PK/PD modelling of subcutaneous glucose dynamics

Alexander Hildenbrand Hansen

Kongens Lyngby 2013
DTU Compute-M.Sc.-2013-19

Supervisor: Henrik Madsen
Co-supervisor: Anne Katrine Duun-Henriksen

Technical University of Denmark
Department of Applied Mathematics and Computer Science
Building 303B, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253031, Fax +45 45882673
compute@compute.dtu.dk
www.compute.dtu.dk DTU Compute-M.Sc.-2013-19

Abstract

This thesis describes the development and improvement of a physiological system consisting of stochastic differential equations on state-space form that models glucose-insulin dynamics in a Type 1 Diabetes patient. The study is a part of the DiaCon collaboration. The system aims at predicting one patient's glucose dynamics while resting, without any exogenous inputs positively exciting the response. Traditionally, ordinary differential equations are used in PK/PD modelling. The incorporation of stochastic differential equations allows a separation of noise into measurement noise, arising from data collection, and system noise such as random biological variation and model deficiencies.

The model parameters are found by *maximum likelihood estimation* using the tool CTSM-R. The *Extended Kalman Filter* is used to calculate the likelihood function in order to estimate the optimum parameter set, $\hat{\boldsymbol{\theta}}$. Having estimated a parameter set, physiological and statistical validations are considered for further improvement.

The CTSM-R tool allows stochastic modelling of continuous time series giving a unique tool to improve existing models. When approaching the system, focus has been on grey-box modelling. Progression is tracked using statistical methods. This thesis revealed an error in the CTSM-R tool, in the calculation of the *initial covariance matrix*. This error is now corrected. Specific focus is devoted to *maximum a posteriori probability* to ensure physiological reliable estimates. Moreover, a combination with *autoregressive processes* has shown to improve predictions, also when adding multiple output sensors.

The final system of non-linear stochastic differential equations has successfully been used to model glucose-insulin dynamics accurately.

Resumé

Dette kandidatspeciale beskriver udviklingen og forbedringen af et fysisk system af stokastiske differentialligninger på state-space form, der vil blive brugt til at modellere glukose-insulin dynamik i en Type 1 Diabetes patient. Specialet er en del af *DiaCon collaboration*. Systemet sigter efter at forudsige en patients glukosedynamik under hvile, uden exogene input der positivt exciterer responset. Traditionelt bruges ordinære differentialligninger til PK/PD modellering. Brugen af stokastiske differentialligninger tillader en opdelingen af støjen i målestøj, der kommer fra data opsamlingen og systemstøj såsom biologisk støj og ufuldkommenhed i modellen.

Ved at bruge værktøjet CTSM-R opnås *maximum likelihood estimater* af parameterne. Det *Extended Kalman Filter* udregner likelihood funktionen for at estimere det optimale parameter set, $\hat{\boldsymbol{\theta}}$. Når dette parameter set er fundet, bruges fysiologisk og statistisk validering i overvejelserne for yderligere forbedring.

Værktøjet CTSM-R tillader stokastisk modellering af kontinuerte tidsserier, og giver dermed et unikt værktøj til forbedring af eksisterende modeller. Ved implementering af systemet har fokus været *grey-box* modellering der tillader fysisk fortolkelige systemer. Fremskridt er overvåget ved brug af statistiske metoder. Dette speciale afslørede en fejl i værktøjet CTSM-R i udregningen *initial covariance matrix*. Denne fejl er nu rettet. Specifik fokus er lagt på *maximum a posteriori probability* der sikrer fysiologiske meningsfyldte estimater. Desuden vises det at i en kombination med autoregressive processer opnåes forbedrede forudsigelser, også når der tilføjes flere output sensorer.

Det endelige system af ikke-linære stokastiske differntialligninger, er succesfuldt blevet brugt til en præcis modellering af glucose-insulin dynamik.

Preface

This thesis was prepared at the Department of Applied Mathematics and Computer Science at the Technical University of Denmark in fulfilment of the requirements for acquiring a M.Sc. in Medicine and Technology. The work started October 1, 2012.

The thesis deals with PK/PD modelling of glucose-insulin dynamics in Type 1 Diabetes patients at rest, using data kindly provided by the DiaCon collaboration. A system of stochastic differential equations on state-space form was investigated and improved based on statistical and physiological considerations.

The reader is assumed to have an insight into modelling of dynamic systems and a general understanding of basic physiology.

Acknowledgements

Special thanks go to my supervisor Henrik Madsen for his support, guidance and inspirational ideas, and to my co-supervisor Anne Katrine Duun-Henriksen for patient guidance in both pharmacological and statistical issues. Thanks to Rune Juhl for his help with CTMS-R and thanks to the involved members in the Section for Dynamical Systems. Finally, I thank Emil Jørgensen for mathematical discussions.

Lyngby, 27-March-2013

Alexander Hildenbrand Hansen

Contents

List of Abbreviations and Symbols

Abbreviations

ACF	Autocorrelation function
ADA	American Diabetes Association
AIC	Akaike information criterion
AR	Autoregressive
C-L	Closed-Loop
CGM	Continuous glucose monitor
CTSM	Continuous Time Stochastic Modelling
DF	Degrees of freedom
MAP	Maximum A Posteriori
MVP	Medtronic Virtual Patient
ODE	Ordinary Differential Equation
PD	Pharmacodynamics
PK	Pharmacokinetics
sCGM	Measurement noise of the CGM device
SDE	Stochastic Differential Equations

sYSI	Measurement noise of the YSI device
T1D	Type 1 Diabetes
YSI	Plasma glucose observation

Symbols

$\boldsymbol{\omega}_t$	Wiener process
$\boldsymbol{\sigma}(\cdot)$	Magnitude of diffusion term
$\boldsymbol{\theta}$	Parameters of the system
$\boldsymbol{e}_j$	White noise
$\boldsymbol{f}(\cdot)$	Drift term of stochastic integral
$\boldsymbol{R}$	Covariance matrix
$\boldsymbol{R}_{\boldsymbol{\theta}}$	Prior correlation matrix
$\boldsymbol{u}_t$	Vector of input variables
$\boldsymbol{x}_t$	State vector at time t
$\boldsymbol{y}_j$	Observation vector at time j
Ω	Parameter set
k_{ij}	Rate parameter
L	Likelihood
p	Probability density function

Chapter 1

Introduction

Patients suffering from diabetes require exogenous insulin delivered by, for example a pump to regulate glucose. A *continuous glucose monitor, CGM,* measures subcutaneous glucose values.

This thesis deals with the creation of a model capable of predicting glucose-insulin dynamics using *stochastic differential equations, SDEs*, in diabetes patients at rest. In experiments of glucose-insulin dynamics patients receive inputs to excitate glucose values such as meals giving a positive excitation and insulin giving a negative excitation. This thesis accommodates the need for an understanding of the dynamics with no positive excitation, to aid in the development of a *glucose monitoring/insulin delivering*, also called a *Closed-Loop, C-L* system. This development is only possible with an substantial understanding of the *pharmacokinetics, PK,* and the *pharmacodynamics, PD,* of insulin.

In PK/PD modelling the properties of a drug from administration to utilisation is described. It will be elaborated later, but in short:

- PK describes *what the body does to the drug.*
- PD describes *what the drug does to the body.*

Various modelling methods can be used [3, 21, 11], but this thesis focuses on stochastic differential equations on state-space form. Stochastic state-space models allow incorporation of observations of the system and makes it possi-

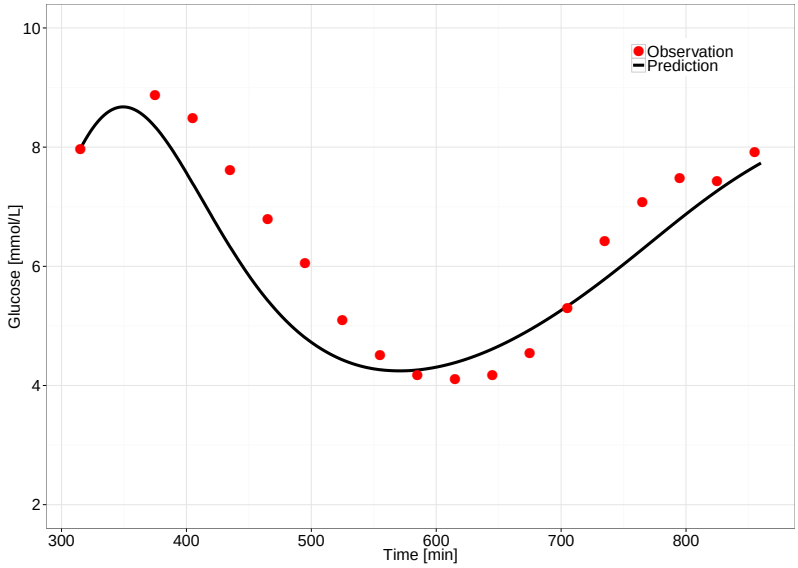


Figure 1.1: Prediction of plasma glucose level using SDEs with negligible diffusion terms thereby resembeling ODEs.

ble to describe model parameters physiologically. Combining information from physiology and information from data is known as *grey-box modelling* [29].

