THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Mixed Effects Modeling of Deterministic and Stochastic Dynamical Systems

Methods and Applications in Drug Development

Jacob Leander

CHALMERS
UNIVERSITY OF TECHNOLOGY

Division of Applied Mathematics and Statistics
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
Gothenburg, Sweden 2021

Mixed Effects Modeling of Deterministic and Stochastic Dynamical Systems
Methods and Applications in Drug Development
Jacob Leander
Gothenburg 2021
ISBN 978-91-7905-450-2

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 4917
ISSN 0346-718X

Division of Applied Mathematics and Statistics
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Sweden
Telephone +46 (0)31 772 1000

Department of Systems and Data Analysis
Fraunhofer-Chalmers Research Centre for Industrial Mathematics
Chalmers Science Park
SE-412 88 Gothenburg
Sweden
Telephone +46 (0)31 309 94 00

Typeset with LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2021

Mixed Effects Modeling of Deterministic and Stochastic Dynamical Systems

Methods and Applications in Drug Development

Jacob Leander

Division of Applied Mathematics and Statistics
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg

Abstract

Mathematical models based on ordinary differential equations (ODEs) are commonly used for describing the evolution of a system over time. In drug development, pharmacokinetic (PK) and pharmacodynamic (PD) models are used to characterize the exposure and effect of drugs. When developing mathematical models, an important step is to infer model parameters from experimental data. This can be a challenging problem, and the methods used need to be efficient and robust for the modeling to be successful. This thesis presents the development of a set of novel methods for mathematical modeling of dynamical systems and their application to PK-PD modeling in drug development.

A method for regularizing the parameter estimation problem for dynamical systems is presented. The method is based on an extension of ODEs to stochastic differential equations (SDEs), which allows for stochasticity in the system dynamics, and is shown to lead to a parameter estimation problem that is easier to solve.

The combination of parameter variability and SDEs are investigated, allowing for an additional source of variability compared to the standard nonlinear mixed effects (NLME) model. For NLME models with dynamics described using either ODEs or SDEs, a novel parameter estimation algorithm is presented. The method is a gradient-based optimization method where the exact gradient of the likelihood function is calculated using sensitivity equations, which is shown to give a substantial improvement in computational speed compared to existing methods. The methods developed have been integrated into NLMEModeling, a freely available software package for mixed effects modeling in Wolfram Mathematica. The package allows for general model specifications and offers a user-friendly environment for NLME modeling of dynamical systems.

The SDE-NLME framework is used in two applied modeling problems in drug development. First, a previously published PK model of nicotinic acid is extended to incorporate SDEs. By extending the ODE model to an SDE model, it is shown that an additional source of variability can be quantified. Second, the SDE-NLME framework is applied in a model-based analysis of peak expiratory flow (PEF) diary data from two Phase III studies in asthma. The established PEF model can describe several aspects of the PEF dynamics, including long-term fluctuations. The association to exacerbation risk is investigated using a repeated time-to-event model, and several characteristics of the PEF dynamics are shown to be associated with exacerbation risk.

The research presented in this doctoral thesis demonstrates the development of a set of methods and applications of mathematical modeling of dynamical systems. In this work, the methods were primarily applied in the field of PK-PD modeling, but are also applicable in other scientific fields.

Keywords: mathematical modeling, dynamical systems, mixed effects, parameter estimation, pharmacokinetics, pharmacodynamics, time-to-event, pharmacometrics, drug development

Acknowledgements

The road toward this thesis has, as most things in life, not been straightforward. I have met many people along the way that have supported me in my development as a researcher, pharmacometrician, and drug developer.

During 2012-2014, I had the opportunity to be one of the first students in the licentiate program Advanced Engineering Mathematics. The research program was a collaboration between Fraunhofer-Chalmers Centre and the Department of Mathematical Sciences at Chalmers University of Technology. In September 2014, I defended my licentiate thesis and later started at AstraZeneca, working with modeling and simulation to support decision making in drug development. In the beginning of 2017, I was given the opportunity to continue my research as a part-time industrial PhD student in a research collaboration with AstraZeneca, Fraunhofer-Chalmers Centre, and the Department of Mathematical Sciences. It has been truly exciting years, working collaboratively with a pharmaceutical company, a research institute, and academia.

First, I would like to thank my main supervisor Mats Jirstrand, who has stood by my side during both licentiate and doctoral thesis. Thanks for all your guidance and for educate me in the wonderful world of mathematical modeling. I also want to thank Torbjörn Lundh, who was my main supervisor during my licentiate thesis. Thank you Marija Cvijovic, who has been my second supervisor during my doctoral thesis. The work presented in this thesis has partially been founded by the Swedish Foundation for Strategic Research and a research grant from AstraZeneca, which is gratefully acknowledged.

I have met a lot of people during my research project, and this thesis would not have been possible without input and guidance from all of you. Thanks to all my colleagues in the Systems and Data Analysis group at Fraunhofer-Chalmers Centre. A special recognition to Joachim Almquist for all the discussions, jokes, manuscripts, code development, and late evenings we have had together during these years. Thanks to my colleagues in the Respiratory & Immunology

Clinical Pharmacology group at AstraZeneca, and a big thank you to all the pharmacometricians at AstraZeneca Gothenburg for all the valuable discussions and support. Thank you Robert Palmér and Ulrika Wählby Hamrén for always being excellent discussion partners and problem solvers. I also want to thank all members of CvijovicLab at the Department of Mathematical Sciences.

Thank you Ziad Taib and Daniel Röshammar who took me under your wings in my master thesis project back in 2012. It was thanks to that master thesis project I ended up at Fraunhofer-Chalmers Centre which paved the way for my current adventures at AstraZeneca.

I also want to thank Ulf Eriksson for always pushing me and letting me complete my research as an industrial PhD student. Combining project work and research in mathematics has not always been an easy ride, but I am grateful and proud of what I have accomplished.

I want to express my deepest gratitude to my parents Kent-Erik and Monica, for your constant support and for always being there for me. Thanks to my sister Lisa who always stand by my side. Finally, thanks to my wonderful wife Annie and our lovely children Elsie and Hugo. You mean the world to me!

List of Publications

This thesis is based on the following appended papers:

Paper I. **Leander J**, Lundh T, and Jirstrand M (2014). Stochastic differential equations as a tool to regularize the parameter estimation problem for continuous time dynamical systems given discrete time measurements. Math. Biosci. 251(1):54-62

Paper II. **Leander J**, Almquist J, Ahlström C, Gabrielsson J, and Jirstrand M (2015). Mixed Effects Modeling Using Stochastic Differential Equations: Illustrated by Pharmacokinetic Data of Nicotinic Acid in Obese Zucker Rats. AAPS J. 17(3):586-596

Paper III. Almquist J, **Leander J**, and Jirstrand M (2015). Using sensitivity equations for computing gradients of the FOCE and FOCEI approximations to the population likelihood. J. Pharmacokinet. Pharmacodyn. 42(3):191-209

Paper IV. Ólafsdóttir HK, **Leander J**, Almquist J, and Jirstrand M (2018). Exact Gradients Improve Parameter Estimation in Nonlinear Mixed Effects Models with Stochastic Dynamics. AAPS J. 20(5):1-13

Paper V. **Leander J**, Almquist J, Johnning A, Larsson J, and Jirstrand M (2020). NLMEModeling: A Wolfram Mathematica Package for Nonlinear Mixed Effects Modeling of Dynamical Systems. Manuscript, arXiv:2011.06879

Paper VI. **Leander J**, Jirstrand M, Eriksson UG, and Palmér R (2021). A stochastic mixed effects model to assess treatment effects and fluctuations in home-measured peak expiratory flow and the association to exacerbation risk in asthma. Manuscript, submitted

The author has contributed to the papers as follows.

Paper I. Designed the research and performed the analysis. Derived the sensitivity equations for the extended Kalman filter and implemented the estimation algorithm in Wolfram Mathematica. Drafted and wrote the manuscript.

Paper II. Designed the research and performed the analysis. Implemented parts of the mixed effects estimation algorithm in Wolfram Mathematica. Drafted and wrote the manuscript.

Paper III. Part of designing the research and analysis. Part of deriving the sensitivity equations for the FOCE algorithm. Implemented parts of the mixed effects estimation algorithm in Wolfram Mathematica. Revised and edited the manuscript.

Paper IV. Part of designing the research and analysis. Part of deriving the sensitivity equations for the SDE-FOCE algorithm. Implemented parts of the mixed effects estimation algorithm in Wolfram Mathematica. Revised and edited the manuscript.

Paper V. Co-developer of the NLMEModeling package. Created all examples in the manuscript and produced all results. Drafted and wrote the manuscript. Note that a shortened version of the manuscript has been accepted as a peer-reviewed conference paper for the 19th IFAC Symposium on System Identification (SYSID 2021).

Paper VI. Part of designing the research and analysis. Implemented the estimation algorithm, developed the model, and performed the post-processing in Wolfram Mathematica. Performed the repeated time-to-event analysis in R. Drafted and wrote the manuscript.

Other relevant publications co-authored by Jacob Leander not included in this thesis.

Tapani S, Almquist J, **Leander J**, Ahlström C, Peletier LA, Jirstrand M, and Gabrielsson J (2014). Joint feedback analysis modeling of nonesterified fatty acids in obese Zucker rats and normal Sprague-Dawley rats after different routes of administration of nicotinic acid. J. Pharm. Sci., 103(8):2571-2584

Hegelund Myrbäck T, Prothon S, Edman K, **Leander J**, Hashemi M, Dearman M, Edenro G, Svanberg P, Andersson EM, Almquist J, Ämmälä C, Hendrickx R, Taib Z, Johansson K, Berggren A, Keen C, Eriksson UG, Fuhr R, and Carlsson B (2020). Effects of a selective glucocorticoid receptor modulator (AZD9567) versus prednisolone in healthy volunteers: two phase 1, single-blind, randomised controlled trials. Lancet Rheumatol. 2(1):31-41

Almquist J, Sadiq MW, Eriksson UG, Hegelund Myrbäck T, Prothon S, and **Leander J** (2020). Estimation of Equipotent Doses for Anti-Inflammatory Effects of Prednisolone and AZD9567, an Oral Selective Nonsteroidal Glucocorticoid Receptor Modulator. CPT Pharmacometrics Syst. Pharmacol., 9(8): 444-455

Rekić D, Johansson S, and **Leander J** (2020). Higher Febuxostat Exposure Observed in Asian Compared with Caucasian Subjects Independent of Bodyweight. Clin. Pharmacokinet. 60(3):319-328

Leander J, Sunnåker M, Rekić D, Aksenov S, Eriksson UG, Johansson S, and Parkinson J (2021). A semi-mechanistic exposure-response model to assess the effects of verinurad, a potent URAT1 inhibitor, on serum and urine uric acid in patients with hyperuricemia-associated diseases. J. Pharmacokinet. Pharmacodyn.

Abbreviations

AIC	Akaike information criterion
BFGS	Broyden-Fletcher-Goldfarb-Shanno
BIC	Bayesian information criterion
EKF	Extended Kalman filter
FOCE	First-order conditional estimation
FOCEI	First-order conditional estimation with interaction
MIDD	Model-informed drug development
NiAc	Nicotinic Acid
NLME	Nonlinear mixed effects
ODE	Ordinary differential equation
PD	Pharmacodynamic(s)
PEF	Peak expiratory flow
PK	Pharmacokinetic(s)
RTTE	Repeated time-to-event
SDE	Stochastic differential equation
SDEMEM	Stochastic differential equation mixed effects model
VPC	Visual predictive check

Contents

1 Introduction

This chapter gives a short introduction to this doctoral thesis. The idea of modeling is described, with emphasis on mathematical modeling. The application of mathematical models in drug development is highlighted, leading into the motivation and aims for this thesis.

In Chapter 2, the most important mathematical concepts considered throughout the appended papers are described. In Chapter 3, the results from the six appended papers are presented and discussed. The thesis ends with further discussions and future perspectives in Chapter 4.

1.1 The Idea of Modeling

Modeling is the process of representing real-world systems in a logical and objective way, where a system can be described as a group of interrelated entities that are integrated to accomplish a common goal. Starting with a problem or question, the system of interest can be translated into a model where the problem can be studied in a simplified context. Hence, a model should always be developed with a purpose in mind. The modeling process requires the modeler to carefully identify relevant aspects of the system, and to transfer the necessary components of the system into the model. Depending on the level of detail of interest, models might have different complexity. There are several different classes of models, depending on how they represent the system of interest. In this work, we will only consider quantitative models, and the interested reader is referred to Gerlee and Lundh (2016) for a broad introduction to modeling and its applications.

A quantitative model, also known as a mathematical model, is a formal representation and description of a system. Eykhoff (1974) described a mathematical model as a representation of the essential aspects of an existing system, which

presents knowledge of that system in usable form. Described in other terms, a mathematical model is a simplification of a system using mathematical concepts and language. Mathematical models can take a variety of forms, including dynamical models, statistical models, and agent-based models. The applications of mathematical models are enormous, spanning from natural sciences (e.g., physics, chemistry, and biology) and engineering (e.g., automotive and electronics) to social sciences (e.g., economics and sociology). From now on, the terms 'mathematical model' and 'model' will be used interchangeably. Note that this should not be confused with animal models, commonly used in preclinical experiments in drug development.

Once the system has been represented as a mathematical model, the problem at hand can be studied in an isolated context. The results from the analysis of the mathematical model then needs to interpreted, providing insights and answers to the questions. As knowledge increases and new insights are gained, the modeling process often needs to be repeated one or several times. A conceptual illustration of the mathematical modeling process, and its iterative nature, is illustrated in Figure 1.1.

