary Science & Technology*, *5*(4), 61–74.

Englezos, P., & Kalogerakis, N. (2000). *Applied parameter estimation for chemical engineers*. CRC Press.

Gillespie, D. T. (2007, apr). Stochastic Simulation of Chemical Kinetics. *Annual Review of Physical Chemistry*, *58*(1), 35–55. Retrieved from `https://doi.org/10.1146/annurev.physchem.58.032806.104637` doi: 10.1146/annurev.physchem.58.032806.104637

Gupta, A. (2013). *Parameter estimation in deterministic and stochastic models of biological systems.* The University of Wisconsin-Madison.

Hahl, S. K., & Kremling, A. (2016). A comparison of deterministic and stochastic modeling approaches for biochemical reaction systems: On fixed points, means,

and modes. *Frontiers in genetics*, *7*, 157.

Hernández-Escoto, H., Prado-Rubio, O. A., & Morales-Rodriguez, R. (2016). Model-based framework for enhanced and controlled operation of a fed-batch bioreactor: xylitol production. In *Computer aided chemical engineering* (Vol. 38, pp. 301–306). Elsevier.

J. Donoso. (2006). *Biopolímeros.* Retrieved from `http://facultatciencies.uib.cat/prof/josefa.donoso/campus/modulos/modulo4/modulo4{_}14.htm`

Kang, H.-W., KhudaBukhsh, W. R., Koeppl, H., & Rempała, G. A. (2019). Quasi-steady-state approximations derived from the stochastic model of enzyme kinetics. *Bulletin of mathematical biology*, *81*(5), 1303–1336.

Koop, L., Corazza, M. L., Voll, F. A. P., & Bonilla-Petriciolet, A. (2017). Optimal control of a fermentation process for xylitol production using differential evolution. *Differential Evolution in Chemical Engineering*, *6*, 321–351.

Lawson, M. J., Petzold, L., & Hellander, A. (2015). Accuracy of the Michaelis–Menten approximation when analysing effects of molecular noise. *Journal of The Royal Society Interface*, *12*(106), 20150054.

Liu, Y., & Gunawan, R. (2017). Bioprocess optimization under uncertainty using ensemble modeling. *Journal of biotechnology*, *244*, 34–44.

Maguire, A., & Rugg-Gunn, A. J. (2003). Xylitol and caries prevention—is it a magic bullet? *British dental journal*, *194*(8), 429–436.

Marchetti, L., Priami, C., & Thanh, V. H. (2017). Stochastic Simulation of Biochemical Reaction Systems. In *Simulation algorithms for computational systems biology* (pp. 7–28). Springer.

Menz, S. (2013). *Hybrid stochastic-deterministic approaches for simulation and analysis of biochemical reaction networks.*

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley Sons.

Myers, R. H., Montgomery, D. C., Vining, G. G., & Robinson, T. J. (2012). *Generalized linear models: with applications in engineering and the sciences* (Vol. 791). John Wiley & Sons.

Nguyen, C. C., Le, L. T., & Boontawan, A. (2020). High Yield of Xylitol Fermentation from Commercial Xylose by Candida guilliermondii TISTR 5068.

Park, S. M., Sang, B. I., Park, D. W., & Park, D. H. (2005). Electrochemical reduction of xylose to xylitol by whole cells or crude enzyme of Candida peltata. *The Journal of Microbiology*, *43*(5), 451–455.

Prado-Rubio, O. A., Hernández-Escoto, H., Rodriguez-Gomez, D., Sirisansaneeyakul,

S., & Morales-Rodriguez, R. (2015). *Enhancing Xylitol Bio-Production by an Optimal Feeding Policy during Fed-Batch Operation* (Vol. 37) (No. June). Elsevier Inc. doi: 10.1016/B978-0-444-63577-8.50138-8

Rao, C. V., & Arkin, A. P. (2003). Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *The Journal of chemical physics*, *118*(11), 4999–5010.