Traditionally PK/PD modelling has been performed using *ordinary differential equations, ODEs*. ODEs have shown to give feasible models, however they are not able to separate measurement noise and system noise, leading to systematic error structure as Figure 1.1 illustrates[1]. Stochastic differential equations are able to make this distinction, enabling an interpretation of model deficiencies. This makes SDEs an attractive extension, since it is not reasonable to assume that models generate perfect predictions for all future time points.

Goal

The main goal of this thesis is to identify a non-linear stochastic state-space model, that based on different types of observations predicts glucose values.

The model development will be an iterative process where validation steps examine statistical properties during the progression. The model is developed and tested on an 9 hour study with patients at rest.

The foundation of the modelling system is the *Medtronic Virtual Patient, MVP*

[1] There is a structure in observations above and below the prediction.

model presented by S. Kanderian in [24]. The MVP model is a system of ODEs. It is extended to a stochastic state-space model and this stochastic model is then the initial point of the investigations. The aim is to achieve significantly better model estimates than the original ODE system.

1.1 Glucose homeostasis

> *"Homeostasis is the property of the body that regulates an internal environment and maintains a relatively constant condition of a property such as glucose."*
>
> Claude Bernard, 1878 in [5]

All cells need fuel, and for most cells this fuel is glucose. Tight regulation of plasma glucose levels is vital since *elevated glucose levels, hyperglycemia,* is toxic and increases the risk of hyperglycemic associated diseases, with the most severe being renal destruction and vascular problems. More dangerous is *low glucose levels, hypoglycemia,* leading to loss of consciousness followed by death if untreated [53]. As stated by Claude Bernard, homeostasis autonomously regulates plasma glucose levels to maintain an ideal range between 4.0 and 8.0 $\mathrm{mmol\,L^{-1}}$ [53][2] for a healthy person. Most tissues in the body, except nervous, can only take up glucose if insulin is present. Therefore insulin secretion is as important as intake of food.

The complete mechanics behind glucose homeostasis is of course very complex and is out of the scope of this thesis. But the primary principles, explained in this section, give an understanding of the challenges.

Plasma glucose regulation is primarily maintained by three pancreatic hormones: 1) *insulin*, 2) *glucagon* and 3) *somatostatin*. In a normally functioning pancreas the regulation is autonomous. The islets of Langerhans is the endocrine tissue regulating secretion of insulin from Beta-cells (β-cells), glucagon from Alpha-cells (α-cells) and somatostatin from Delta-cells (δ-cells) - with insulin being the primary regulator[3]. Only the effect of insulin is examined in this thesis.

There are many factors governing the secretion of insulin[4]. Known factors include: stress level, exercise, neural stimulation, hormonal control and plasma

[2] [43] defines an ideal range between 4.0 and 7.0 $\mathrm{mmol\,L^{-1}}$

[3] Glucagon increases the breakdown of glycogen in the liver. Somatostatin inhibits the secretion of glucagon and insulin

[4] Not all factors are known [45].

glucose levels. Likewise there is a large family of *Glucose Transporters, GLUTs*, that catalyse the transportation of glucose. GLUTs for the nervous tissue, GLUT-1 and GLUT-3, do not need insulin. Others, GLUT-4 and GLUT-8, are insulin dependent [45]. This thesis focuses on the insulin-dependent glucose uptake. The main concept of insulin-dependent glucose uptake is easily described:

1. An incline in plasma glucose, often caused by intake of a meal, causes β-cells to secrete insulin.
2. This facilitates an uptake of glucose in insulin-dependent tissue making plasma glucose levels decline. The half-life of insulin is short, around 10 min [43].
3. If plasma glucose falls below the ideal range, secretion of insulin from the β-cells is inhibited, and the liver starts synthesising glucose from different processes (hepatic processes)[5] [53].

1.1.1 Diabetes mellitus

The total number of diabetes patients in the year 2000 was 171 million and this number is expected to increase to 366 million in 2030 [60]. *Type 1 Diabetes, T1D*, consists of around 5% of the total number of diabetes patients. The cost of diabetes is high, and since T1D is incurable patients rely on a lifetime of treatment. Patients suffering from diabetes need a tight regulation of glucose values in the range 5.0 and 7.2 $\mathrm{mmol\,L^{-1}}$ according to the recommendations of the *American Diabetes Association, ADA,*.

In T1D patients, production of insulin from β-cells diminishes or ceases. The destruction of the islets of Langerhans is autoimmune, and manifests when about 10% of the islets are left [53]. As made clear before, insulin secretion has an essential role in plasma glucose regulation, so a malfunction is severe. Untreated, a patients plasma glucose increases as the tissue is unable to take up glucose. The patient quickly looses weight and will eventually pass away.

The current treatment for T1D is exogenous delivery of insulin. The amount of insulin to be delivered depends on the glucose level. For this thesis it is coarse-grained assumed that glucose levels can be divided in two groups: 1) Blood glucose, more precisely *plasma glucose* and 2) *subcutaneous glucose*.

In a clinical setting, plasma glucose can be measured in veins, by YSI2300

[5] The complete glucose-insulin dynamics is out of the scope of this thesis, therefore only this very simplified fundamental concept is given.

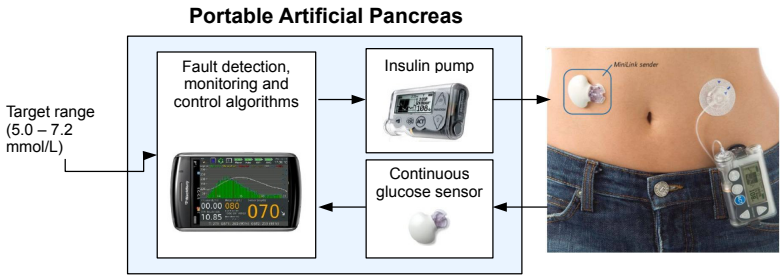


Figure 1.2: Illustration of the concept of an artificial pancreas from [12].

STAT Plus[6], also considered as "gold standard". In every-day-life, patients can use fingerstick measurement and obtain plasma glucose values or use a CGM device, and obtain a subcutaneous glucose value. Based on this information the desired amount of insulin can be delivered.

Traditionally there are two ways to deliver insulin:

1. Manual pen-injections of fast or slow acting insulin four to six times a day.
2. Automatical injections using a pump, to continuously deliver small amounts of fast acting insulin and insulin boluses when appropriate.

1.2 Motivation

The holy grail of diabetes treatment is an artificial pancreas, that without influence from the user, delivers the correct amount of insulin based on subcutaneous glucose measurements.

An artificial pancreas is essentially a C-L system that, based on a model and a controller, delivers rapid acting insulin to regulate glucose levels, as illustrated by Figure 1.2.

A proper modelling of glucose-insulin dynamics is essential for a high-quality artificial pancreas because it relies on observations of the subcutaneous glucose and not plasma glucose levels. An identification of the dynamics, will aid in the improvement of life quality of diabetes patients.

[6]Yellow Springs Instruments, Yellow Springs, OH.

Before an artificial pancreas can be developed the relationship between plasma glucose and subcutaneous glucose must first be identified. Traditionally, subcutaneous glucose levels has been regarded as a delay of plasma glucose levels, but the relationship may be more complicated [63].

The DiaCon collaboration has, to test a C-L system, conducted a study measuring plasma glucose, subcutaneous glucose and plasma insulin in T1D-patients at rest. Data from this study is used in this thesis. This data is presented in the following section.

1.3 Data

The patients examined in this thesis are a part of a larger study conducted by the DiaCon collaboration. A detailed description can be found in [52], but a brief summary and characteristics of relevant patients is given here.

The study consists of two randomised crossover[7] studies. Study I compared a C-L system with Open-Loop insulin pump treatment. Study II tested performance of a C-L system with euglycemic and hyperglycemic ranges at the start of the study.

For this thesis, patients from Study II were selected for further investigation. Characteristics of the six patients (three men) where: age 35 ± 11 years, BMI 25.2 ± 4.6 kg/m^2 and total daily insulin dose 0.6 ± 0.1 IU/kg/day treated with insulin Aspart[8]. Patients did not use medicine affecting plasma glucose, other than insulin.