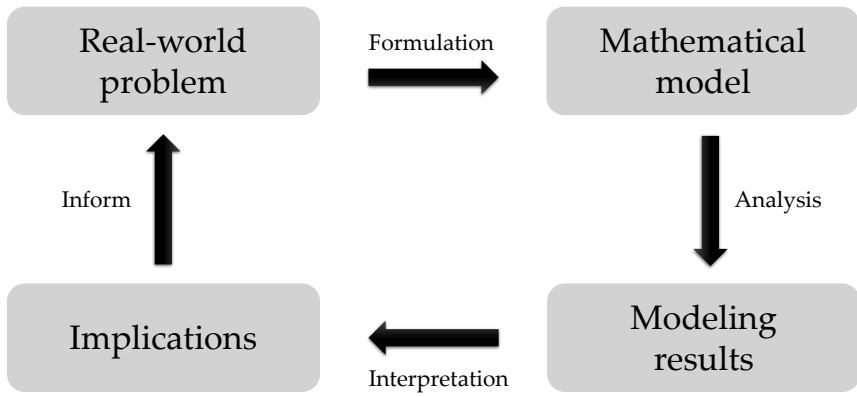


Figure 1.1: Conceptual illustration of the mathematical modeling process. Starting with a problem in the real world, a mathematical model is created. The model produces certain results, which then needs to be interpreted to inform the initial problem. As knowledge increase, several iterations of the modeling process might be necessary.

1.2 Mathematical Modeling in Drug Development

Drug development is the process by which new therapies are created and brought to market to treat diseases (Wood 2006). The development of new drugs is a long and expensive process (DiMasi et al. 2016), involving many different development phases such as discovery of a suitable target, compound development, pre-clinical experiments, clinical studies in humans, and post-marketing studies.

Pharmacology, named from the Greek words *pharmakon* (drug/poison) and *-logia* (study of/knowledge), is a cornerstone of drug development which aims to understand how chemical substances (man-made or natural) interacts with living organisms. Two important concepts in pharmacology encountered throughout this thesis is *pharmacokinetics* and *pharmacodynamics*. Pharmacokinetics (PK), often described as what the body does to the drug, deals with how the concentration of drugs at various sites in the body changes over time. The PK encompasses several different phases, such as drug absorption, distribution, metabolism, and elimination. Pharmacodynamics (PD), often described as what the drug does to the body, deals with the time course of drug action and effect and is (in most cases) closely linked to the PK. The PD effect could for example be changes in glucose concentration, blood pressure, or disease severity. These two important concepts are often studied simultaneously, in a framework named pharmacokinetic-pharmacodynamic (PK-PD) modeling. Several different examples of PK-PD models will be encountered throughout this thesis.

The application of mathematical models in general, and PK-PD models in particular, has increased in popularity in drug development during the last decades. It has evolved into its own discipline named pharmacometrics, which has been defined as the branch of science concerned with mathematical models of biology, pharmacology, disease, and physiology to describe and quantify interactions between drugs and patients (Barrett et al. 2008; Ette and Williams 2007).

Today, a wide spectrum of quantitative methods are used on a regular basis to facilitate decision making in the drug development process, often referred to as model-informed drug development (MIDD) (Helmlinger et al. 2017; Kimko and Pinheiro 2015; Marshall et al. 2016; Milligan et al. 2013). MIDD can be used to weigh risks and benefits throughout the development phases, and to provide answers to important questions: How does the drug behave in the body? What is the best dose? Which patients are likely to benefit from the drug? What is the probability of success?

1.3 Every Patient is Different

As humans are different, we also respond differently to drug therapies. The differences might be attributed to known *covariates*, such as gender, ethnicity, or weight. There might also be other sources of variation that are not known explicitly. Often, the aim is to understand the sources of variation and quantify the degree of variability within a population of interest.

A statistical framework suitable for incorporating variability between individuals is the mixed effects model. The 'mixed' term refers to the fact that the model incorporates both *fixed effects* (parameters assumed to be the same for all individuals) and *random effects* (parameters assumed to be different between individuals). The models encountered throughout this thesis deal with mixed effects models where the model parameters occur nonlinearly, resulting in the nonlinear mixed effects (NLME) model (Davidian and Giltinan 1995; Lindstrom and Bates 1990). NLME modeling has been applied in several scientific fields, including image analysis (Bilgel et al. 2016; Chen et al. 2013), forestry (Sirkiä et al. 2015), and not least pharmacometrics.

In pharmacometrics, the NLME model has become the standard approach for analyzing data from multiple individuals, where it is used to quantify inter-individual variability in for example drug exposure and effect. In 1977, Lewis Sheiner published a seminal paper, describing the estimation of inter-individual variability in PK parameters (Sheiner et al. 1977). The paper was followed by a series of papers by Lewis Sheiner and Stuart Beal (Sheiner and Beal 1980; Sheiner and Beal 1981; Sheiner and Beal 1983), and their models and estimation methods are now incorporated in the NONMEM program, one of the most popular software for NLME modeling in drug development (Beal et al. 2017).

1.4 Describing the Dynamics

In PK-PD modeling, the underlying system of interest is often described mathematically using ordinary differential equations (ODEs). The application of ODEs in PK-PD modeling is diverse, spanning from rather simple PK models describing change in concentration in different compartments to comprehensive quantitative systems pharmacology models consisting of a large number of interacting entities (Coletti et al. 2020).

Using ODEs, the evolution of the system dynamics is by definition determin-

istic, and any deviations between model predictions and observed data are usually described using a random variable (typically referred to as the measurement error). As the deterministic model does not account for uncertainty in the underlying dynamical system, it attempts to represent the average behavior of the system (Irurzun-Arana et al. 2020). The deterministic description might be an appropriate model in many cases, while in some cases the introduction of stochasticity in the dynamical model is deemed more appropriate.

One way of introducing stochasticity in the dynamical model is to consider the extension to stochastic differential equations (SDEs). SDEs are a flexible class of models describing the evolution of a stochastic process. SDEs have successfully been applied in several fields, including finance, electrical engineering, and control theory (Black and Scholes 1973; Jazwinsky 1970; Åström 1970). One of the first applications of SDEs in PK-PD modeling was published by Kristensen et al. (2005), where the authors proposed using SDEs to improve the structure of the dynamical model.

The combination of SDEs and NLME is an intriguing approach, as it enables characterization of three sources of variability in the observed data: inter-individual variability, measurement error, and stochasticity in the system dynamics. The stochastic differential equation mixed effects model (SDEMEM) framework was published during 2000s (Ditlevsen and De Gaetano 2005; Overgaard et al. 2005; Tornøe et al. 2005). Since its introduction, it has been used in several applications, including PK modeling, neuronal signaling, and oncology (Berglund et al. 2012; Matzuka et al. 2016; Picchini et al. 2008; Picchini and Forman 2019).

1.5 Motivation and Aims of This Thesis

The development and application of mathematical models is important in many scientific areas. Once a model has been developed, it can be used to shed light upon new questions and hypotheses. Hence, a model is a powerful tool as it not only can be used retrospectively, but also to inform and guide the design of future experiments. In drug development, the application of modeling and simulation can accelerate the development of new drugs, reduce patient burden, and build new knowledge.

The research in this thesis considers several aspects of mathematical modeling of dynamical systems. The main focus has been on NLME modeling, both in terms of method development and application of novel methodology. The research addresses the following three aims:

Aim 1: To investigate the use of stochastic dynamical models to improve understanding and description of the underlying system of interest.

Aim 2: To contribute to the development of methodology for parameter estimation in dynamical systems, with focus on computational performance and robustness.

Aim 3: To apply novel methodology to address relevant questions within the area of drug development.

2 Methods

This chapter describes the most important mathematical concepts used throughout the appended papers. In the following sections, bold font is used to denote vectors and matrices.

2.1 Ordinary Differential Equations

One of the most frequently used tools for mathematical modeling in engineering and life sciences is differential equations. An *ordinary* differential equation is a differential equation containing one or more functions of one independent variable (which here will be time) and the derivatives of those functions. To describe the underlying dynamical system of interest we will in this thesis consider a set of first-order ODEs of the form

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t, \boldsymbol{\theta}), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0(\boldsymbol{\theta}), \tag{2.1}$$

where t denotes time, $\boldsymbol{x}$ is the vector of state variables, $\boldsymbol{u}$ is the known system input, and $\boldsymbol{\theta}$ is the vector of model parameters. The function $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t, \boldsymbol{\theta})$ describes the dynamics of the system. The initial condition, given at time $t = t_0$, is described by $\boldsymbol{x}_0(\boldsymbol{\theta})$, which may depend on the model parameters. Examples of inputs, $\boldsymbol{u}$, in PK-PD modeling applications could be different dosing regimens, such as infusion or oral dosing.

Furthermore, the system is assumed to be observed at discrete time points. The observation model is described by

$$\boldsymbol{y}_j = \boldsymbol{h}(\boldsymbol{x}(t_j), \boldsymbol{u}(t_j), t_j, \boldsymbol{\theta}) + \boldsymbol{e}(t_j), \quad j = 1, \ldots, J, \tag{2.2}$$

where $\boldsymbol{y}_j$ is a vector of output variables at time point t_j and $\boldsymbol{h}(\boldsymbol{x}(t_j), \boldsymbol{u}(t_j), t_j, \boldsymbol{\theta})$ is a function describing the observation model. The observation error $\boldsymbol{e}(t_j)$ is

assumed to be normally distributed with mean zero and covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{x}(t_j), \boldsymbol{u}(t_j), t_j, \boldsymbol{\theta})$.

2.2 Sensitivity Analysis

Sensitivity analysis is the study of how perturbations affect the output of a mathematical model or system. Different types of perturbations might be considered, including perturbations in the input, initial conditions, and/or the model parameters.

Parameter sensitivity analysis can serve as tool to investigate how sensitive the model is with respect to different model parameters. This can for example be used to guide experimental design and to reduce model complexity. For gradient-based parameter optimization problems, parameter sensitivity analysis can be utilized to calculate the gradient of the objective function, which is considered throughout the appended papers.

The first-order sensitivity equations for an ODE are obtained by differentiating the ODE with respect to the parameters. This leads to a set of new differential equations describing the evolution of the sensitivity of the states (Dickinson and Gelinas 1976). The first-order sensitivities, $\frac{d\boldsymbol{x}}{d\theta_k}$, of the state variables $\boldsymbol{x}$ with respect to a parameter θ_k, where $k = 1, \ldots, p$, with p being the number of model parameters, is given by the solution to the differential equation

$$\frac{d}{dt}\frac{d\boldsymbol{x}}{d\theta_k} = \frac{\partial \boldsymbol{f}}{\partial \theta_k} + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\frac{d\boldsymbol{x}}{d\theta_k}, \quad \frac{d\boldsymbol{x}}{d\theta_k}(t_0) = \frac{\partial \boldsymbol{x}_0(\boldsymbol{\theta})}{\partial \theta_k}. \tag{2.3}$$

For an ODE model with n state variables, depending on p parameters, the first-order state sensitivities lead to a set of $n \cdot p$ sensitivity equations. In general, the sensitivity equations depend on the state variables $\boldsymbol{x}$. This requires the sensitivity equations to be solved simultaneously as the original set of ODEs, leading to a set of $n\,(p+1)$ differential equations. In the same fashion, higher order parameter sensitivities can be obtained by differentiating the first-order sensitivity equations.

2.3 The Nonlinear Mixed Effects Model

Using ODEs, the evolution of the state variables is completely determined by the initial condition $\boldsymbol{x}(t_0)$, the input $\boldsymbol{u}(t)$, and the model parameters $\boldsymbol{\theta}$. Hence, the only source of variability in the observed data arises from the error in the observation model. Since different individuals are likely to show different dynamic behavior, it is motivated to consider an additional source of variability in the model, namely the inter-individual variability.

In an NLME model, the model for individual i is described by

$$\frac{d\boldsymbol{x}_i}{dt} = \boldsymbol{f}(\boldsymbol{x}_i, \boldsymbol{u}_i, t, \boldsymbol{\phi}_i), \quad \boldsymbol{x}_i(t_0) = \boldsymbol{x}_{i,0}(\boldsymbol{\phi}_i), \tag{2.4}$$

$$\boldsymbol{y}_{ij} = \boldsymbol{h}(\boldsymbol{x}_i(t_j), \boldsymbol{u}_i(t_j), t_j, \boldsymbol{\phi}_i) + \boldsymbol{e}(t_j), \tag{2.5}$$

where $\boldsymbol{\phi}_i$ denotes the parameters for individual i. The individual parameters are related to the population parameters $\boldsymbol{\theta}$ by the relationship

$$\boldsymbol{\phi}_i = \boldsymbol{\phi}(\boldsymbol{\theta}, \boldsymbol{Z}_i, \boldsymbol{\eta}_i), \tag{2.6}$$

where $\boldsymbol{\theta}$ denotes the fixed effects, $\boldsymbol{Z}_i$ denotes the known covariates for individual i, and $\boldsymbol{\eta}_i$ denotes the random effects. The random effects are assumed to be multivariate normally distributed

$$\boldsymbol{\eta}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Omega}), \tag{2.7}$$

where $\boldsymbol{\Omega}$ denotes the covariance matrix for the random effects.

The NLME model, also known as multilevel or hierarchical model, introduces an additional level in the statistical description of the data. By transformation of the random effects, via the functional relationship ϕ, different distributions of the individual parameters can be obtained, such as log-normal and logit-normal.

2.4 The Stochastic Differential Equation Mixed Effects Model

We will also consider dynamical models described by SDEs, allowing for stochasticity in the dynamics. In a stochastic differential equation mixed effects model (SDEMEM), the dynamics for individual i is described by an SDE of the

following form

$$d\boldsymbol{x}_i = \boldsymbol{f}(\boldsymbol{x}_i, \boldsymbol{u}_i, t, \boldsymbol{\phi}_i)dt + \boldsymbol{G}(\boldsymbol{x}_i, \boldsymbol{u}_i, t, \boldsymbol{\phi}_i)d\boldsymbol{W}_i, \quad \boldsymbol{x}_i(t_0) = \boldsymbol{x}_{i,0}(\boldsymbol{\phi}_i). \tag{2.8}$$

Here, $\boldsymbol{W}_i$ is a standard Wiener process, also known as Brownian motion, with increments $d\boldsymbol{W}_i \sim \mathcal{N}(0, dt\boldsymbol{I})$ where $\boldsymbol{I}$ denotes the identity matrix. The matrix $\boldsymbol{G}$ weights the increments, and we will refer to $\boldsymbol{G}(\boldsymbol{x}_i, \boldsymbol{u}_i, t, \boldsymbol{\phi}_i)d\boldsymbol{W}_i$ as the system noise. For SDEs, the function $\boldsymbol{f}(\boldsymbol{x}_i, \boldsymbol{u}_i, t, \boldsymbol{\phi}_i)$ is usually referred to as the drift. Note that in the case of $\boldsymbol{G} = \boldsymbol{0}$, the SDE reduces to an ODE. The solution $\boldsymbol{x}_i$ to the SDE is a stochastic process and lacks, in most cases, a closed form solution. For additional information regarding SDEs, see Øksendal (2003).