Rogers, A., & Gibon, Y. (2009). Enzyme kinetics: theory and practice. In *Plant metabolic networks* (pp. 71–103). Springer.

Ruiz Paredes, C. F. (2019). *Métodos de punto interior - Un nuevo método de punto interior* (Unpublished doctoral dissertation). Universidad Tecnológica de Pereira.

Sirisansaneeyakul, S., Wannawilai, S., & Chisti, Y. (2013). Repeated fed-batch production of xylitol by Candida magnoliae TISTR 5663. *Journal of Chemical Technology and Biotechnology*, *88*(6), 1121–1129. doi: 10.1002/jctb.3949

Soize, C. (2013). Stochastic modeling of uncertainties in computational structural dynamics—recent theoretical advances. *Journal of Sound and Vibration*, *332*(10), 2379–2395.

Tochampa, W., Sirisansaneeyakul, S., Vanichsriratana, W., Srinophakun, P., Bakker, H. H. C., & Chisti, Y. (2005). A model of xylitol production by the yeast Candida mogii. *Bioprocess and Biosystems Engineering*, *28*(3), 175–183. doi: 10.1007/s00449-005-0025-0

Tochampa, W., Sirisansaneeyakul, S., Vanichsriratana, W., Srinophakun, P., Bakker, H. H. C., Wannawilai, S., & Chisti, Y. (2015). Optimal control of feeding in fed-batch production of xylitol. *Industrial & Engineering Chemistry Research*, *54*(7), 1992–2000.

Villadsen, J., Nielsen, J., & Lidén, G. (2011). *Bioreaction engineering principles*. Springer Science & Business Media.

Warne, D. J., Baker, R. E., & Simpson, M. J. (2019). Simulation and inference algorithms for stochastic biochemical reaction networks: from basic concepts to state-of-the-art. *Journal of the Royal Society Interface*, *16*(151), 20180943.

Chapter 4

Conclusions and perspectives

In this research, it was possible to develop a stochastic model based on the SSA and to demonstrate the ability of these to predict the behavior of enzymatic systems with and without inhibition from the kinetic mass action law and Michaelis-Menten kinetics. Through, it was established that the size of the system would directly determine the uncertainty of the model and therefore the ability to reproduce or not the uncertainty of experimental measurements. Besides, it was determined that the uncertainty of the model tends to remain constant after a certain number of realizations, which allows optimizing the execution time of the stochastic model.

The stochastic model of xylitol production shows that despite the lack of knowledge of the elemental reactions that define microbial growth, it is possible to apply the stochastic QSSA method and the SSA to simulate this process. The results show the same predictive power of the deterministic model, besides, to provide information on the uncertainty of the model for the different system states. On the other hand, the parameter re-estimation for the deterministic model with the interior-point algorithm allowed to obtain better predictions of xylitol and thus increase model predictive power. From the stochastic calibration method was possible to obtain a more robust description of the uncertainty of the model by providing reliable confidence intervals for the parameters, compared to the deterministic calibration method.

Among the major contributions of this work are the extension of the stochastic chemical kinetics and stochastic QSSA methods for the cases of enzymatic reactions with inhibition. As well as, the uncertainty analyzes obtained from the modeling for each state of the system and its relationship with the relative uncertainty of the experimental measurements. Another contribution of this work was to adapt the stochastic QSSA methodology to a highly complex case such as xylitol bioproduction and couple

it with a stochastic parameter estimation method. It is noteworthy that to date only applications of these methodologies appear in the literature for generic cases or with the kinetics of enzymatic reactions without inhibition.

This work is encouraging to continue delving into the application of stochastic methods for cases in specific bioprocesses in the area of enzymatic reactions with microbial growth, which also include inhibition and mass transport processes. By modeling the system directly with a method of random nature, the modeling and optimization of these systems under uncertainty is possible, also because the applicability of Bayesian methods (ABC) for parameter estimation was shown.