Blood samples for plasma glucose and insulin Aspart analysis was drawn every 30 min, during all studies. Plasma glucose was analysed using YSI2300. Observations of plasma glucose levels obtained with the YSI2300 will be referred to as *YSI observations*. Patients did not self-administer carbohydrates or correction boluses, but alarming values or prominent symptoms of hypoglycemia resulted in administration of glucose intravenously by medical professionals.

Every 5 min a subcutaneous glucose value was obtained using two DexCom SevenPLUS[9] CGM devices[10] placed at different locations on the body. Obser-

[7] A study in which subjects are randomly assigned to different treatments and then switched halfway in the treatment.

[8] Novorapid, Novo Nordisk, Bagsværd, Denmark.

[9] DexCom, San Diego, CA.

[10] Using an electrochemical difference to calculate the subcutaneous glucose value.

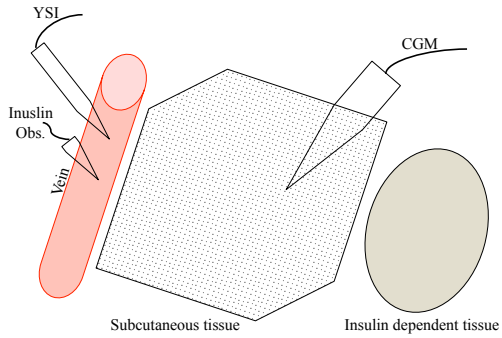


Figure 1.3: Schematic overview of measurement setup. YSI measures plasma glucose levels, G_p. The CGM device measures subcutaneous glucose levels, G_{sc}. Insulin dependent tissue symbolises tissue requiring insulin to take up glucose.

vations of subcutaneous glucose levels are referred to as *CGM observations*. A schematic overview of the setup is seen in Figure 1.3.

Insulin was administered using Paradigm Veo[11] insulin pump. A small continuous basal insulin infusion of 0.025 IU/h is given, because this pump could not be turned completely off. Every 15 min a dose suggestion is given by the Closed-loop system that is approved and administered by a medical professional.

Patients were in bed during the study from 22:00 to 07:00, with the C-L system being initiated at study start. In Study II, two different types of time series were at hand. One with euglycemic values and another with hyperglycemic values at study start. A patient with euglycemic values at study start was chosen in order to get optimal conditions in the modelling and to best describe glucose-insulin dynamics.

	Study start	Study End
24 hour scale	22:00	07:00
Minute scale	315 min	860 min

Table 1.1: Link between 24 hour scale and study time index. As there has been pre-recording in the study, the study time index does not start in zero. The duration of the study is approximatlty 9 hours.

Table 1.1 shows the link between 24 hour scale and the study time index used in this thesis. The recording has commenced before study start, thus 22:00 is not

[11] Medtronic, Northridge, CA

0 min. This pre-recording is cut from the data before modelling, and is therefore not shown in the figures presented throughout this thesis.

The patients glucose dynamic is seen in Figure 1.4, showing the observations of the two CGM devices. It is obvious that the CGM observations differ. For most of this thesis the CGM device shown in the top plot, referred to as the original CGM device, is used. It was assumed that this CGM device had the most correct representation of subcutaneous glucose dynamics. In the final stages the alternate device is also examined. It is evident that glucose dynamics are almost inside the ideal range (shown by the green ribbon). Even though there is no positive excitation in the study period, some dynamics inside the ideal range exists, making modelling of the dynamics possible. No correction bolus or extra carbohydrates were given during the study.

Throughout this thesis YSI observations are represented by a red dot. CGM observations are represented by a blue dot. Predictions are a black line. The ideal range of glucose levels is represented by a green ribbon.

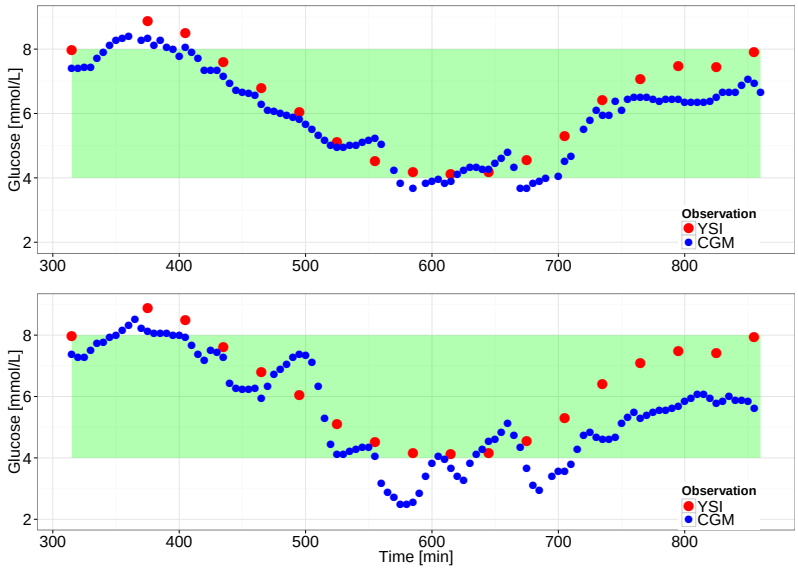


Figure 1.4: The YSI observations (red) and CGM observations (blue) used to estimate parameters in this study. The ideal range of glucose levels is indicated with the green ribbon.
TOP PLOT) The original CGM device.
BOTTOM PLOT) The alternate CGM device.

It is noted that the second YSI observation is missing. The initial investigation concluded that this observation was erroneous as it was measured to 7.2 mmol/l (previous value 8.0 mmol/l and following value 8.9 mmol/l). There are no indications that this sudden "decrease-increase" should have occurred naturally. Because YSI is considered gold standard that observation is deemed an error, and it has been removed to increase the validity and reliability of the model. This removal is caused by the initial investigation in Chapter 5. This removal is of course an assumption and there may exist an explanation for the decrease-increase behaviour.

1.4 Background

Many different models have been suggested to describe glucose-insulin dynamics. During the 1970s Albisser [1] and Pfeiffer [44] worked on designs using intravenous glucose sampling and intravenous insulin infusion, developing the first glucose controlled insulin infusion system [21]. Control algorithms based on intravenous insulin delivery has been developed through the 1980s and 1990s. Parker provides an extensive overview of these algorithms in [41].

Due to complications of the invasive treatments, less invasive and self-treatable methods have been investigated. Subcutaneous delivery of insulin fulfils these requirements. The fundamental concepts used in glucose-insulin dynamic modelling was introduced by Bergman in 1981 [4]. Bergman developed three equations describing the system as simple as possible, using ordinary differential equations.

Bergman's model described glucose-insulin dynamics of the healthy human body. Extending on this concept, Fisher replaced the insulin secretion term with an insulin infusion term in order to accommodate the development of an artificial pancreas [14] for T1D. There are, however, problems regarding glucose control, as an extra delay of insulin absorption is introduced when administering insulin subcutaneously, Cobelli [11]. Insulin kinetics can to some extent be mimicked, but the process is complicated. Wilinska examined 11 different linear and non-linear compartment models for insulin PK, concluding that model selection is highly dependable on available data [62].

The development of real-time devices monitoring subcutaneous glucose values, has since 2005 undergone a tremendous evolution [58], with many different companies producing CGM devices. A comprehensive overview of the mathematical algorithms currently used has been made by Bequette in [3]. The use of CGM devices introduced a lag between plasma glucose levels and subcutaneous glu-

cose levels. Keenan investigated this delay for different CGM devices in 2009, and found a lag of $3 - 14$min, depending both on physiological lag, sensor lag and lag from filtering algorithms [25]. Other studies [55, 31] have found lags of $3 - 5$min.

Many different types of compartment models have been suggested to model glucose-insulin dynamics [28, 61]. The Medtronic Virtual Patient model [24] builds on the concept developed by Bergman. This model is based on five linear and non-linear ordinary differential equations. As shown in Figure 1.1 ordinary differential equations are not always sufficient. This thesis investigates the use of non-linear stochastic differential equations on state-space form. An area only touched lightly upon in the current literature [12, 56, 37].

Stochastic differential equations, though they have not yet gained popularity in PK/PD modelling, have proven very useful in many other different kinds of applications. SDEs have been used to model bacterial growth [22, 38], heat dynamics in buildings [2, 32], evaluation of flow in sewers [7], estimation of electricity markets [17] and financial modelling [18] for example. Hopefully this thesis will contribute to the further exploitation of the use of SDEs in PK/PD modelling.

1.5 Outline of the thesis

Chapter 2 covers concepts of pharmacokinetic and pharmacodynamic modelling and the PK/PD principles used in this thesis are introduced. The chapter also introduces the mathematical theories of: stochastic differential equations on state-space form, Maximum Likelihood principles and Extended Kalman Filtering.