In models governed by SDEs, the underlying state is stochastic. The problem of inferring the underlying state from observations is commonly referred to as the filtering problem. To solve the filtering problem for models governed by SDEs, we will consider the extended Kalman filter (EKF) (Jazwinsky 1970). For linear dynamical system models, the Kalman filter provides an optimal state estimator for a given parameter vector $\boldsymbol{\phi}_i$. For nonlinear dynamical system models, the EKF uses a first-order linearization around the model prediction and provides estimates of the conditional expectation and covariance of the underlying state and output. Additional details regarding the EKF can be found in Paper I, Paper II, and Paper IV.

In contrast to the classical NLME framework which only considers two sources of variability in the observed data, the SDEMEM considers three sources of variability: inter-individual variability, system noise, and measurement error.

2.5 Estimating Model Parameters

One key step in the mathematical modeling process is the problem of inferring model parameters from experimental data, usually referred to as the parameter estimation problem. A popular method for estimation of parameters in a statistical model is the maximum likelihood approach. Maximum likelihood estimation aims to find the values of the model parameters such that, given the statistical model, the observed data is most probable. This is achieved by defining a parametrized probabilistic model for the data and maximizing the likelihood function with respect to the parameters.

For a statistical model, not necessarily a mixed effects model, the likelihood function is defined as the joint probability distribution of the observed data $\mathcal{D}$, but viewed and used as a function of the model parameters. Hence, the

likelihood function is defined by

$$\mathcal{L}(\boldsymbol{\theta}) = \mathcal{L}(\boldsymbol{\theta}|\mathcal{D}) \triangleq p(\mathcal{D}|\boldsymbol{\theta}). \tag{2.9}$$

In many cases, it is more convenient to work with the log-likelihood defined by

$$\ell(\boldsymbol{\theta}) = \log \mathcal{L}(\boldsymbol{\theta}). \tag{2.10}$$

For NLME models, the random effects are unobserved entities which complicates the parameter inference. To formulate the likelihood function for an individual given the individual data $\mathcal{D}_i$, the joint distribution of the data and the random effects are marginalized with respect to the random effects. To simplify the notation, we let $\boldsymbol{\theta}$ denote all model parameters of interest, including parameters in $\boldsymbol{\Sigma}$, $\boldsymbol{\Omega}$, and, in the case of stochastic dynamics, also the matrix $\boldsymbol{G}$. Hence, the likelihood for individual i can be expressed as

$$\mathcal{L}_i(\boldsymbol{\theta}|\mathcal{D}_i) = \int p(\mathcal{D}_i, \boldsymbol{\eta}_i|\boldsymbol{\theta}) d\boldsymbol{\eta}_i = \int p(\mathcal{D}_i|\boldsymbol{\theta}, \boldsymbol{\eta}_i) p(\boldsymbol{\eta}_i|\boldsymbol{\theta}) d\boldsymbol{\eta}_i. \tag{2.11}$$

In most cases, the integral over the random effects lacks a closed-form solution. To deal with the integral several approaches exist, including sampling-based methods or closed-form approximations, which will be further described in the next chapter.

The goal of maximum likelihood estimation is to find the values of the model parameters that maximize the likelihood over the parameter space $\boldsymbol{\Theta}$. The maximum likelihood estimate (MLE) is given by

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\arg\max}\, \mathcal{L}(\boldsymbol{\theta}). \tag{2.12}$$

Since the logarithm is a monotone function, maximizing the likelihood is equivalent to minimizing the negative log-likelihood. This is of statistical and computational convenience and will be used throughout the appended papers.

The MLE, $\hat{\boldsymbol{\theta}}$, is a point estimate, for which the observed data have the highest probability to occur. To assess the precision in $\hat{\boldsymbol{\theta}}$ several approaches exist, including bootstrapping (DiCiccio and Efron 1996), profile likelihood (Raue et al. 2009), and observed Fisher information (Lehman and Casella 1998). Bootstrapping and profile likelihood are computationally more expensive than the observed Fisher information, which will be used throughout the appended papers.

The observed Fisher information matrix $\boldsymbol{I}(\hat{\boldsymbol{\theta}})$, or the observed information, is given by the negative Hessian of the log-likelihood evaluated at $\hat{\boldsymbol{\theta}}$. Hence,

element $\boldsymbol{I}(\hat{\boldsymbol{\theta}})_{i,j}$ of the observed Fisher information matrix is given by

$$\boldsymbol{I}(\hat{\boldsymbol{\theta}})_{i,j} = -\frac{\partial^2}{\partial\theta_i \partial\theta_j}\ell(\boldsymbol{\theta})\bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \tag{2.13}$$

The inverse of the observed Fisher information matrix at the MLE is an estimator of the asymptotic covariance matrix of the estimated parameters. Hence, the estimated parameters are asymptotically normal distributed

$$\hat{\boldsymbol{\theta}} \rightarrow^d \mathcal{N}(\boldsymbol{\theta}_0, [\boldsymbol{I}(\boldsymbol{\theta}_0)]^{-1}), \tag{2.14}$$

where $\boldsymbol{\theta}_0$ denote the true parameter value. Here, asymptotically normal means that the distribution tends to a normal distribution when the size of the data increases. The standard errors of the MLE are calculated as the square root of the diagonal of the covariance matrix.

Although not part of this thesis, another important parameter estimation method is the Bayesian approach. Unlike maximum likelihood, where the parameters are non-random, Bayesian inference treats the parameters as random variables as well. This leads to a flexible approach, where prior information regarding the parameters easily can be incorporated using Bayes' theorem. When Bayesian approaches is combined with random effects models, this is usually referred to as Bayesian hierarchical modeling (Gelman et al. 2013).

2.6 Model Selection and Evaluation

In any regression modeling problem, it is necessary to investigate the ability of the model to describe the observed data. In addition, having one or more criteria for selecting a model among several competing models is an important step in the model building process. Here, a few of the most important methods for model selection and evaluation used throughout the appended papers are highlighted.

Model Selection

As the log-likelihood measures the goodness-of-fit of a statistical model, it also plays an important role when selecting between competing models. The likelihood-ratio test assesses the goodness-of-fit of two competing models, and

is based on the test statistic

$$\lambda_{LR} = -2\big(\ell(\boldsymbol{\theta}_0) - \ell(\hat{\boldsymbol{\theta}})\big). \tag{2.15}$$

In the test statistic above, $\ell(\boldsymbol{\theta}_0)$ is the log-likelihood under the null hypothesis $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ and $\ell(\hat{\boldsymbol{\theta}})$ is the log-likelihood under the alternative hypothesis $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Asymptotically, the test statistic is χ^2-distributed under the null hypothesis with degrees of freedom equal to the difference in dimensionality between $\boldsymbol{\Theta}$ and $\boldsymbol{\Theta}_0$ (Wilks 1938). Hence, the likelihood-ratio test can be used to investigate whether a model is significantly different from the null hypothesis. The likelihood-ratio test requires that the two models are nested, which also implies that $\Theta_0 \subseteq \Theta$.

When there are several competing models, the Akaike information criterion (AIC) can be used for model selection. The AIC is defined as

$$AIC = 2p - 2\ell(\hat{\boldsymbol{\theta}}), \tag{2.16}$$

where $\ell(\hat{\boldsymbol{\theta}})$ is the log-likelihood evaluated at the MLE and p is the number of estimated parameters (Akaike 1974). Hence, as AIC weights goodness-of-fit and model complexity, a model with lower AIC should be preferred over a model with higher AIC.

The Bayesian information criterion (BIC) is closely related to the AIC, but with a modified penalty term (Schwarz 1978). The BIC for an NLME model is given by

$$BIC = 2p \log N - 2\ell(\hat{\boldsymbol{\theta}}), \tag{2.17}$$

where N is the number of individuals, with p and $\ell(\hat{\boldsymbol{\theta}})$ defined as previously.

Model Evaluation

As previously mentioned, it is important to investigate the ability of the estimated model to describe the observed data and to determine that the underlying model assumptions are appropriate. This exercise, usually referred to as model evaluation, is a key step in the model building process.

For NLME models, graphical analysis is a popular tool as it can illustrate several complex aspects of the model's ability to describe the observed data, including description on both the population and the individual level. Commonly used goodness-of-fit plots, such as individual predictions versus observations and individual residuals versus time, can be used to guide model development and detect incorrect model assumptions. For a comprehensive overview of

graphical model evaluation techniques for NLME models, see Nguyen et al. (2017).

For simulation-based diagnostics, the visual predictive check (VPC) is one of the most popular tools in pharmacometrics. The VPC is a simulation-based graphical tool, with the purpose to assess graphically whether simulations from a model can describe the trend and variability in the observed data. This is achieved by simulating replicates of the observed data using the estimated model and comparing different percentiles of the simulated data with corresponding percentiles in the observed data. Several different types of VPCs of different complexity exist, including scatter VPC, confidence interval VPC, and prediction-corrected VPC (Bergstrand et al. 2011; Nguyen et al. 2017).

For NLME models described by ODEs, the VPC simulation incorporates inter-individual variability and measurement error. The VPC method can, as shown in Papers V and VI, be extended to NLME models described by SDEs. By simulating realizations of the stochastic dynamics, for example using the Euler-Maruyama scheme (Kloeden and Platen 1992), the stochasticity in the individual dynamics can be incorporated in the simulations. Hence, the VPC for SDEMEMs includes an additional source of variability when comparing the distribution of the model simulations with observed data.

2.7 Survival Analysis

The models considered so far describes a continuous response, modelled using either ODEs or SDEs, observed at discrete time points. Another type of model that will be encountered in Paper VI is so called *time-to-event* models, which considers the time until the occurrence of an event. In life sciences, the term survival analysis is commonly used, while the term reliability analysis is used in engineering applications. In this thesis, the term survival analysis will be used, although the event of interest might not always be a true survival time, as in the sense of 'time-to-death'. Events in clinical trials that are of interest could for example be onset of disease or withdrawal from a study (often referred to as dropout). Here, the most important concepts in survival analysis are described, and the interested reader is referred to Klein and Moeschberger (2003) and Therneau and Grambsch (2000) for further details.

A time-to-event model considers a continuous, non-negative random variable T, representing the time until the occurrence of an event. The random variable T has a probability density function $f(t)$ with cumulative distribution function $F(t) = P(T \leq t)$. One important concept in survival analysis is the survival

function, defined as the probability that the event has not yet occurred at time t. The survival function $S(t)$ is given by

$$S(t) = p(T > t) = \int_t^\infty f(x)dx = 1 - F(t). \tag{2.18}$$

Another characterization of the random variable T is given by the hazard function. The hazard function describes the instantaneous rate of occurrence of the event and is defined by

$$h(t) = \lim_{dt \to 0} \frac{p(t \leq T < t + dt \mid T \geq t)}{dt}. \tag{2.19}$$

In the equation above, the numerator describes the probability of an event occurring in the interval $[t, t + dt)$, given that no event has occurred up to time t. By dividing with the interval length dt we obtain the rate of occurrence per unit of time. Note that the hazard function in a time-to-event model is neither a probability nor a probability density but an auxiliary entity, which turns out to be very useful for modeling purposes.

The survival function $S(t)$ can be expressed in terms of the hazard function $h(t)$. Expanding the numerator in the definition of the hazard function and passing to the limit we obtain

$$h(t) = \lim_{dt \to 0} \frac{p(t \leq T < t + dt)}{dt\, p(T \geq t)} = \frac{1}{S(t)} \lim_{dt \to 0} \frac{F(t + dt) - F(t)}{dt} = \frac{f(t)}{S(t)}. \tag{2.20}$$

The relationship derived above can be interpreted in the sense that the hazard rate at time t is equal to the density of events at time t divided by the probability of surviving to time t without experiencing an event. Using the relationship $S'(t) = -f(t)$ we arrive at

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log S(t). \tag{2.21}$$

Integrating both sides from 0 to t and using the fact that no event has occurred before time zero we obtain a useful formula describing the probability of surviving to time t expressed in terms of the hazard function

$$S(t) = \exp\left(- \int_0^t h(x)dx \right). \tag{2.22}$$

The integral $\int_0^t h(x)dx$ in the exponent is often referred to as the cumulative hazard.

One of the most popular time-to-event models is the Cox proportional hazards model (Cox 1972). The Cox proportional hazards model is specified through a parametric hazard function $h(t)$ for individual i

$$h_i(t) = h_0(t) \exp(\boldsymbol{X}_i^T \boldsymbol{\beta}), \tag{2.23}$$

where $h_0(t)$ denotes the baseline hazard, $\boldsymbol{X}_i$ is a vector of individual covariates, and $\boldsymbol{\beta}$ a vector of regression coefficients.

The Cox proportional hazards model can be extended to include random effects, which turns out to be especially useful for *recurrent events*. Recurrent events are events that might occur several times within the same individual (in contrast to death, which only occur once). Examples of recurrent events in medical applications are seizures, hospitalizations, or exacerbations (a period of asthma worsening which is considered in Paper VI). For recurrent events, the dependence between the events is incorporated through a non-negative random effect. The random effect, usually referred to as the frailty, extends the standard Cox model to a 'shared frailty model' (Balan and Putter 2020). The term 'shared' refers to the fact that events within the same individual shares the same frailty. The frailty u_i acts multiplicative on the hazard according to

$$h_i(t) = h_0(t) u_i \exp(\boldsymbol{X}_i^T \boldsymbol{\beta}). \tag{2.24}$$

The frailty can be used to model inter-individual variability in the underlying risk that is not explained by the known covariates. Common distributions of the frailty include the log-normal and the Gamma distribution. Recurrent events are commonly defined in the framework of counting processes and martingale theory (Andersen and Gill 1982), for which the underlying mathematical theory is beyond the scope of this work.