Chapter 3 explains the model systems used to describe glucose-insulin dynamics. Two different systems: the Medtronic Virtual Patient model and the Steil model, are presented and formulated as a state-space system of stochastic differential equations.

Chapter 4 gives an introduction to Chapters 5 to 8 that present and analyse the models.

Chapter 5 describes the discovery of a potential computational flaw in the CTSM-R tool. A visualisation of the problem and a solution to the deficiency is presented.

Chapter 6 is the initial investigation of the Medtronic Virtual Patient model using stochastic differential equations on state-space form. A simplification of the model is established to improve insulin kinetics.

Chapter 7 tests if Steil's, more complex, model describes the variation in the data better than the Medtronic Virtual Patient model, based on a statistical analysis.

Chapter 8 makes a thorough investigation of possible improvements using Maximum A Posteriori probability, autoregressive modelling and a combination of the two approaches.

Chapter 9 is a discussion of the results obtained from the models presented in Chapters 5 to 8.

Chapter 10 summarises the experiences of the thesis into a set of recommendations for future work.

Chapter 11 contains a conclusion of this work.

Chapter 2

Modelling methodology

Many different approaches can be taken when modelling complex data. Before the actual modelling is commenced, it is useful to reflect on the methodology, since the approach depends on available data and the final goal. It is, mostly, desired to build a model as simple as possible, that at the same time describes all variation in the data. In general, there are two different approaches [36]:

Backward selection: Start with the full parameter set Ω_{full} and simplify the model.

Forward selection: Start with the null parameter set Ω_{null} and extend the model.

Backward selection is often preferred when there is an abundance of data and only a small risk of overparameterisation. Forward selection is a popular approach when little or no knowledge about the parameters is available or when a limited amount of data is accessible. Extending the model makes it possible to track changes throughout the development and thereby minimising the risk of errors and overparameterisation. A model is clearly overparameterised if the number of parameters to be estimated is greater than or equal to the number of observations, but smaller models can also be overparameterised [46]. A large parameter set may introduce difficulties with the identifiability of the model. This issue is addressed in a later chapter.

A limited amount of data with small excitations is available. Therefore, it is reasonable to use forward selection as the key approach in this thesis - keeping in mind that the limited number of data enforces a restriction on the complexity of the models. Furthermore the limited amount of data makes the modelling sensitive to biological factors. If the experiment is repeated on a different day, the result is not necessarily the same. This comes from the unavoidable biological factors that influence experiments[12]. This influence also justifies the use of SDEs.

This chapter gives an overview of six important concepts used in this thesis. The first part is based on [50] and explains the general concepts in PK/PD modelling. The second part elaborates the choice of SDEs over ODEs and is based on [35] by Henrik Madsen. The third part gives an insight to the maximum likelihood method and likelihood testing, used to asses improvements of the models, and is based on [36]. The fourth part, based on [34, 35, 30], introduces the Extended Kalman Filter, used to find the optimal parameter estimates: $\hat{\boldsymbol{\theta}}$. Finally, the fifth part describes validation and the sixth part describes the CTSM-R tool.

2.1 PK/PD modelling

Pharmacokinetics describes how adjustable elements as *dose* and *route* of administration are related to drug level-time relationships of the body. In a popular term *"what the body does to the drug"*. Pharmacodynamics describes the relation between the concentration of the drug and the magnitude of the effect. In a popular term *"what the drug does to the body"* [50]. In this thesis, PK describes the bodies influence on insulin: How it is absorbed, distributed and eliminated. The PD part describes the effect of insulin on glucose dynamics.

A popular method in PK/PD modelling is the use of compartment models described by first- or second-order kinetics [50]. There are four fundamental concepts in PK/PD modelling:

Absorption: The process of drug entering the bloodstream.

Distribution: The reversible[13] transfer of drug from one location to another within the body.

Metabolism: The process of a drug being broken down to smaller compounds.

Excretion: Irreversible loss of the drug from the body.

[12] This influence also imposes challenges when developing a global model for an entire population, since there exist a fundamental biological difference between people.

[13] Reversible since the drug can be redistributed.

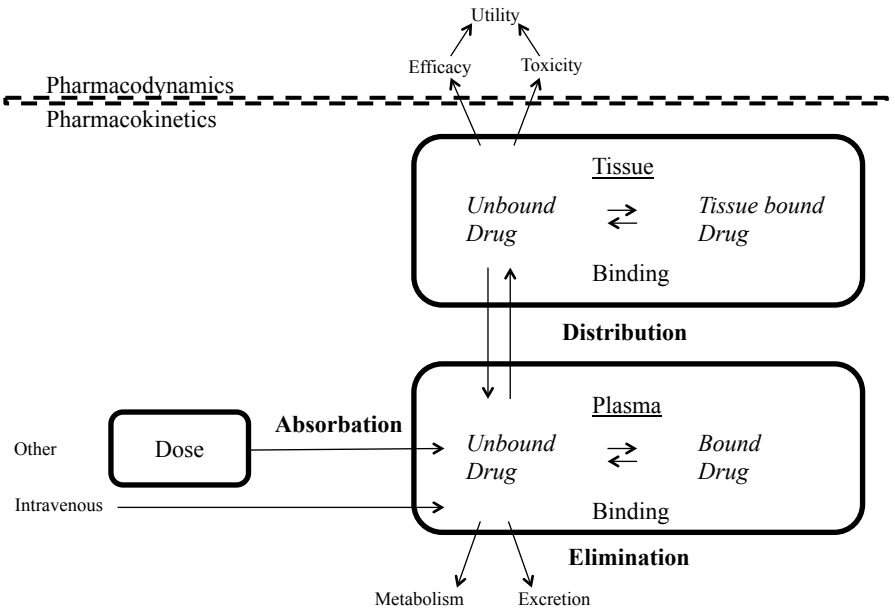


Figure 2.1: Schematic presentation of PK/PD dynamics from [16]

Elimination is the result of metabolism and excretion. A schematic representation is seen in Figure 2.1 [16].

An important assumption in compartment models is mass-balance considerations. The total dose needs to be accounted for in all steps of the process, so that drug does not disappear. This also implies that the sum of rate of change must be zero in the compartments.

Example 2.1 (Compartment model using ODEs) *Drug dynamics can be modelled using compartment models. A simple illustration is seen in Figure 2.2, showing a two-compartment model.*

This simple system can be described by two ordinary differential equations, using first-order kinetics, as:

$$d\boldsymbol{x}_t = d\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} k_{21} \cdot C_2 - (k_{12} + k_e) \cdot C_1 \\ k_{12} \cdot C_1 - k_{21} \cdot C_2 \end{bmatrix} dt \tag{2.1}$$

where C_1 and C_2 are the concentrations in compartment 1 and 2, respectively, and k_{ij} is a rateparameter where: k_{12} is the rate of change from C_1 to C_2 , k_{21} is the rate of change from C_2 to C_1 and k_e is the rate of elimination of the drug. □

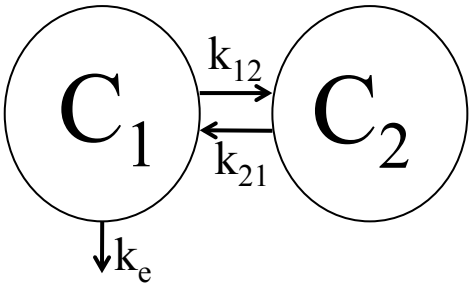


Figure 2.2: Scematic representation of a two-compartment model, illustrating the fundamental compartment model concept.

The models used in the thesis are more complicated. But the compartment approach from Example 2.1 is the fundamental concept.

2.1.1 Grey-box modelling

In grey-box modelling, model structure comes from physical and statistical considerations. Based on these considerations the parameters of the model are estimated. The parameter estimates are validated statistically and physically and compared to available results in the literature for a physiological interpretation of the model.

The concept of grey-box modelling is easily linked to Example 2.1. There are restrictions on the concentration (it can not exceed 100 %) and the estimated rate constants can be compared to known rate constants for that particular drug.

Example 2.1 uses ODEs to model the data. A physical system using ODEs can be written on state-space form. State-space allows an easily interpreted and physiologically meaningful representation of a physiological system. A state-space representation is shown in Equation (2.2):

$$d\boldsymbol{x}_t = \boldsymbol{f}\left(\boldsymbol{x}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta}\right) dt \tag{2.2a}$$

$$\boldsymbol{y}_j = \boldsymbol{h}\left(\boldsymbol{x}_j, \boldsymbol{u}_j, t_j, \boldsymbol{\theta}\right) + \boldsymbol{e}_j \tag{2.2b}$$

where $\boldsymbol{x}_t$ is the state vector, $\boldsymbol{y}_j$ observation vector, $\boldsymbol{u}_t$ is a vector of input variables, $\boldsymbol{\theta}$ is the parameter vector of the system, t is time, $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ are functions and $\boldsymbol{e}_j$ is white noise. Equation (2.2a) is the model for the states, *the state equation*, and Equation (2.2b), defines how the states are observed, *the observation equation*. State-space models are very useful for describing time

varying systems, due to the first-order Markov property of the state vector, since all the information about the future is contained in just $\boldsymbol{x}_t$ [35].

Example 2.2 (Example 2.1 on state-space form) *The state equation for the two-compartment model Example 2.1 can now easily be written on state-space from as shown in Equation (2.3).*

$$\begin{aligned} d\boldsymbol{x}_t &= d\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} k_{21} \cdot C_2 - (k_{12} + k_e) \cdot C_1 \\ k_{12} \cdot C_1 - k_{21} \cdot C_2 \end{bmatrix} dt \\ \boldsymbol{y_j} &= h(\boldsymbol{x}_j) + \boldsymbol{e}_j \end{aligned} \tag{2.3}$$

where $\boldsymbol{y}_j$ is the observation equation at discrete time instances j. □

There is, however, some critical points where the use of ODEs are problematic. As shown in the Chapter 1, ODEs are sometimes not sufficient enough to describe the variation in the data and they assume perfect predictions. It is desired that incorrect model specifications, random variation or uncertainties from measurements can be described as a source of error. If the ODE models do not give a sufficient description of the error structure, then SDEs should be considered.

2.2 Stochastic differential equations

By using SDEs, noise is allowed into the state equations. The states are therefore predicted with uncertainty. This is particularly useful when fluctuations or disturbances are expected to influence the states. SDEs also allow a split between correlated system noise and uncorrelated observation noise [56]. A SDE can be written as:

$$d\boldsymbol{x}_t = \boldsymbol{f}\left(\boldsymbol{x}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta}\right) dt + \boldsymbol{\sigma}\left(\boldsymbol{u}_t, t, \boldsymbol{\theta}\right) d\boldsymbol{\omega}_t \tag{2.4}$$

where $t \in \mathbb{R}$ is time, $\boldsymbol{u}_t$ is a vector of input variables and $\boldsymbol{\theta}$ is the parameter vector of the model. $\boldsymbol{f}(\cdot)$ and $\boldsymbol{\sigma}(\cdot)$ are non-linear functions. $\boldsymbol{f}(\cdot)$ is the drift term and $\boldsymbol{\sigma}\left(\boldsymbol{u}_t, t, \boldsymbol{\theta}\right) d\boldsymbol{\omega}_t$ is the diffusion term, where $\{\boldsymbol{\omega}_t\}$ is a Wiener process and $\boldsymbol{\sigma}(\cdot)$ is the magnitude of the Wiener process.

Two fundamental concepts in SDE theory is the Wiener process and the Itô integral. The following account is based on [40, 6]. A process is an indexed family of random variables where $T = [0, \infty)$ is often used to represent a systems evolution over time. One stochastic process is [49]:

Definition 2.1 (Gaussian process) A *Gaussian process* is a stochastic process $X_t, t \in T$, such that any finite-dimensional marginal distribution is jointly Gaussian. □

One well-known Gaussian process is the Wiener process. The Wiener process is not stationary, but has stationary and independent increments.

Definition 2.2 (Wiener process) The *Wiener process* is a Gaussian stochastic process $\{\boldsymbol{\omega}_t \geq 0\}$, with the following mathematical properties

i $\boldsymbol{\omega}_0 = 0$

ii The function $t \to \boldsymbol{\omega}_t$ is almost surely everywhere continuous

iii The increment $\boldsymbol{\omega}_t - \boldsymbol{\omega}_s$ is independent with $N(0, t-s)$ for any $0 \leq s < t$

The Wiener process is also known as a standard Brownian motion. □

An important property of the Wiener process is that it is almost everywhere non-differentiable, furthermore the Wiener process has unbounded variation. This leads to the the following notation of stochastic integral:

Definition 2.3 (Itô SDE) An *Itô stochastic differential equation* is an integral equation of the form:

$$\boldsymbol{x}_t = \boldsymbol{x}_0 + \int_0^t \boldsymbol{f}\left(\boldsymbol{x}_s, s\right) ds + \int_0^t \boldsymbol{\sigma}\left(\boldsymbol{x}_s, s\right) d\boldsymbol{\omega}_s \tag{2.5}$$

where $\boldsymbol{f}(\cdot)$ is the *drift term* and $\boldsymbol{\sigma}\left(\cdot\right) d\boldsymbol{\omega}_s$ is the diffusion term.

The first integral can be interpreted in a Riemannian sense, the second as an Itô integal.

A shorthanded Itô notation of Equation (2.5) is:

$$\begin{aligned} d\boldsymbol{x}_t &= \boldsymbol{f}\left(\boldsymbol{x}_t, t\right) dt + \boldsymbol{\sigma}\left(\boldsymbol{x}_t, t\right) d\boldsymbol{\omega}_t \\ \boldsymbol{x}_0 &= a \end{aligned} \tag{2.6}$$

The solution to Equation 2.5 is a stochastic process, $\boldsymbol{x}_t, t \geq 0$.

Equation (2.4) is now formally justified. □

With the knowledge about SDEs it is possible to extend the state-space model from Section 2.1, Equation (2.2), to a (non)-linear stochastic differential equation on state-space form:

$$d\boldsymbol{x}_t = \boldsymbol{f}\left(\boldsymbol{x}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta}\right) dt + \boldsymbol{\sigma}\left(\boldsymbol{u}_t, t, \boldsymbol{\theta}\right) d\boldsymbol{\omega}_t \tag{2.7a}$$

$$\boldsymbol{y}_j = \boldsymbol{h}\left(\boldsymbol{x}_j, \boldsymbol{u}_j, t_j, \boldsymbol{\theta}\right) + \boldsymbol{e}_j \tag{2.7b}$$

where $t \in \mathbb{R}$ is time, $\boldsymbol{x}_t$ is the state vector, $\boldsymbol{y}_j$ is the observation vector, $\boldsymbol{u}_t$ is a vector of input variables and $\boldsymbol{\theta}$ is the parameter vector. $\boldsymbol{f}(\cdot)$, $\boldsymbol{\sigma}(\cdot)$ and $\boldsymbol{h}(\cdot)$ are non-linear functions, $\{\boldsymbol{\omega}_t\}$ is a Wiener process and $\{e_j\}$ is a white noise process.

Example 2.3 (Example 2.2 on stochastic form) *It is easily seen that the two-compartment state space representation in Equation (2.3), can be written on stochastic form by adding a diffusion term, $\boldsymbol{\sigma} d\omega_t$:*

$$\begin{aligned} d\boldsymbol{x}_t &= d\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} k_{21} \cdot C_2 - (k_{12} + k_e) \cdot C_1 \\ k_{12} \cdot C_1 - k_{21} \cdot C_2 \end{bmatrix} dt + \boldsymbol{\sigma} \begin{bmatrix} d\omega_1 \\ d\omega_2 \end{bmatrix} \\ \boldsymbol{y_j} &= h(\boldsymbol{x}_j) + \boldsymbol{e}_j \end{aligned} \tag{2.8}$$

where $\boldsymbol{y}_j$ is the observation equation at discrete time instances j. □

When a non-linear model is at hand then the *Extended Kalman Filter, EKF* and the Maximum Likelihood principle can be used to estimate the parameters $\boldsymbol{\theta}$. The concept behind these are explained in the following two sections.

2.3 The Maximum Likelihood principle

Before the EKF is presented, knowledge about *the Maximum Likelihood principle* is necessary. Given a set of data and a model with a number of parameters, it is often of interest to find the parameter estimates so that the model describes the data best. The likelihood function has been identified as the key inferential quantity conveying all information in statistical modelling including uncertainty [36, 15]. The entire likelihood function captures all the information in the data about the parameters. Specifically the maximum likelihood estimate is of interest, as it gives the parameters that are most likely to explain the variation in the data [36].

The data for one patient from a study has the following structure:

$$\boldsymbol{y}_j, \quad j = 1, \ldots, N \tag{2.9}$$

where N is the number of measurements the patient.

Introducing the notation $\mathcal{Y}_N = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N - 1, \boldsymbol{y}_N]$ to describe the data it is possible to express the likelihood function as a product of conditional probability densities:

$$L\left(\boldsymbol{\theta}; \mathcal{Y}_N\right) = \left(\prod_{j=1}^{N} p\left(\boldsymbol{y}_j | \mathcal{Y}_{j-1}, \boldsymbol{\theta}\right)\right) p\left(\boldsymbol{y}_0 | \boldsymbol{\theta}\right) \tag{2.10}$$

where p is the probability density function.