Another important concept in survival analysis is *censoring*. An individual is referred to as censored when the true survival time is not known. If an individual has not experienced an event before the observation period ends, they are described as right-censored. Other types of censoring in survival analysis include left-censoring (the event occurred prior to the observation period) and interval-censoring (the event occurred during a known interval).

3 Summary of Papers

In this chapter the six appended papers are summarized and discussed. First, a short description of the six papers is given, followed by a more detailed presentation.

Paper I illustrates how the extension of an ODE model to an SDE model can be used to regularize the likelihood function for single-subject data. Sensitivity equations for the EKF is derived and used in the parameter estimation.

Paper II extends the NLME model to incorporate SDEs. The method is used both on simulated data and data from a preclinical PK experiment. A novel method for the parameter estimation problem is used, which is described in Paper III and Paper IV.

Paper III introduces the use of sensitivity equations to calculate the exact gradient needed in the parameter estimation problem for NLME models described by ODEs.

Paper IV extends the exact gradient method presented in Paper III to NLME models with dynamics governed by SDEs.

Paper V presents NLMEModeling, a Wolfram Mathematica package for NLME modeling of dynamical systems.

Paper VI develops a novel stochastic mixed effects model to analyze home-measured lung function data in asthmatic patients and investigates the association to exacerbation risk using a repeated time-to-event model.

3.1 Regularization of the Likelihood Using Stochastic Differential Equations

Estimation of model parameters in ODEs given discrete time measurement data is a complex problem. There are a number of possible difficulties, including convergence to local minima, non-identifiability, and non-differentiable terms in dynamical system models (Schittkowski 2002). Existing methods for parameter estimation in dynamical systems include least-square minimization, multiple shooting methods (Bock 1983), stochastic methods (Moles et al. 2003), and hybrid methods (Rodriguez-Fernandez et al. 2006).

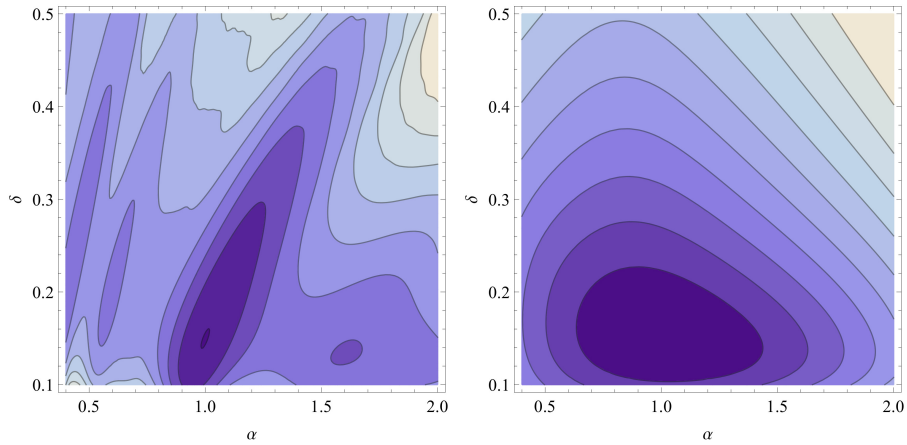
In Paper I, a novel method to regularize the likelihood function in the parameter estimation problem for ODE models is presented. We consider an extension of the dynamical model to be described by SDEs. SDEs serve as a natural way of introducing stochasticity in the dynamics, providing a more flexible model structure to account for deviations between model predictions and observations.

To estimate the model parameters from observations, the probabilistic approach is considered with the EKF used for the state estimation. To maximize the likelihood with respect to the model parameters, a gradient-based search method is used. In contrast to the commonly used finite difference approximation of the gradient, sensitivity equations for the EKF prediction and updating equations are utilized for a robust and exact gradient calculation.

Using two different models from mathematical biology, namely the FitzHugh-Nagumo model for excitable media (FitzHugh 1961; Nagumo et al. 1962) and the Lotka-Volterra predator-prey model (Volterra 1926; Lotka 1925), the impact of system noise on the likelihood function is investigated.

By considering a likelihood function depending on two model parameters, the likelihood function can easily be visualized. For the two models considered, the ODE description gives likelihood functions with several local minima. By extending the dynamical model to incorporate system noise, the state variables are attracted towards the observed data and the number of local minima is shown to be reduced, as depicted in Figure 3.1. Similar results have recently been shown in a discrete time setting, using a multiple shooting approach (Ribeiro et al. 2020).

For future work, we suggest an extended analysis to high dimensional problems that are known to exhibit problems with local minima. Another potential extension of the method presented in Paper I could be to use the system noise



(a) ODE **(b)** SDE

Figure 3.1: Illustration of the regularization using SDEs in the Lotka-Volterra predator-prey model. The two plots show the level curves of the likelihood function in the parameter estimation problem using the ODE setting (a) and the SDE setting (b). In (a) there are several local minima to where a local optimization algorithm might converge.

as a "turning knob" during the optimization of model parameters. Once the optimization algorithm is close to the minima, the magnitude of the system noise could be decreased which eventually would make the SDE model correspond to the ODE model (with system noise set to zero).

3.2 Nonlinear Mixed Effects with Stochastic Dynamics

As previously discussed, the combination of stochastic differential equations and mixed effects considers three sources of variability in the observations: inter-individual variability, system noise, and measurement error.

Paper II explores the use of SDEMEMs for PK applications. Two different PK modeling examples are presented. First, a simulation-estimation exercise is conducted with the aim of investigating whether the model parameters and the three sources of variability can reliably be estimated from observed data. The impact of assuming a deterministic model and neglecting the system noise is also investigated. Secondly, a previously published PK model of nicotinic acid (NiAc) in obese Zucker rats is extended to incorporate SDEs (Ahlström

et al. 2013).

The parameter estimation method used in Paper II considers the combination of the first-order conditional estimation (FOCE) method (Beal et al. 2017; Wang 2007) and the EKF. The FOCE method is used for approximating the intractable likelihood, while the EKF is used for state estimation in the stochastic dynamical model. The combination of FOCE and EKF was introduced by Henrik Madsen and colleagues, see for instance Klim et al. (2009), Mortensen et al. (2007), Overgaard et al. (2005), and Tornøe et al. (2005), which in this work has been further extended.

Paper II utilize a novel method for calculating the gradient in the parameter estimation problem. The method is based on an exact derivation of the gradient, calculated using sensitivity equations. The exact gradient method for NLME models governed by ODEs is presented in Paper III, and in Paper IV the corresponding algorithm for NLME models with SDEs is presented.

The results from the simulation model in Paper II show that the model parameters as well as the three sources of variability can be reliably estimated from the data. If the system noise is neglected, the estimated residual error is significantly higher than the value used for simulations. Similar results have recently been reported by others (Wiqvist et al. 2021). The results for the NiAc disposition model show that the error previously described as pure measurement error can be divided into a reduced measurement error and a significant system noise term. The comparison of the approaches is illustrated in Figure 3.2, where the estimated concentration of NiAc for six rats are shown for the ODE case (left panels) and SDE case (right panels).

The significant system noise implies that the deterministic structure of the NiAc disposition model potentially could be improved, with the aim of reducing the uncertainty in the underlying dynamics. This was not in scope of the current investigation but may be a suitable problem for future work.

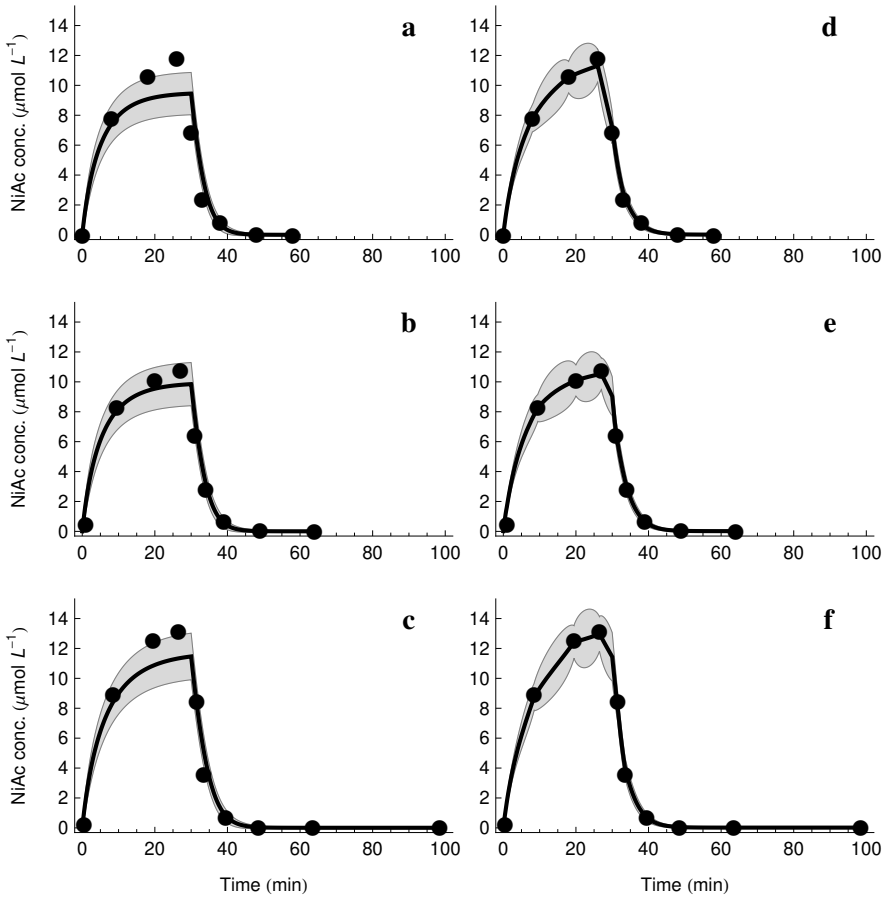


Figure 3.2: ODE modeling (left panels) compared to SDE modeling (right panels) for the NiAc PK model in Paper II.

3.3 Exact Gradients for Nonlinear Mixed Effects Models

One of the challenges developing NLME models is the estimation of the model parameters from observed data, and several computer programs exist for this purpose (Beal et al. 2017; Certara 2020; Fidler et al. 2019; Lixoft SAS 2020). For parameter estimation based on the likelihood approach, the FOCE approximation and stochastic approximation expectation maximization (SAEM) are two popular methods (Beal et al. 2017; Wang 2007; Kuhn and Lavielle 2005).

Both the FOCE and SAEM methods approximate the likelihood function, and the difference lies in how the approximation is performed. While the FOCE approximates the integral with a closed-form expression, SAEM instead aims to maximize the integral using a sampling-based approach, which (at least in theory) can be made to be arbitrarily precise. However, for many PK-PD problems, FOCE is still one of the most popular methods used. The FOCE approximation and the calculation of the gradient are illustrated in Figure 3.3, and briefly explained as follows.

We consider a set of experimental observations $\boldsymbol{d}_{ij}$, where the index notation ij is used to denote the jth observation for the ith individual, with $i = 1, \ldots, N$ and $j = 1, \ldots, n_i$. Here, we make a distinction between the actual observation $\boldsymbol{d}_{ij}$, the model for the observation $\boldsymbol{y}_{ij}$ (a random variable), and the model prediction $\hat{\boldsymbol{y}}_{ij}$. The residuals are given by the deviation from the model prediction

$$\boldsymbol{\epsilon}_{ij} = \boldsymbol{d}_{ij} - \hat{\boldsymbol{y}}_{ij}, \tag{3.1}$$

where

$$\hat{\boldsymbol{y}}_{ij} = E[\boldsymbol{y}_{ij}|\mathcal{D}_{i(j-1)}] \tag{3.2}$$

is the predicted model output, conditioned on $\mathcal{D}_{i(j-1)} = \{\boldsymbol{d}_{ij}, \boldsymbol{d}_{i2}, \ldots, \boldsymbol{d}_{i(j-1)}\}$ which is the information available up to time $t_{i(j-1)}$, with corresponding conditional covariance matrix

$$\boldsymbol{R}_{ij} = Cov[\boldsymbol{y}_{ij}|\mathcal{D}_{i(j-1)}]. \tag{3.3}$$

Furthermore, we let $\mathcal{D}_i$ denote the collection of all data for individual i and $\mathcal{D} = \{\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_N\}$ denote the complete set of observations.

Assuming independence between individuals, i.e., the individual data and random effects $(\mathcal{D}_i, \boldsymbol{\eta}_i), i = 1, \ldots, N$ are assumed to be independent between individuals, the population likelihood for all data can be written as a product over individual likelihoods. Since the random effects are unobserved, the individual joint likelihoods (Figure 3.3A) are marginalized over the random effects, to obtain the contribution to the population likelihood (Figure 3.3B). Hence, the expression for the population likelihood can be written as

$$\mathcal{L}(\boldsymbol{\theta}|\mathcal{D}) = \prod_{i=1}^{N} \int p(\mathcal{D}_i|\boldsymbol{\theta}, \boldsymbol{\eta}_i) p(\boldsymbol{\eta}_i|\boldsymbol{\theta}) d\boldsymbol{\eta}_i = \prod_{i=1}^{N} \int \exp(l_i) d\boldsymbol{\eta}_i. \tag{3.4}$$

In the expression above, $l_i = l_i(\boldsymbol{\eta}_i)$ is the individual joint log-likelihood. Since both the observations and random effects are assumed to be normal distributed,

the expression for l_i is given by

$$l_i = -\frac{1}{2}\sum_{j=1}^{n_i}\left(\boldsymbol{\epsilon}_{ij}^T\boldsymbol{R}_{ij}^{-1}\boldsymbol{\epsilon}_{ij} + \log\det\left(2\pi\boldsymbol{R}_{ij}\right)\right) - \frac{1}{2}\boldsymbol{\eta}_i^T\boldsymbol{\Omega}^{-1}\boldsymbol{\eta}_i - \frac{1}{2}\log\det\left(2\pi\boldsymbol{\Omega}\right). \tag{3.5}$$

The integral with respect to $\boldsymbol{\eta}_i$ does not have a closed-form solution, except in trivial cases. To approximate the integral, a second-order Taylor expansion of l_i around the point $\boldsymbol{\eta}_i^*$ which maximizes l_i is used, often referred to as the Laplacian approximation (Vonesh 1996). The first step of the FOCE method is to find $\boldsymbol{\eta}_i^*$ for each individual (Figure 3.3C). The second-order Taylor expansion of l_i at $\boldsymbol{\eta}_i^*$ (Figure 3.3D) gives a closed-form expression of the marginalization, which yields the following approximate expression for the population log-likelihood

$$\hat{\ell}(\boldsymbol{\theta}) = \log\hat{\mathcal{L}}(\boldsymbol{\theta}) = \sum_{i=1}^{N}\left(l_i(\boldsymbol{\eta}_i^*) - \frac{1}{2}\log\det\left[\frac{-\boldsymbol{H}_i(\boldsymbol{\eta}_i^*)}{2\pi}\right]\right). \tag{3.6}$$

In the expression above, $\boldsymbol{H}_i$ is the Hessian of l_i, evaluated at $\boldsymbol{\eta}_i^*$. Depending on the number of terms kept in an approximated expression of $\boldsymbol{H}_i$, the Laplacian method, FOCE, or the FOCE with interaction (FOCEI) are obtained. Further details can be found in Paper III.