An exact evaluation to the likelihood function is only possible if the initial probability density is known as all subsequent conditional densities then are available. For the non-linear models used in this thesis, an approximate method based on Extended Kalman Filtering, explained in the next section, can be used.

The stochastic differential equations, introduced in the previous section, are driven by Wiener processes with Gaussian increments. Under some regularity conditions[14], it is reasonable to assume that the conditional densities can be approximated by Gaussian densities, completely characterized by its mean, its covariance, and an error term:

$$\hat{\boldsymbol{y}}_{j|j-1} = E\{\boldsymbol{y}_j | \mathcal{Y}_{j-1}, \boldsymbol{\theta}\} \tag{2.11}$$

$$\boldsymbol{R}_{j|j-1} = V\{\boldsymbol{y}_j | \mathcal{Y}_{j-1}, \boldsymbol{\theta}\} \tag{2.12}$$

$$\epsilon_j = \boldsymbol{y}_j - \hat{\boldsymbol{y}}_{j|j-1} \tag{2.13}$$

where $\hat{\boldsymbol{y}}_{j|j-1}$ is the mean, $\boldsymbol{R}_{j|j-1}$ is the covariance and ϵ_j is the error term.

By conditioning on $\boldsymbol{y}_0$ and taking the negative logarithm the Maximum Log-Likelihood is obtained:

$$\begin{aligned} -\ln\left(L\left(\boldsymbol{\theta}; \mathcal{Y}_N | \boldsymbol{y}_0\right)\right) =& \frac{1}{2} \sum_{k=1}^{N} \left(\ln\left(\det\left(\boldsymbol{R}_{k|k-1}\right)\right) + \epsilon_k^T \boldsymbol{R}_{k|k-1}^{-1} \epsilon_k\right) \\ &+ \frac{1}{2} \left(\sum_{k=1}^{N} l\right) \ln\left(2\pi\right) \end{aligned} \tag{2.14}$$

This means that the Maximum Likelihood estimate, $\hat{\boldsymbol{\theta}}$, satisfies the non-linear optimisation problem:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} \in \Theta} \{-\ln\left(L\left(\boldsymbol{\theta}; \mathcal{Y}_N | \boldsymbol{y}_0\right)\right)\} \tag{2.15}$$

[14] Regularity conditions can be found in [30].

2.3.1 Statistical comparison

As mentioned, the limited amount of data means that forward selection is used in model building. When a model has been extended it can be compared to previous models in two ways [36]. If the models are nested Wilk's Likelihood ratio test can be used, otherwise *Akaike information criterion, AIC,* can be used.

Let Ω_{null} and Ω_{full} be two hypothesis, where $\Omega_{null} \subseteq \Omega_{full}$. Then the likelihood ratio becomes [36]:

Definition 2.4 (Likelihood ratio) Consider the hypothesis $\mathcal{H}_0 : \theta \in \Omega_{null}$, against the alternative $\mathcal{H}_1 : \theta \in \Omega_{full} \setminus \Omega_{null}$, where $\dim(\Omega_{null}) = m$ and $\dim(\Omega_{full}) = k$. For given observations $\boldsymbol{y} = [y_1, y_2, ..., y_n]$ the *likelihood ratio* is defined as:

$$\lambda(\boldsymbol{y}) = \frac{\sup_{\boldsymbol{\theta} \in \Omega_{null}} L(\boldsymbol{\theta}; \boldsymbol{y})}{\sup_{\theta \in \Omega_{full}} L(\boldsymbol{\theta}; \boldsymbol{y})} \tag{2.16}$$

□

With the above in mind Wilk's Likelihood ratio test is written as [36]:

Theorem 2.5 (Wilk's Likelihood Ratio Test) *For $\lambda(\boldsymbol{y})$, as in Definition 2.4, then under the null-hypothesis $\mathcal{H}_0$, the random variable $-2log\lambda(\boldsymbol{Y})$ converges in law to a χ^2 random variable with $(k - m)$ degrees of freedom:*

$$-2log\lambda(\boldsymbol{Y}) \to \chi^2(k - m) \tag{2.17}$$

under $\mathcal{H}_0$. □

The evidence against $\mathcal{H}_0$ is measured by the p-value. For hypothesis testing this means that if the calculated p-value is smaller than the significance level of $\alpha = 0.05$ in the null model, $\mathcal{H}_0$, is rejected.

When models are not nested the Akaike information criterion can be used [34]:

Definition 2.6 (Akaike information criterion) Given a set of candidate models for the data the preferred model is the one with the minimum *Akaike information criterion* value, calculated as:

$$AIC = 2 \cdot n - 2 \cdot \ln\left(L\left(\boldsymbol{\theta}; \mathcal{Y}_N | \boldsymbol{y}_0\right)\right)) \tag{2.18}$$

where n is the number of parameters, and the remaining parts follow the definitions in Section 2.3. □

The forward selection procedure using Wilk's Likelihood ratio test is visualised in Figure 2.3. It shows how for each calculated p-value, a decision of what diffusion term to include is made. This procedure continues until $p > 0.05$. When presenting results in the following chapters, this selection procedure is presented like Figure 2.3, to easily visualise the included diffusion terms, where $l(L_{\Omega_{ij}})$ is the loglikelihood, from now referred to as *the likelihood*, and p_{ij} is the corresponding p-value.

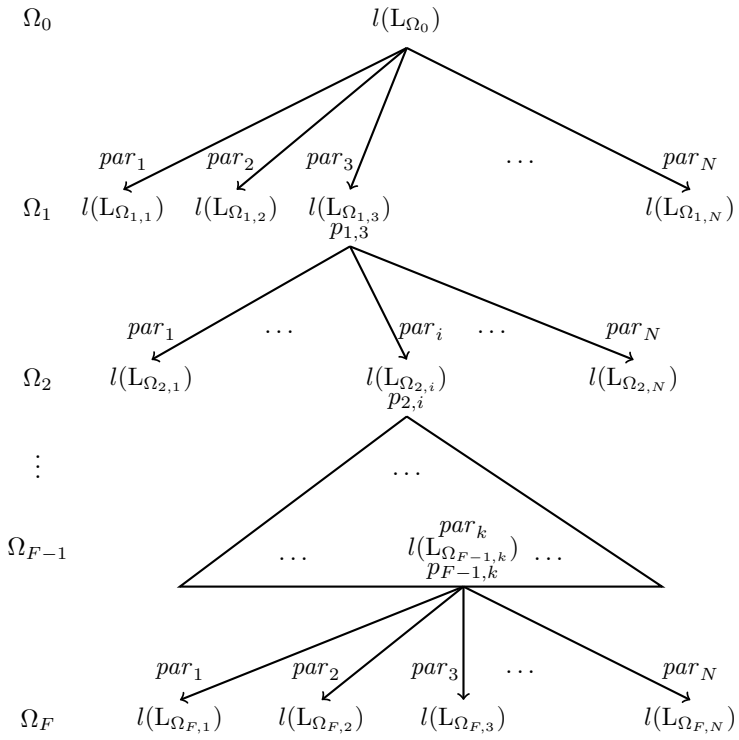


Figure 2.3: Forward selection procedure represented by a tree plot explaining the selection procedure. For each parameter the loglikelihood is calculated. Using Wilks likelihood ratio a statistical decision is made based on the p-value.

2.3.2 Maximum A Posteriori probability

Prior knowledge about parameters can be expressed by some prior probability density function, $p(\boldsymbol{\theta})$. In a Bayesian approach the prior information can be mixed with the information in the data. This gives a *Maximum A Posteriori,*

MAP, probability.

Using Bayes Theorem [36] the posterior probability density function can be formed:

$$p(\boldsymbol{\theta}|\mathcal{Y}_N) = \frac{p(\mathcal{Y}_N|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathcal{Y})} \propto p(\mathcal{Y}_N|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{2.19}$$

where all parameters have been defined previously.

Equation (2.19) is reduced to the likelihood function when prior information is available, making maximum likelihood a special case of MAP.