Due to the nature of the conditional estimation methods, the maximization of $\hat{\ell}(\boldsymbol{\theta})$ is a nested optimization problem and a computationally expensive task. For each evaluation of the population log-likelihood $\hat{\ell}(\boldsymbol{\theta})$, N individual optimizations of l_i are required to find the points $\boldsymbol{\eta}_i^*, i = 1, \ldots, N$. We will refer to the optimization of the population log-likelihood $\hat{\ell}(\boldsymbol{\theta})$ with respect to the model parameters as the outer optimization problem and the optimization of individual joint log-likelihoods l_i with respect to the random effects parameters as the inner optimization problem.

The maximizations of $\hat{\ell}(\boldsymbol{\theta})$ and l_i (in fact, the minimizations of $-\hat{\ell}(\boldsymbol{\theta})$ and $-l_i$) are typically solved using a gradient-based optimization method, such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method (Nocedal and Wright 2006). The BFGS algorithm is an iterative method for unconstrained nonlinear optimization problems and approaches a (local) minimum by combining information regarding the descent direction with curvature information. Both the descent direction and the curvature information are obtained from the gradient. To obtain the gradients needed in the inner and outer problem, as well as the first-order derivatives in the approximation of $\boldsymbol{H}_i$, a finite difference approach

is usually adopted (Figure 3.3E). In this work, we instead consider the use of sensitivity equations to obtain the necessary expressions.

Paper III concerns the derivation of the exact gradient method for NLME models described by ODEs. The exact gradient method is compared with a finite difference approach, using a two-compartmental PK model of increasing complexity. The performance of the method is investigated in terms of accuracy and precision in the gradient calculation, as well as the computational time.

In Paper IV, the exact gradient method presented in Paper III is extended to NLME models with dynamics described using SDEs. Expanding on the combination of FOCE and EKF, the exact gradient method for SDEMEMs is derived. The derivation builds on the results from Paper III, with the additional requirement of differentiating the prediction and updating equations of the EKF. This was partly done in Paper I, where the first-order sensitivities for the EKF were derived. The method is evaluated on simulated data from three common PK and PK-PD models extended to stochastic models, describing uncertainty in the absorption kinetics. As in Paper III, the performance of the exact gradient method is compared to a finite difference approach.

The main result from Paper III and IV is the derivation of how the exact gradients can be computed for the FOCE and FOCEI approximations of the likelihood for NLME models. The derivation requires up to second-order sensitivity equations of the underlying system equations, which is used to form the needed gradient expressions. One key result in the derivation is the computation of the matrix $d\boldsymbol{\eta}_i^*/d\boldsymbol{\theta}$, describing how the point used in the Taylor expansion depends on the model parameters (Figure 3.3F).

The exact gradient method is shown to have several advantages over the finite difference approach. First, the exact gradient method is shown to be faster than using finite differences. The speed-up varies depending on the model complexity and the number of model parameters. Second, the exact gradient method is shown to have an improved precision and accuracy in the gradient calculation. The reason for this is most likely that in the exact gradient approach, the numerical precision only depends on the precision used to solve the underlying system of ODEs. In that sense, the exact gradient method is also more convenient from a user perspective since the modeler does not have to consider defining a step-length used in a finite difference approach (one value for each dimension of the parameter space).

Expanding on the work presented in Paper III and IV, there are a few possible extensions. In contrast to the FOCE approximation, which only considers first-order terms in the approximation of the Hessian matrix $\boldsymbol{H}_i$, the Laplacian approximation includes second-order terms. These could, at least in theory,

be retrieved using sensitivity equations. However, second-order terms in the Hessian of l_i would require up to third-order sensitivity equations to obtain the gradient, which might be computationally demanding to solve. It is also important to note that in Paper III and Paper IV, we only show that the exact gradient method improves speed and numerical properties of the gradient. It remains to be shown that the improved numerical properties lead to an increased robustness of the optimization problem. This has to some extent been evaluated in a master thesis project, but a more thorough investigation is warranted (Ólafsdóttir 2016).

The exact gradient method presented in Paper III has been adopted and is available in version 7.4 and higher of the computer program NONMEM (option FAST), one of the most popular programs for NLME model estimation in the pharmaceutical industry. In NONMEM, the exact gradient method was shown to improve the computational speed for both the estimation and the covariance step for ODE-based models (Beal et al. 2017).

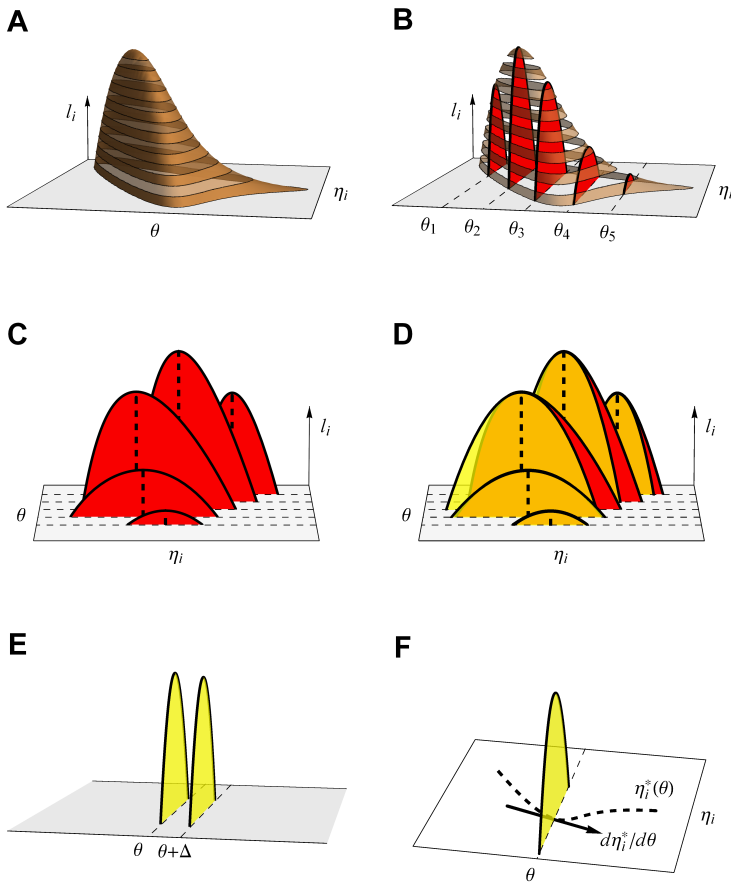


Figure 3.3: Approximation of population likelihood and calculation of the gradient. In **(A)** an individual joint log-likelihood l_i is illustrated, depending on a fixed effect parameter θ and a random effect parameter η_i. In **(B)** the integration over the random effect is illustrated (strictly speaking the marginalization is done on the individual likelihoods), which calculates the contribution of each l_i to the population (log)-likelihood. **(C)** shows the first step of the FOCE method, which is to find the value of η_i which maximizes l_i. **(D)** illustrate the second step of the FOCE approximation, where a second-order Taylor expansion is done at η_i^* to obtain an approximate expression for the contribution to the population (log)-likelihood. **(E)** and **(F)** shows the concept of calculating the gradient with either a finite difference approach **(E)** or an exact gradient approach **(F)**. Figure re-used with permission from the originator (Almquist 2017).

3.4 NLMEModeling: A Wolfram Mathematica Package

The methods and algorithms presented in this thesis have been implemented in Wolfram Mathematica, a programming platform well suited for modeling and scientific computation (Wolfram Research, Inc. 2020). Wolfram Mathematica was primarily used due to its powerful and user-friendly combination of symbolic computation, matrix operations, and efficient algorithms. The methods developed were later refined, and together with user-friendly functionality deployed into NLMEModeling, a Wolfram Mathematica package for NLME modeling of dynamical systems. The package have previously been used in several NLME modeling applications, including oncology (Cardilin et al. 2017; Cardilin et al. 2018; Cardilin et al. 2019), single-cell experiments (Almquist et al. 2015a), and PK-PD modeling (Andersson et al. 2016; Andersson et al. 2017; Andersson et al. 2019; Tapani et al. 2014).

Paper V serves as a tutorial on NLMEModeling. The package is relevant for both current users of Wolfram Mathematica that want the ability to perform NLME modeling, but also to modelers who seek a convenient and streamlined modeling environment. The current version of the package supports NLME models where the dynamical model is defined using either ODEs or SDEs, together with a flexible observation model.

The package is demonstrated using three PK-PD modeling examples. The modeling workflow, including model definition, parameter estimation, and model evaluation are illustrated. A minimal example is depicted in Figure 3.4, which shows how an NLME model can be defined and estimated using only a few lines of code.

```
(*Define dynamic model*)
sys = {
   A1'[t] == -ka * A1[t],
   A2'[t] == ka * A1[t] - CL * Exp[η1] / V * A2[t],
   A1[0] == Dose,
   A2[0] == 0,
   c[t] == A2[t] / V};
(*Define observation model*)
obs = {c[t]};
(*Estimate the NLME model, using a proportional error model (and diagonal Ω matrix)*)
modelObject = NLMEDynamicalModelFit[data, {sys, obs}, {{ka, 0.9}, {CL, 5}, {V, 30}}, η1, Sigma → "Proportional"];
(*Perform goodness-of-fit analysis*)
GoodnessOfFitAnalysis[modelObject];
(*Do a visual predictive check, based on 200 simulated datasets*)
VisualPredictiveCheck[modelObject, 200, Quantiles → {0.1, 0.5, 0.9}, ConfidenceInterval → 90];
```

Figure 3.4: Illustration of the syntax used to define, estimate, and perform additional model diagnostics.

The estimation functionality returns a model object, which further can be used for model evaluation and additional analyses. Here, the model object is passed into a goodness-of-fit analysis and a VPC, resulting in the output depicted in Figure 3.5.

NLMEModeling is a freely available package, providing the modeler with an integrated Wolfram Mathematica environment for NLME modeling. In addition, the user of the package can develop additional functionality tailored to their own needs.

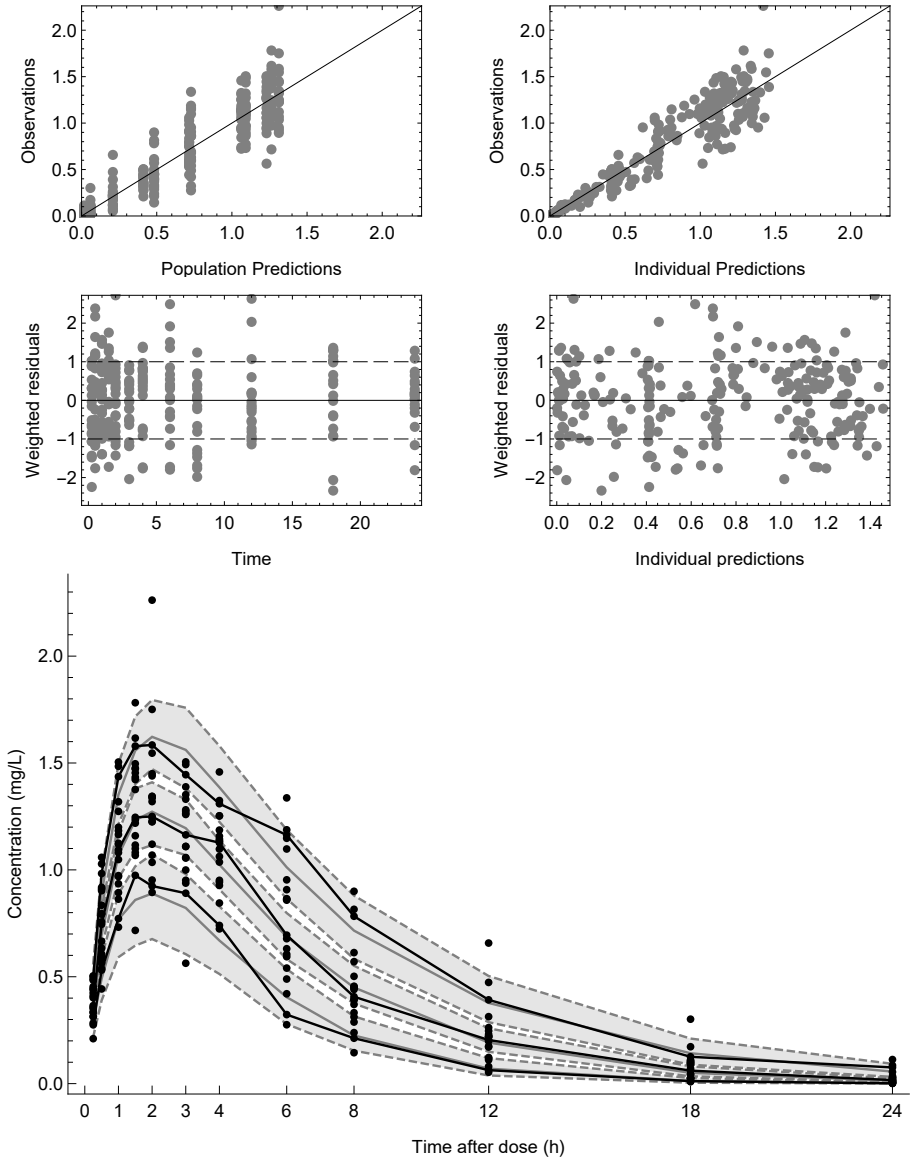


Figure 3.5: Example of the default goodness-of-fit analysis and a VPC plot using the NLMEModeling package.