If only a subset of parameters have a prior probability density function[15] MAP estimates can still be performed. Assuming the prior probability density of the parameters are Gaussian:

$$\boldsymbol{\mu}_{\boldsymbol{\theta}} = E\{\boldsymbol{\theta}\} \tag{2.20}$$
$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = V\{\boldsymbol{\theta}\} \tag{2.21}$$
$$\boldsymbol{\epsilon}_{\boldsymbol{\theta}} = \boldsymbol{\theta} - \boldsymbol{\mu}_{\boldsymbol{\theta}} \tag{2.22}$$

the posterior probability density function, when conditioning on $\boldsymbol{y}_0$ and taking the negative logarithm, can be written:

$$\begin{aligned} -\ln\left(L\left(\boldsymbol{\theta};\mathcal{Y}_N|\boldsymbol{y}_0\right)\right) =& \frac{1}{2}\sum_{k=1}^{N}\left(\ln\left(\det\left(\boldsymbol{R}_{k|k-1}\right)\right) + \epsilon_k^T\boldsymbol{R}_{k|k-1}^{-1}\epsilon_k\right) \\ &+ \frac{1}{2}\left(\left(\sum_{k=1}^{N} l + p\right)\right)\ln\left(2\pi\right) \\ &+ \frac{1}{2}\ln\left(\det\left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\right)\right) + \frac{1}{2}\boldsymbol{\epsilon}_{\boldsymbol{\theta}}^T\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\epsilon}_{\boldsymbol{\theta}} \end{aligned} \tag{2.23}$$

The estimates can now be found by solving a non-linear optimisation problem similar to Equation (2.15).

2.4 Extended Kalman Filter

The Kalman Filter gives the optimal[16] reconstruction and prediction of the state vector $\boldsymbol{x}_t$ given some observations of the input $\{\boldsymbol{u}_t\}$ and observations of the output $\{\boldsymbol{y}_j\}$ in the linear state space model given in Equation (2.2) [34].

[15] $p(\boldsymbol{\theta})$ partly uniform.
[16] Minimum variance

A common approach for non-linear models is based on linearisation of the functions around the nominal trajectory. This means that the linear Kalman filter is usable[17]. Using the current estimate of the state is a good choice for the linearisation trajectory.

A linearisation about the current state estimate at every sampling time, followed by applying the Kalman filter to the resulting model, gives the *Extended Kalman Filter, EKF* [35].

The EKF can be used to solve non-linear SDEs as Equation (2.4) numerically, since the likelihood principle gives the non-linear optimisation problem given by Equation (2.15). The EKF is an approximative method. It gives a near-optimal estimator as it is a linearisation of a non-linear model [35].

For a non-linear model, a given set of parameters $\boldsymbol{\theta}$ and initial states $\boldsymbol{x}_0$, the *one-step predictions*[18] of the covariance $\boldsymbol{R}_{k|k-1}$, Equation (2.12), and the error term $\boldsymbol{\epsilon}_k$, Equation (2.13) can be computed by the EKF. The EKF consists of steps divided in a prediction part and an updating part[19]. A schematic drawing of this is shown in Figure 2.4.

When the EKF has minimised Equation (2.15), the near-optimal parameter estimate $\hat{\boldsymbol{\theta}}$ has been identified.

2.5 Validation

When one model is achieved, a systematic examination of the model is made, inspired by the framework of Kristensen [29]. A schematic representation of the approach is seen in Figure 2.5.

In brief: *model (re)formulation* is the initial model structure and applied modifications. *Parameter estimation* is estimation of unknown parameters based on data and *Residual analysis* is the evaluation of the quality of the resulting model. In *model falsification or unfalsification* it is decided whether the model sufficiently serves the intended purpose. If the model is unfalsified the development terminates. If falsified the cycle must be repeated by reformulating the model. The *statistical tests* may give indications of deficient model parts and the *non-parametric modelling* provides indications of unmodelled structure.

[17] For a detailed description of the linear Kalman filter, see [34].
[18] A k-step prediction can be obtained by skipping the updating step.
[19] For each step see the corresponding Equation in [30].

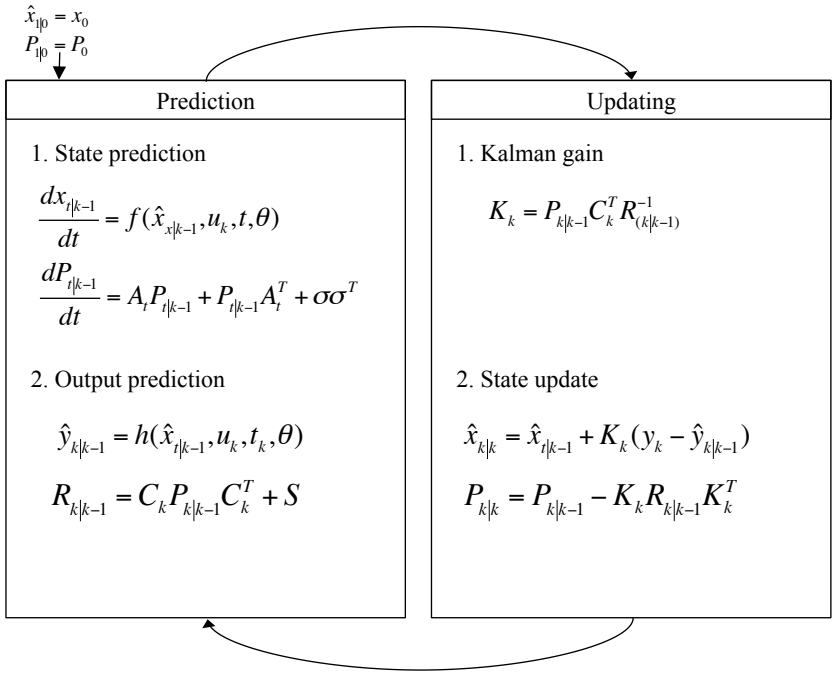


Figure 2.4: Schematic overview of the Extended Kalman Filter based on [59].

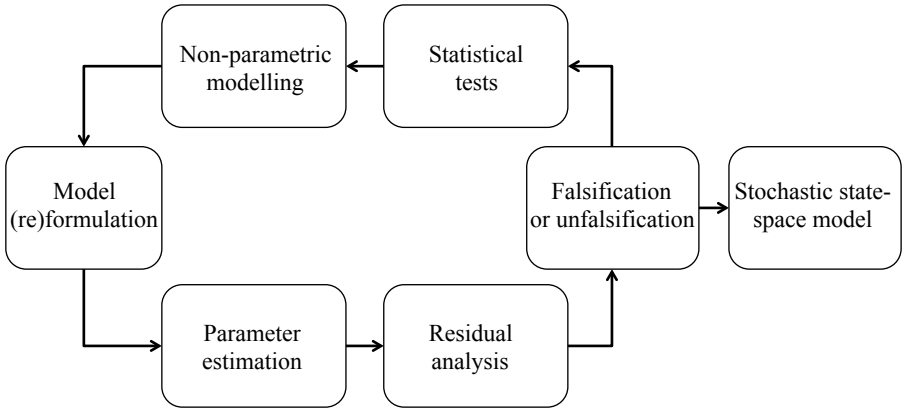


Figure 2.5: Framework for systematic improvement of stochastic grey-box models based on [29].

The overall goal in model building is to construct a model where the residuals resemble a discrete white noise process, so that the following holds:

Definition 2.7 (White noise) A process $\{\varepsilon\}$ is said to be *white noise*, if $\{\varepsilon\}$ is a sequence of mutually uncorrelated identically distributed random variables with mean value 0 and constant variance σ_ε^2. This implies that:

$$\mu_t = \mathbb{E}[\varepsilon_t] = 0 \tag{2.24}$$
$$\sigma_t^2 = \mathrm{Var}[\varepsilon_t] = \sigma_\varepsilon^2 \tag{2.25}$$
$$\gamma_\varepsilon = \mathrm{Cov}[\varepsilon_t, \varepsilon_{t+k}] = 0 \ , \ \text{for } k \neq 0 \tag{2.26}$$

□

If the residuals can be regarded as white noise then the model is adequate. There are several ways to examine the residuals: *autocorrelation function, ACF*, *(sample) autocorrelation function, SACF*, *partial autocorrelation function, PACF* [9, 36] and the *lag-dependence function, LDF* and *partial lag-dependence function, PLDF* introduced by Nielsen and Madsen in [39] for non-linear systems.

2.5.1 Statistical tests

By using the central limit theorem [47], the estimator in Equation (2.15) is asymptotically Gaussian with mean $\boldsymbol{\theta}$ and covariance [29]:

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}} = \boldsymbol{H}^{-1}$$

where $\boldsymbol{H}$ is the Hessian evaluated at the minimum of the objective function in Equation (2.15).

A measure of uncertainty of the individual parameter estimate is obtained by decomposing the covariance matrix as follows [39]

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}} = \boldsymbol{\sigma}_{\hat{\boldsymbol{\theta}}} \boldsymbol{R} \boldsymbol{\sigma}_{\hat{\boldsymbol{\theta}}}$$

where $\boldsymbol{\sigma}_{\hat{\boldsymbol{\theta}}}$ is the diagonal matrix of the standard deviations of the parameter estimates and $\boldsymbol{R}$ is the corresponding correlation matrix. The correlation matrix $\boldsymbol{R}$ can be used to indicate whether a parameter is redundant.