3.5 Stochastic Mixed Effects Modeling of Peak Expiratory Flow

In clinical trials, home-based measurements can allow for frequent monitoring of study participants. In respiratory clinical trials, a patient's lung function can be captured at home using a hand-held device. An example of such a measurement is peak expiratory flow (PEF), which measures the maximal flow of air during exhalation. In contrast to previously considered applications where the collection of data is sparse, home-measured variables is collected frequently (sometimes several times a day). With that level of detail in the observed data, several interesting aspects of the PEF dynamics can be studied.

Another important variable in respiratory clinical trials is exacerbations, which currently is the registrational endpoint in Phase III trials. An exacerbation is a period of disease worsening, that requires additional treatment and/or hospital admission.

Paper VI describes the development of a novel longitudinal mixed effects model to describe home-measured PEF. The combination of NLME modeling and stochastic differential equations is used, which allows for quantification of several properties related to the PEF dynamics: the longitudinal trend, long-term fluctuations, and day-to-day variability. The concept is depicted in Figure 3.6, which illustrates the decomposition of the different components of the PEF time series.

The three components are combined to describe a patient's PEF observation according to

$$PEF(t_j) = x(t_j) + v(t_j) + \sigma e(t_j), \tag{3.7}$$

where $x(t_j)$, $v(t_j)$, and $\sigma e(t_j)$ are used to describe the treatment response, the long-term fluctuations, and the day-to-day variability, respectively. The treatment response is modelled by an indirect response model governed by the following ODE

$$\frac{dx(t)}{dt} = k_{tr}\big(PEF_{base}(1+eff) - x(t)\big), \quad x(0) = PEF_{base}, \tag{3.8}$$

where k_{tr} describes the rate of onset of treatment effect, PEF_{base} is the baseline PEF level, and eff is an asymptotic treatment effect. The long-term fluctuations are modeled using an SDE governed by

$$dv(t) = -k_v\, v(t)\, dt + g\, dW(t), \quad v(0) = 0, \tag{3.9}$$

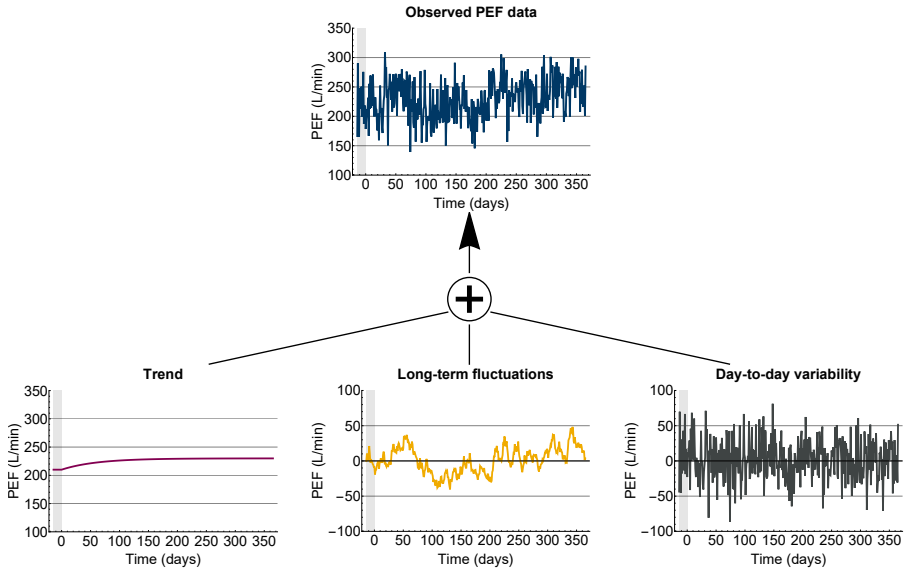


Figure 3.6: Conceptual illustration of the modeling approach used in Paper VI. The time series is decomposed into three components describing different characteristics of the PEF response: trend, long-term fluctuations, and day-to-day variability.

commonly known as the Ornstein-Uhlenbeck model. In addition to the two dynamical components, an additive observation noise, $\sigma e(t_j)$, is used to capture the day-to-day variability.

The proposed modeling approach is used to analyze PEF data from two previously reported clinical trials in asthma (Bleecker et al. 2016; FitzGerald et al. 2016). Using the mixed effects approach, individual parameters describing baseline PEF, asymptotic treatment effect, and the size of the long-term fluctuations and the day-to-day variability are estimated. The individual model parameters together with known covariates are used in a repeated time-to-event (RTTE) analysis to investigate the association to exacerbation risk. The hazard function for individual i is described by

$$h_i(t) = h_0(t)u_i \exp(\boldsymbol{X}_i^T \boldsymbol{\beta}), \tag{3.10}$$

where $h_0(t)$ denotes the baseline hazard, u_i is a gamma-distributed frailty, $\boldsymbol{X}_i$ denotes the individual covariates, and $\boldsymbol{\beta}$ are the corresponding regression coefficients. Covariates considered in the RTTE analysis were the factors study, treatment group, age, and sex, as well as the individual model parameters base-

line PEF (PEF_{base}), asymptotic treatment effect (eff), day-to-day variability (σ), and long-term fluctuations (g).

The model enables characterization of multiple statistical properties of PEF time series data to support better estimation and understanding of treatment effects, disease stability, and exacerbation risk. Interestingly, several of the model parameters are shown to be associated to the risk of experiencing an exacerbation. The long-term fluctuations show a significant positive association with exacerbation risk, implying that patients that have a higher degree of fluctuations have a higher risk of experiencing an exacerbation.

The stochastic PEF model presented in Paper VI could be used as an informative way of analyzing the home-measured PEF data. This might lead to better designed clinical trials, as well as serve as a tool for finding patients that has a high risk of experiencing an exacerbation. The modeling approach is generalizable and could also be extended to other types of measurements or disease areas.

4 Discussion and Conclusions

The research presented in this doctoral thesis aims to contribute to the development of methods for mathematical modeling of dynamical systems, with applications in drug development. The research is presented as six appended papers and in the previous sections of this thesis.

4.1 Main Contributions and Research Aims

The three research aims presented in the introduction cover the investigation of stochastic models (A1), the development of parameter estimation methods (A2), and the application of mathematical modeling in drug development (A3).

In Paper I, the problem of parameter estimation in dynamical system given discrete time observations was considered. The impact of going from an ODE model to an SDE description was investigated as a tool to overcome the problem of local minima in the likelihood function for single-subject data. In Paper II, the combination of SDEs and NLME modeling was used and applied to both simulated and experimental PK data. In this paper, a novel method for parameter estimation in NLME models was developed, combining the FOCE approximation with exact gradients using sensitivity equations. The details of the exact gradient method is presented in Paper III and Paper IV. Furthermore, the methods developed have been integrated into NLMEModeling, a Wolfram Mathematica package for NLME modeling of dynamical systems, which is presented in Paper V. Finally, the methodology was used in an applied setting in Paper VI, where an NLME model governed by SDEs was used to describe home-measured PEF data and to investigate the association to exacerbation risk in asthma patients.

The appended papers contribute to the aims as follows. The investigation of stochastic models (A1) has been addressed in the majority of the papers, except

for Paper III. The development of parameter estimation methods (A2) was the focus of Paper I, Paper III and Paper IV. The application of mathematical modeling in drug development (A3) was addressed partly in Paper II and even more extensively in Paper VI. Paper V is a result of the methods developed in close collaboration with co-authors throughout this doctoral thesis, and connects to all the aims.

4.2 The Modeling Process

The application of mathematical modeling presented in Paper VI is a clear example of how the modeling process, depicted in Figure 1.1, can be used to provide insight into a real-world problem. The system of interest, in this case the PEF response over time, was translated to a mathematical model with the purpose of describing home-based PEF measurements to better understand the characteristics of the PEF dynamics and the association to exacerbation risk. The model was able to describe the most important characteristics of the underlying system, such as the trend and fluctuations, which then was used to provide answers to the questions at hand. The modeling results showed that several factors could be contributed to exacerbation risk, including the stochastic behavior of the dynamical process. This knowledge can potentially be used in the future to design better clinical trials and to improve the understanding of the patient population.

An important part of the modeling process is the estimation of model parameters from experimental data. For NLME models, this is a computationally demanding task. In terms of parameter estimation methods, two novel methods have been presented as part of this thesis; the regularization method using SDEs for single-subject data and the exact gradient method for NLME models described by either ODEs or SDEs. The exact gradient method presented in Paper III and IV can replace the use of finite differences and is a significant improvement of the established FOCE method. The exact gradient method for NLME models with ODEs has been adopted by the developers of the NONMEM software and is available from version 7.4, providing faster parameter estimation compared to a finite difference approach. Having access to efficient and robust algorithms is important, as it enables modelers to spend more time on scientific questions and linking back the model results to the actual problem.

In pharmacometrics, the models have traditionally been described using ODEs. It is, however, important to note that all models are simplifications of the real world. As discussed in the introduction chapter, some aspects of the

system that are deemed not relevant for the question might be removed in the modeling process. To account for the fact that not only observations are uncertain (as in the standard NLME model), uncertainty in the underlying dynamics can be considered. The extension to stochastic models leads to a more general class of models compared to ODEs, with the downside of being more computationally demanding. The reward, however, is the ability to separate and quantify different types of uncertainty. As mathematical models are an important part of the decision-making process in modern drug development, accurate quantification of the uncertainty in model predictions are particularly important.

4.3 Future Work and Open Problems

The research conducted during this PhD project poses several new questions. Here, a few potential extensions and important remarks are discussed, which could be of interest for future investigations.

Exact Gradient Calculation

The computation of exact gradients is an important contribution to the parameter estimation problem. In this work, the computation was done using sensitivity equations, obtained by differentiating the original system of ODEs with respect to the model parameters, which is sometimes denoted forward sensitivities (Serban and Hindmarsh 2005). For parameter estimation in NLME models using the FOCE method, the sensitivity approach requires up to second-order sensitivity equations. As the number of sensitivity equations required grows with model complexity the size of the extended ODE system can become large.

An alternative strategy for computing the gradient is to consider so called adjoint sensitivities, where the adjoint-state method can be used to find the gradient of a functional of the solution to the ODEs, such as a sum of squares or an individual likelihood (Serban and Hindmarsh 2005). The adjoint-state method is an intriguing approach, as it reduces the number of equations that needs to be solved to obtain the gradient. In contrast to the sensitivity approach used in this work which requires $n\,(1+p)$ equations, with n being the number of state variables and p the number of parameters, the adjoint-state method only requires $2\,n$ equations. The adjoint sensitivity method has successfully been applied to parameter estimation in biochemical reaction networks modeling,

where it was shown to improve computational efficiency (Fröhlich et al. 2017). Currently, the use of adjoint sensitivities in the parameter estimation for NLME models are limited. Depending on the complexity of the functional (e.g., outer or inner optimization problem in the FOCE approximation), the application of adjoint sensitivities might be more or less complex. For the inner optimization problem (to find the point used in the Laplace approximation), the use of adjoint sensitivities should be straightforward. For the outer optimization problem (the estimation of the NLME model parameters), on the other hand, the FOCE algorithm requires second-order sensitivities, for which a potential strategy could be to consider higher order adjoint sensitivities (Stapor et al. 2018).

Identifiability Analysis

Another important topic closely related to the problem of parameter estimation is identifiability analysis, which aims to assess whether the parameters in a model can be inferred from experimental data. Identifiability analysis is further divided into structural and practical identifiability. Structural identifiability is related to the model structure independent of experimental data (Bellman and Åström 1970), while practical identifiability considers the amount and quality of the observed data used for parameter estimation (Raue et al. 2009). Although identifiability analysis has not been the topic of this research project, it is an important part of the mathematical modeling process.

For single-subject data, several approaches exist to assess structural identifiability including, e.g., Exact Arithmetic Rank, input-output approaches, and profile likelihood (Karlsson et al. 2012; Bearup et al. 2013; Raue et al. 2009; Raue et al. 2014). For PK-PD applications, the identifiability of fundamental PD models has been investigated (Janzén et al. 2016). For dynamical models described by SDEs, identifiability analysis has been applied for a range of system biology models (Browning et al. 2020).

In recent years, the concept of identifiability in mixed effects models has been an area of increasing research interest (Lavielle and Aarons 2016). Several methods has been extended to the mixed effects case, including the Taylor series expansion approach and the input-output approach, with special focus on PK-PD applications (Janzén et al. 2018). In addition to assessing structural identifiability of the fixed effects parameters, the methods enable assessment of structural identifiability of the random effects covariance matrix.

When mixed effects models with dynamics governed by SDEs are considered, an additional source of variability has to be identified. As the applications of

SDEs and mixed effects models increase, identifiability analysis for SDEMEMs needs to be further investigated. Ultimately, assessment of identifiability (both structural and practical) should be an important step of the modeling process.

Inferring the Underlying State

Throughout this work, the EKF has been used to infer the underlying state of the system in dynamical models described by SDEs. For nonlinear systems, the EKF uses a first-order linearization around the state trajectory to approximate the probability densities for the state estimates conditioned on available data using Gaussian distributions. However, for highly nonlinear models, the linearization of the system dynamics around the state trajectory might not be an adequate approximation. To allow for a more general representation of the state estimates conditioned on available data other types of estimators can be considered. One class of such methods is sequential Monte Carlo methods, also known as particle filters (Schön et al. 2018). Applications of particle filters within the field of PK-PD modeling have been considered (Krengel et al. 2013), and recent developments include the use of particle methods for SDEMEMs (Botha et al. 2020; Wiqvist et al. 2021).

Frequently Sampled Observations

Traditionally, PK-PD models utilize sparsely sampled data to infer the underlying model parameters. As digital solutions are becoming more common in clinical trials, including home-based sampling and sensor devices, the resolution of the emerging clinical data is constantly improving (James et al. 2020). To deal with high-frequent data, as in Paper VI, SDEs are an intriguing approach as they enable description of stochastic behaviours not captured by a deterministic model. If the sampling frequency is high enough (in relation to the time-scale of interest), the transition to so called discrete-time dynamical models models could be possible (Ljung 1999). The discrete-time framework is commonly used in technical applications, such as automotive and manufacturing processes, and could significantly reduce the computational cost since no integration of the ODE system is needed.

Joint Inference of Longitudinal and Event Data

Although only considered in the Paper VI, time-to-event data are important in many clinical trials. In Paper VI, the longitudinal model and the time-to-event

model were estimated separately. Recent developments on this topic is the concept of joint modeling, where a joint probability distribution of the longitudinal data and the event data are considered (Król et al. 2017; Rizopoulos 2016). Extensions of the SDEMEM framework to include time-to-event data is an intriguing approach, as it would allow for simultaneous estimation of the longitudinal model and the time-to-event model. This could potentially lead to an improved understanding of the association between the longitudinal response and the time-to-event process. Joint modeling has also been successfully applied to account for study dropout in the analysis of exacerbation risk (Król et al. 2020). The impact of study dropout was not considered in the current investigation, but is an interesting extension for future work.

A Last Remark

To conclude, the research presented in this doctoral thesis has contributed to the development of methods and applications of mathematical modeling of dynamical systems. Although the methods primarily have been applied to problems within the field of drug development, they are most likely applicable in many other scientific fields.

Bibliography

Ahlström C, Kroon T, Peletier LA, and Gabrielsson J (2013). Feedback modeling of non-esterified fatty acids in obese Zucker rats after nicotinic acid infusions. J. Pharmacokinet. Pharmacodyn. 40 (6): 623–638. DOI: 10.1007/s10928-013-9335-z.

Akaike H (1974). A New Look at the Statistical Model Identification. IEEE Trans. Automat. Contr. 19 (6): 716–723. DOI: 10.1109/TAC.1974.1100705.

Almquist J (2017). Kinetic Models in Life Science - Contributions to Methods and Applications. PhD thesis. Chalmers University of Technology.

Almquist J, Bendrioua L, Adiels CB, Goksör M, Hohmann S, and Jirstrand M (2015a). A nonlinear mixed effects approach for modeling the cell-to-cell variability of Mig1 dynamics in yeast. PLoS One 10 (4): 1–32. DOI: 10.1371/journal.pone.0124050.

Almquist J, Leander J, and Jirstrand M (2015b). Using sensitivity equations for computing gradients of the FOCE and FOCEI approximations to the population likelihood. J. Pharmacokinet. Pharmacodyn. 42 (3): 191–209. DOI: 10.1007/s10928-015-9409-1.

Almquist J, Sadiq MW, Eriksson UG, Hegelund Myrbäck T, Prothon S, and Leander J (2020). Estimation of equipotent doses for anti-inflammatory effects of prednisolone and AZD9567, an oral selective non-steroidal glucocorticoid receptor modulator. CPT Pharmacometrics Syst. Pharmacol. 9 (8): 444–455. DOI: 10.1002/psp4.12536.

Andersen PK and Gill RD (1982). Cox's Regression Model for Counting Processes: A Large Sample Study. Ann. Stat. 10 (4): 1100–1120. DOI: 10.1214/aos/1176345976.

Andersson R, Jirstrand M, Almquist J, and Gabrielsson J (2019). Challenging the dose-response-time data approach: Analysis of a complex system. Eur. J. Pharm. Sci. 128: 250–269. DOI: 10.1016/j.ejps.2018.11.015.

Andersson R, Jirstrand M, Peletier L, Chappell MJ, Evans ND, and Gabrielsson J (2016). Dose-response-time modelling: Second-generation turnover model

with integral feedback control. Eur. J. Pharm. Sci. 81: 189–200. DOI: 10.1016/j.ejps.2015.10.018.

Andersson R, Kroon T, Almquist J, Jirstrand M, Oakes ND, Evans ND, Chappel MJ, and Gabrielsson J (2017). Modeling of free fatty acid dynamics: insulin and nicotinic acid resistance under acute and chronic treatments. J. Pharmacokinet. Pharmacodyn. 44 (3): 203–222. DOI: 10.1007/s10928-017-9512-6.

Åström KJ (1970). Introduction to Stochastic Control Theory. Vol. 70. Mathematics in Science and Engineering. Academic Press.

Balan TA and Putter H (2020). A tutorial on frailty models. Stat. Methods Med. Res. 29 (11): 3424–3454. DOI: 10.1177/0962280220921889.

Barrett JS, Fossler MJ, Cadieu KD, and Gastonguay MR (2008). Pharmacometrics: A multidisciplinary field to facilitate critical thinking in drug development and translational research settings. J. Clin. Pharmacol. 48 (5): 632–649. DOI: 10.1177/0091270008315318.

Beal SL, Sheiner LB, Boeckmann A, and Bauer R (2017). NONMEM 7.4 User's Guides (1989-2017).

Bearup DJ, Evans ND, and Chappell MJ (2013). The input-output relationship approach to structural identifiability analysis. Comput. Methods Programs Biomed. 109 (2): 171–181. DOI: 10.1016/j.cmpb.2012.10.012.

Bellman R and Åström KJ (1970). On structural identifiability. Math. Biosci. 7 (3-4): 329–339. DOI: 10.1016/0025-5564(70)90132-X.

Berglund M, Sunnåker M, Adiels M, Jirstrand M, and Wennberg B (2012). Investigations of a compartmental model for leucine kinetics using non-linear mixed effects models with ordinary and stochastic differential equations. Math. Med. Biol. 29 (4): 361–384. DOI: 10.1093/imammb/dqr021.

Bergstrand M, Hooker AC, Wallin JE, and Karlsson MO (2011). Prediction-corrected visual predictive checks for diagnosing nonlinear mixed-effects models. AAPS J. 13 (2): 143–151. DOI: 10.1208/s12248-011-9255-z.

Bilgel M, Prince JL, Wong DF, Resnick SM, and Jedynak BM (2016). A multivariate nonlinear mixed effects model for longitudinal image analysis: Application to amyloid imaging. Neuroimage 134: 658–670. DOI: 10.1016/j.neuroimage.2016.04.001.

Black F and Scholes M (1973). The Pricing of Options and Corporate Liabilities. J. Polit. Econ. 81 (3): 637–654.

Bleecker ER, FitzGerald JM, Chanez P, Papi A, Weinstein SF, Barker P, Sproule S, Gilmartin G, Aurivillius M, Werkström V, and Goldman M (2016). Efficacy and safety of benralizumab for patients with severe asthma uncontrolled with high-dosage inhaled corticosteroids and long-acting β2-agonists (SIROCCO): a randomised, multicentre, placebo-controlled phase 3 trial. Lancet 388 (10056): 2115–2127. DOI: 10.1016/S0140-6736(16)31324-1.

Bock HG (1983). Recent Advances in Parameteridentification Techniques for O.D.E. *Numerical treatment of inverse problems in differential and integral equations*. Ed. by P Deuflhard and E Hairer. Boston, MA: Birkhäuser Boston: 95–121. DOI: 10.1007/978-1-4684-7324-7_7.

Botha I, Kohn R, and Drovandi C (2020). Particle Methods for Stochastic Differential Equation Mixed Effects Models. Bayesian Anal. 1–35. DOI: 10.1214/20-BA1216.

Browning AP, Warne DJ, Burrage K, Baker RE, and Simpson MJ (2020). Identifiability analysis for stochastic differential equation models in systems biology: Identifiability analysis for stochastic differential equation models in systems biology. J. R. Soc. Interface 17 (173): 37–44. DOI: 10.1098/rsif.2020.0652rsif20200652.

Cardilin T, Almquist J, Jirstrand M, Sostelly A, Amendt C, El Bawab S, and Gabrielsson J (2017). Tumor Static Concentration Curves in Combination Therapy. AAPS J. 19 (2): 456–467. DOI: 10.1208/s12248-016-9991-1.

Cardilin T, Almquist J, Jirstrand M, Zimmermann A, Bawab SE, and Gabrielsson J (2018). Model-Based Evaluation of Radiation and Radiosensitizing Agents in Oncology. CPT Pharmacometrics Syst. Pharmacol. 7 (1): 51–58. DOI: 10.1002/psp4.12268.

Cardilin T, Almquist J, Jirstrand M, Zimmermann A, Lignet F, El Bawab S, and Gabrielsson J (2019). Modeling long-term tumor growth and kill after combinations of radiation and radiosensitizing agents. Cancer Chemother. Pharmacol. 83 (6): 1159–1173. DOI: 10.1007/s00280-019-03829-y.

Certara (2020). Phoenix NLME. Version 8.3. Princetown, USA.

Chen G, Saad ZS, Britton JC, Pine DS, and Cox RW (2013). Linear mixed-effects modeling approach to FMRI group analysis. Neuroimage 73: 176–190. DOI: 10.1016/j.neuroimage.2013.01.047.

Coletti R, Leonardelli L, Parolo S, and Marchetti L (2020). A QSP model of prostate cancer immunotherapy to identify effective combination therapies. Sci. Rep. 10 (1): 1–18. DOI: 10.1038/s41598-020-65590-0.

Cox DR (1972). Regression Models and Life-Tables. J. R. Stat. Soc. Ser. B 34 (2): 187–220.

Davidian M and Giltinan DM (1995). Nonlinear Models for Repeated Measurement Data. Routledge. DOI: 10.1201/9780203745502.

DiCiccio TJ and Efron B (1996). Bootstrap confidence intervals. Stat. Sci. 11 (3). DOI: 10.1214/ss/1032280214.

Dickinson RP and Gelinas RJ (1976). Sensitivity Analysis of Ordinary Differential Direct Method. J. Comput. Phys. 21: 123–143. DOI: 10.1016/0021-9991(76)90007-3.

DiMasi JA, Grabowski HG, and Hansen RW (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. J. Health Econ. 47: 20–33. DOI: 10.1016/j.jhealeco.2016.01.012.

Ditlevsen S and De Gaetano A (2005). Mixed effects in stochastic differential equation models. REVSTAT-Statistical J. 3 (2): 137–153.

Ette EI and Williams PJ (2007). Pharmacometrics: The Science of Quantitative Pharmacology. Hoboken, NJ, USA: John Wiley & Sons, Inc. DOI: 10.1002/0470087978.

Eykhoff P (1974). System identification : parameter and state estimation. Wiley-Interscience.

Fidler M, Wilkins JJ, Hooijmaijers R, Post TM, Schoemaker R, Trame MN, Xiong Y, and Wang W (2019). Nonlinear Mixed-Effects Model Development and Simulation Using nlmixr and Related R Open-Source Packages. CPT Pharmacometrics Syst. Pharmacol. 8 (9): 621–633. DOI: 10.1002/psp4.12445.

FitzGerald JM, Bleecker ER, Nair P, Korn S, Ohta K, Lommatzsch M, Ferguson GT, Busse WW, Barker P, Sproule S, Gilmartin G, Werkström V, Aurivillius M, and Goldman M (2016). Benralizumab, an anti-interleukin-5 receptor α monoclonal antibody, as add-on treatment for patients with severe, uncontrolled, eosinophilic asthma (CALIMA): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet 388 (10056): 2128–2141. DOI: 10.1016/S0140-6736(16)31322-8.

FitzHugh R (1961). Impulses and Physiological States in Theoretical Models of Nerve Membrane. Biophys. J. 1 (6): 445–466. DOI: 10.1016/S0006-3495(61)86902-6.

Fröhlich F, Kaltenbacher B, Theis FJ, and Hasenauer J (2017). Scalable Parameter Estimation for Genome-Scale Biochemical Reaction Networks. PLoS Comput. Biol. 13 (1): 1–18. DOI: 10.1371/journal.pcbi.1005331.

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, and Rubin DB (2013). Bayesian Data Analysis. Chapman Hall/CRC.

Gerlee P and Lundh T (2016). Scientific models: Red atoms, white lies and black boxes in a yellow book. Springer International Publishing Switzerland: 1–96. DOI: 10.1007/978-3-319-27081-4.

Hegelund Myrbäck T, Prothon S, Edman K, Leander J, Hashemi M, Dearman M, Edenro G, Svanberg P, Andersson EM, Almquist J, Ämmälä C, Hendrickx R, Taib Z, Johansson KA, Berggren AR, Keen CM, Eriksson UG, Fuhr R, and Carlsson BC (2020). Effects of a selective glucocorticoid receptor modulator (AZD9567) versus prednisolone in healthy volunteers: two phase 1, single-blind, randomised controlled trials. Lancet Rheumatol. 2 (1): e31–e41. DOI: 10.1016/S2665-9913(19)30103-1.

Helmlinger G, Al-Huniti N, Aksenov S, Peskov K, Hallow KM, Chu L, Boulton D, Eriksson U, Hamrén B, Lambert C, Masson E, Tomkinson H, and Stanski D (2017). Drug-disease modeling in the pharmaceutical industry - where mechanistic systems pharmacology and statistical pharmacometrics meet. Eur. J. Pharm. Sci. 109: S39–S46. DOI: 10.1016/j.ejps.2017.05.028.

Irurzun-Arana I, Rackauckas C, McDonald TO, and Trocóniz IF (2020). Beyond Deterministic Models in Drug Discovery and Development. Trends Pharmacol. Sci. 41 (11): 882–895. DOI: 10.1016/j.tips.2020.09.005.

James CA, Barfield MD, Maass KF, Patel SR, and Anderson MD (2020). Will patient-centric sampling become the norm for clinical trials after COVID-19? Nat. Med. 26 (12): 1810. DOI: 10.1038/s41591-020-01144-1.

Janzén DLI, Bergenholm L, Jirstrand M, Parkinson J, Yates J, Evans ND, and Chappell MJ (2016). Parameter Identifiability of Fundamental Pharmacodynamic Models. Front. Physiol. 7. DOI: 10.3389/fphys.2016.00590.

Janzén DLI, Jirstrand M, Chappell MJ, and Evans ND (2018). Extending existing structural identifiability analysis methods to mixed-effects models. Math. Biosci. 295: 1–10. DOI: 10.1016/j.mbs.2017.10.009.

Jazwinsky AH (1970). Stochastic Processes and Filtering Theory. Academic Press.

Karlsson J, Anguelova M, and Jirstrand M (2012). An Efficient Method for Structural Identifiability Analysis of Large Dynamic Systems. IFAC Proc. Vol. 45 (16): 941–946. DOI: 10.3182/20120711-3-BE-2027.00381.

Kimko H and Pinheiro J (2015). Model-based clinical drug development in the past, present and future: A commentary. Br. J. Clin. Pharmacol. 79 (1): 108–116. DOI: 10.1111/bcp.12341.

Klein JP and Moeschberger ML (2003). Survival Analysis. Statistics for Biology and Health. New York, NY: Springer New York. DOI: 10.1007/b97377.

Klim S, Mortensen SB, Kristensen NR, Overgaard RV, and Madsen H (2009). Population stochastic modelling (PSM)—An R package for mixed-effects models based on stochastic differential equations. Comput. Methods Programs Biomed. 94 (3): 279–289. DOI: 10.1016/j.cmpb.2009.02.001.

Kloeden PE and Platen E (1992). Numerical Solution of Stochastic Differential Equations. Springer. DOI: 10.1007/978-3-662-12616-5.

Krengel A, Hauth J, Taskinen MR, Adiels M, and Jirstrand M (2013). A continuous-time adaptive particle filter for estimations under measurement time uncertainties with an application to a plasma-leucine mixed effects model. BMC Syst. Biol. 7: 1–33. DOI: 10.1186/1752-0509-7-8.

Kristensen NR, Madsen H, and Ingwersen SH (2005). Using Stochastic Differential Equations for PK/PD Model Development. J. Pharmacokinet. Pharmacodyn. 32 (1): 109–141. DOI: 10.1007/s10928-005-2105-9.

Król A, Mauguen A, Mazroui Y, Laurent A, Michiels S, and Rondeau V (2017). Tutorial in joint modeling and prediction: A statistical software for correlated longitudinal outcomes, recurrent events and a terminal event. J. Stat. Softw. 81 (3). DOI: 10.18637/jss.v081.i03.

Król A, Palmér R, Rondeau V, Rennard S, Eriksson UG, and Jauhiainen A (2020). Improving the evaluation of COPD exacerbation treatment effects by accounting for early treatment discontinuations: a post-hoc analysis of

randomized clinical trials. Respir. Res. 21 (1): 158. DOI: 10.1186/s12931-020-01419-8.

Kuhn E and Lavielle M (2005). Maximum likelihood estimation in nonlinear mixed effects models. Computational Statistics Data Analysis 49 (4): 1020–1038. DOI: 10.1016/j.csda.2004.07.002.

Lavielle M and Aarons L (2016). What do we mean by identifiability in mixed effects models? J. Pharmacokinet. Pharmacodyn. 43 (1): 111–122. DOI: 10.1007/s10928-015-9459-4.

Leander J, Almquist J, Ahlström C, Gabrielsson J, and Jirstrand M (2015). Mixed Effects Modeling Using Stochastic Differential Equations: Illustrated by Pharmacokinetic Data of Nicotinic Acid in Obese Zucker Rats. AAPS J. 17 (3): 586–596. DOI: 10.1208/s12248-015-9718-8.

Leander J, Almquist J, Johnning A, Larsson J, and Jirstrand M (2020). NLMEModeling: A Wolfram Mathematica Package for Nonlinear Mixed Effects Modeling of Dynamical Systems. arXiv: 2011.06879 [stat.CO].

Leander J, Lundh T, and Jirstrand M (2014). Stochastic differential equations as a tool to regularize the parameter estimation problem for continuous time dynamical systems given discrete time measurements. Math. Biosci. 251 (1): 54–62. DOI: 10.1016/j.mbs.2014.03.001.

Leander J, Sunnåker M, Rekić D, Aksenov S, Eriksson UG, Johansson S, and Parkinson J (2021). A semi-mechanistic exposure-response model to assess the effects of verinurad, a potent URAT1 inhibitor, on serum and urine uric acid in patients with hyperuricemia-associated diseases. J. Pharmacokinet. Pharmacodyn. DOI: 10.1007/s10928-021-09747-y.

Lehman E and Casella G (1998). Theory of Point Estimation. Springer, New York, NY. DOI: https://doi.org/10.1007/b98854.

Lindstrom MJ and Bates DM (1990). Nonlinear Mixed Effects Models for Repeated Measures Data. Biometrics 46. DOI: 10.2307/2532087.

Lixoft SAS (2020). Monolix. Version 2020R1. Antony, France.

Ljung L (1999). System Identification: Theory for the User. 2nd ed. Prentice Hall PTR.

Lotka AJ (1925). Elements of physical biology. Williams & Wilkins.

Marshall SF, Burghaus R, Cosson V, Cheung S, Chenel M, DellaPasqua O, Frey N, Hamrén B, Harnisch L, Ivanow F, Kerbusch T, Lippert J, Milligan PA, Rohou S, Staab A, Steimer JL, Tornøe C, and Visser SA (2016). Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation. CPT Pharmacometrics Syst. Pharmacol. 5 (3): 93–122. DOI: 10.1002/psp4.12049.

Matzuka B, Chittenden J, Monteleone J, and Tran H (2016). Stochastic nonlinear mixed effects: a metformin case study. J. Pharmacokinet. Pharmacodyn. 43 (1): 85–98. DOI: 10.1007/s10928-015-9456-7.

Milligan PA, Brown MJ, Marchant B, Martin SW, Graaf PH van der, Benson N, Nucci G, Nichols DJ, Boyd RA, Mandema JW, Krishnaswami S, Zwillich S, Gruben D, Anziano RJ, Stock TC, and Lalonde RL (2013). Model-Based Drug Development: A Rational Approach to Efficiently Accelerate Drug Development. Clin. Pharmacol. Ther. 93 (6): 502–514. DOI: 10.1038/clpt.2013.54.

Moles CG, Mendes P, and Banga JR (2003). Parameter estimation in biochemical pathways: A comparison of global optimization methods. Genome Res. 13 (11): 2467–2474. DOI: 10.1101/gr.1262503.

Mortensen SB, Klim S, Dammann B, Kristensen NR, Madsen H, and Overgaard RV (2007). A matlab framework for estimation of NLME models using stochastic differential equations: Applications for estimation of insulin secretion rates. J. Pharmacokinet. Pharmacodyn. 34 (5): 623–642. DOI: 10.1007/s10928-007-9062-4.

Nagumo J, Arimoto S, and Yoshizawa S (1962). An Active Pulse Transmission Line Simulating Nerve Axon. Proc. IRE 50 (10): 2061–2070. DOI: 10.1109/JRPROC.1962.288235.

Nguyen TT, Mouksassi MS, Holford N, Al-Huniti N, Freedman I, Hooker AC, John J, Karlsson MO, Mould DR, Perez Ruixo JJ, Plan EL, Savic R, Van Hasselt JG, Weber B, Zhou C, Comets E, and Mentre F (2017). Model evaluation of continuous data pharmacometric models: Metrics and graphics. CPT Pharmacometrics Syst. Pharmacol. 6 (2): 87–109. DOI: 10.1002/psp4.12161.

Nocedal J and Wright S (2006). Numerical Optimization. New York: Springer-Verlag. DOI: 10.1007/978-0-387-40065-5.

Øksendal B (2003). Stochastic Differential Equations: An Introduction with Applications. Universitext. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-14394-6.

Ólafsdóttir HK (2016). Sensitivity-based Gradients for Parameter Estimation in Nonlinear Mixed Effects Models with Deterministic and Stochastic Dynamics. Master's thesis. Chalmers University of Technology.

Ólafsdóttir HK, Leander J, Almquist J, and Jirstrand M (2018). Exact Gradients Improve Parameter Estimation in Nonlinear Mixed Effects Models with Stochastic Dynamics. AAPS J. 20 (5): 1–13. DOI: 10.1208/s12248-018-0232-7.

Overgaard RV, Jonsson N, Tornøe CW, and Madsen H (2005). Non-linear mixed-effects models with stochastic differential equations: Implementation of an estimation algorithm. J. Pharmacokinet. Pharmacodyn. 32 (1): 85–107. DOI: 10.1007/s10928-005-2104-x.

Picchini U, Ditlevsen S, De Gaetano A, and Lansky P (2008). Parameters of the Diffusion Leaky Integrate-and-Fire Neuronal Model for a Slowly Fluc-

tuating Signal. Neural Comput. 20 (11): 2696–2714. DOI: 10.1162/neco.2008.11-07-653.

Picchini U and Forman JL (2019). Bayesian inference for stochastic differential equation mixed effects models of a tumour xenography study. J. R. Stat. Soc. C. 68 (4): 887–913. DOI: 10.1111/rssc.12347.

Raue A, Karlsson J, Saccomani MP, Jirstrand M, and Timmer J (2014). Comparison of approaches for parameter identifiability analysis of biological systems. Bioinformatics 30 (10): 1440–1448. DOI: 10.1093/bioinformatics/btu006.

Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, and Timmer J (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics 25 (15): 1923–1929. DOI: 10.1093/bioinformatics/btp358.

Rekić D, Johansson S, and Leander J (2020). Higher Febuxostat Exposure Observed in Asian Compared with Caucasian Subjects Independent of Bodyweight. Clin. Pharmacokinet. 60 (3): 319–328. DOI: 10.1007/s40262-020-00943-6.

Ribeiro AH, Tiels K, Umenberger J, Schön TB, and Aguirre LA (2020). On the smoothness of nonlinear system identification. Automatica 121: 109158. DOI: 10.1016/j.automatica.2020.109158.

Rizopoulos D (2016). The R Package JMbayes for Fitting Joint Models for Longitudinal and Time-to-Event Data Using MCMC. J. Stat. Softw. 72 (7). DOI: 10.18637/jss.v072.i07.

Rodriguez-Fernandez M, Mendes P, and Banga JR (2006). A hybrid approach for efficient and robust parameter estimation in biochemical pathways. Biosystems 83 (2-3): 248–265. DOI: 10.1016/j.biosystems.2005.06.016.

Schittkowski K (2002). Numerical Data Fitting in Dynamical Systems. Vol. 77. Applied Optimization. Boston, MA: Springer US. DOI: 10.1007/978-1-4419-5762-7.

Schön TB, Svensson A, Murray L, and Lindsten F (2018). Probabilistic learning of nonlinear dynamical systems using sequential Monte Carlo. Mech. Syst. Signal Process. 104: 866–883. DOI: 10.1016/j.ymssp.2017.10.033.

Schwarz G (1978). Estimating the Dimension of a Model. Ann. Stat. 6 (2): 461–464. DOI: 10.1214/aos/1176344136.

Serban R and Hindmarsh AC (2005). CVODES: The Sensitivity-Enabled ODE Solver in SUNDIALS. Proceedings of the ASME 2005 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 6: 5th International Conference on Multibody Systems, Nonlinear Dynamics, and Control, Parts A, B, and C. ASME: 257–269. DOI: 10.1115/DETC2005-85597.

Sheiner LB and Beal SL (1980). Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-menten model: Routine clinical pharmacokinetic data. J. Pharmacokinet. Biopharm. 8 (6): 553–571. DOI: 10.1007/BF01060053.

Sheiner LB and Beal SL (1981). Evaluation of methods for estimating population pharmacokinetic parameters II. Biexponential model and experimental pharmacokinetic data. J. Pharmacokinet. Biopharm. 9 (5): 635–651. DOI: 10.1007/BF01061030.

Sheiner LB and Beal SL (1983). Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model: Routine clinical pharmacokinetic data. J. Pharmacokinet. Biopharm. 11 (3): 303–319. DOI: 10.1007/BF01061870.

Sheiner LB, Rosenberg B, and Marathe VV (1977). Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. J. Pharmacokinet. Biopharm. 5 (5): 445–479. DOI: 10.1007/BF01061728.

Sirkiä S, Heinonen J, Miina J, and Eerikäinen K (2015). Subject-Specific prediction using a nonlinear mixed model: Consequences of different approaches. For. Sci. 61 (2): 205–212. DOI: 10.5849/forsci.13-142.

Stapor P, Fröhlich F, and Hasenauer J (2018). Optimization and profile calculation of ODE models using second order adjoint sensitivity analysis. Bioinformatics 34 (13): i151–i159. DOI: 10.1093/bioinformatics/bty230.

Tapani S, Almquist J, Leander J, Ahlström C, Peletier LA, Jirstrand M, and Gabrielsson J (2014). Joint feedback analysis modeling of nonesterified fatty acids in obese zucker rats and normal sprague-dawley rats after different routes of administration of nicotinic acid. J. Pharm. Sci. 103 (8): 2571–2584. DOI: 10.1002/jps.24077.

Therneau TM and Grambsch PM (2000). Modeling Survival Data: Extending the Cox Model. Springer. DOI: 10.1007/978-1-4757-3294-8.

Tornøe CW, Overgaard RV, Agersø H, Nielsen HA, Madsen H, and Jonsson EN (2005). Stochastic differential equations in NONMEM®: Implementation, application, and comparison with ordinary differential equations. Pharm. Res. 22 (8): 1247–1258. DOI: 10.1007/s11095-005-5269-5.

Volterra V (1926). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. Mem. Accad. Lincei 6: 31–113.

Vonesh E (1996). A note on the use of Laplace's approximation for nonlinear mixed-effects models. Biometrika 83 (2): 447–452. DOI: 10.1093/biomet/83.2.447.

Wang Y (2007). Derivation of various NONMEM estimation methods. J. Pharmacokinet. Pharmacodyn. 34 (5): 575–593. DOI: 10.1007/s10928-007-9060-6.

Wilks SS (1938). The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. The Annals of Mathematical Statistics 9 (1): 60–62. DOI: 10.1214/aoms/1177732360.

Wiqvist S, Golightly A, McLean AT, and Picchini U (2021). Efficient inference for stochastic differential equation mixed-effects models using correlated particle pseudo-marginal algorithms. Comput. Stat. Data Anal. 157: 107151. DOI: 10.1016/j.csda.2020.107151.

Wolfram Research, Inc. (2020). Mathematica. Version 12.1. Champaign, IL.

Wood AJ (2006). A Proposal for Radical Changes in the Drug-Approval Process. N. Engl. J. Med. 355 (6): 618–623. DOI: 10.1056/nejmsb055203.