The asymptotic Gaussianity of the estimator in Equation (2.15) also allows marginal t-tests to be performed to test the hypothesis [29]:

Definition 2.8 (Parameter selection)

$$\mathcal{H}_0: \quad \theta_j = 0$$

against:

$$\mathcal{H}_1: \quad \theta_j \neq 0$$

to test whether a given parameter θ_j is marginally insignificant or not.

The test quantity is the parameter estimate θ_j divided by $\sigma_{\widehat{\theta}_j}$, and under $\mathcal{H}_0$ this quantity is asymptotically t-distributed with $N-p$ degrees of freedom where N is the number of data points and p is the number of parameters. This gives:

$$z(\widehat{\theta}_j) = \frac{\widehat{\theta}_j}{\sigma_{\widehat{\theta}_j}} \in t\,(N-p) \tag{2.27}$$

□

2.6 The CTSM-R tool

The tool used to estimate models, is the *Continuous Time Stochastic Modelling, CTSM*-R tool. This is an R-package[20] developed by Rune Juhl[21] based the original CTSM tool developed by Niels Rode Kristensen and Henrik Madsen at the Technical University of Denmark [30, 48, 23].

A detailed description of the CTSM-R tool will not be given, but a brief summery is necessary to conceptually understand the modelling approach.

Parameter estimation is carried out by maximising the likelihood function with an optimisation scheme, that requires initial values for states and parameters together with lower and upper bounds.

Values of a parameter can be estimated between lower and upper bounds or fixed to a constant value. Initial state values can likewise be estimated between boundaries or fixed to a value. While estimating the models, the parameters are transformed to an optimisation space, and when the optimum has been found, they are transformed back to the parameter space.

[20] http://www.ctsm.info/

[21] Technical University of Denmark - Department of Applied Mathematics and Computer Science

The tool supports statistical analyses: Estimated values, standard error and probability of the t-distribution are given as standard. Furthermore the derivative of the objective function with respect to each initial state or parameter and the derivative of the penalty function is given. Finally the correlation matrix of the parameters is available. If the off-diagonal of $\boldsymbol{R}$ has values $\geq |0.99|$, there is indication of over-parametrisation.

Observations are expressed as functions of states and inputs and the variance of the observation noise is also added to the system. The variance of the observation is provided and used for interpretation of the estimated model.

2.7 Summary

The main principles used in models of PK/PD systems and the statistical methods and tools used to validate the models have been introduced.

Most physical systems are appropriately described in continuous-time with discrete-time measurements. This is also the main assumption in this thesis, and the aim is to identify the relationship between the subcutaneous glucose levels and the plasma glucose levels. This is done by:

1. Using PK/PD modelling to identify a suitable model structure to the describe the relationship.
2. Formulating the model structure on state-space form using SDEs.
3. Identifying the parameters, $\boldsymbol{\theta}$, of the model, using the likelihood principle and Extended Kalman Filter.
4. Using methods and tools for statistical comparison (Section 2.3.1) it is possible to examine possible improvements.

The obtained parameters of the final model should, as grey-box modelling is used, then have a physical interpretation that easily is related to the true physiological parameters. The next chapter will explain the model used in the thesis and describe the parameters that are to be estimated.

Chapter 3

Modelling glucose-insulin dynamics

Glucose-insulin dynamics can be modelled as a physiological system. The foundation of these systems is often based on the Bergman Minimal Model [33]. Two systems extending the Bergman Minimal model are: the Medtronic Virtual Patient model [24] and the *Steil* model [54]. Both models are used in this thesis, and this chapter will give a summary of them.

The MVP model has been chosen because it is one of the state-of-the-art models for modelling glucose-insulin dynamics. The MVP model is based on ordinary differential equations. For the purpose of this thesis the MVP model will be formulated as SDEs on state-space form.

An alternative to the MVP model, the *Steil* model, will also be presented. The Steil model differs from the MVP model in the description of the subcutaneous glucose dynamics having a more physiologically correct interpretation. A detailed description of the Steil model with emphasis on the difference to the MVP model will be given. Finally the advantages and disadvantages of each model will be discussed.

3.1 Medtronic Virtual Patient model

The metabolic equations of the MVP model can be regarded as seven compartments - one compartment corresponding to each equation. The essence of the model is based on the Bergman Minimal Model [4, 33] with modifications by Fisher [14] replacing the insulin secretion term with an insulin infusion term. The seven equations describing the physiological system are:

$$\frac{dI_{sc}(t)}{dt} = -\frac{1}{\tau_1} \cdot I_{sc}(t) + \frac{1}{\tau_1} \cdot \frac{ID(t)}{C_I} \tag{3.1a}$$

$$\frac{dI_P(t)}{dt} = -\frac{1}{\tau_2} I_p(t) + \frac{1}{\tau_2} I_{sc}(t) \tag{3.1b}$$

$$\frac{dI_{Eff}(t)}{dt} = -p_2 \cdot I_{Eff}(t) + p_2 \cdot S_I \cdot I_P(t) \tag{3.1c}$$

$$\frac{dG_p(t)}{dt} = -(GEZI + I_{Eff}) \cdot G_p(t) + EGP + \frac{D_2}{\tau_{max}} + \frac{G_{iv}}{Vg} \tag{3.1d}$$

$$\frac{dG_{sc}(t)}{dt} = -\frac{1}{\tau_3} \cdot G_{sc}(t) + \frac{1}{\tau_3} \cdot G_P(t) \tag{3.1e}$$

$$\frac{dD_1(t)}{dt} = A_g \cdot CHO(t) - \frac{D_1}{\tau_{max}} \tag{3.1f}$$

$$\frac{dD_2(t)}{dt} = \frac{D_1}{\tau_{max}} - \frac{D_2}{\tau_{max}} \tag{3.1g}$$

The system is constructed so it captures the most essential dynamics. Each state has an elimination part and an absorption part. In the elimination part, the concentration is lowered in the compartment, and in the absorption part the concentration is increased in the compartment. An analysis of the influence of each compartment is given to explain the physical interpretation of each parameter. A short explanation of these equations will be given in the following sections.

3.1.1 Insulin dynamics

Equations (3.1a) to (3.1c) relate to insulin. Combined, these equations describe the dynamic of insulin from entering the body in the subcutaneous tissue until an effect is produced in the plasma.

The states in these three equations are:

State	Unit	Description
I_{sc}	IU/L	Amount of insulin in the subcutaneous tissue
I_P	IU/L	Amount of insulin in the plasma
I_{Eff}	1/min	Effect of insulin on glucose transportation

The input is described by:

Input	Unit	Description
ID	IU/min	Insulin delivery to the subcutaneous tissue

The parameters are:

Parameter	Unit	Description
C_I	L/min	Clearance of insulin in the subcutaneous tissue
τ_1	min	Time constant associated with insulin movement in the subcutaneous tissue
τ_2	min	Time constant associated with insulin movement in plasma
p_2	1/min	p_2 is the delay in insulin action following an increase in plasma insulin
S_I	L/IU/min	Insulin sensitivity

3.1.2 Glucose dynamics

Equations (3.1d) and (3.1e) describe the dynamics of glucose entering the plasma and the subcutaneous tissue.

The states in these two equations are:

State	Unit	Description
G_P	mmol/L	Amount of glucose in the plasma
G_{sc}	mmol/L	Amount of glucose in the subcutaneous tissue

The input is described by:

Input	Unit	Description
G_{iv}	mg/min	Glucose injected directly into the plasma in case of too low plasma glucose

The parameters are:

Parameter	Unit	Description
$GEZI$	1/min	Effect of glucose per se to increase glucose uptake into cells and lower endogenous glucose production at zero insulin
EGP	mmol/L	Endogenous glucose production rate, estimated at zero insulin
Vg	L kg	Distribution volume of glucose
τ_3	min	Time constant associated with glucose movement in the subcutaneous tissue

3.1.3 Meal

Equations (3.1f) and (3.1g) model the dynamics in relation to the meal. When a meal is consumed, the absorption of glucose from the meal is modelled by two compartments.

The states in these two equations are:

State	Unit	Description
D_1	mg	Carbohydrate absorption
D_2	mg	A delay compartment of carbohydrate absorption

The input is described by:

Input	Unit	Description
CHO	mg/min	Digested carbohydrate

The parameters are:

Parameter	Unit	Description
Ag	unitless	Carbohydrate bioavailability
τ_{max}	min	Time constant associated with carbohydrate movement

Kanderian investigated glucose-insulin dynamics in MATLAB[22] using ODEs, estimating parameters in four steps [24] . First insulin related parameters were identified. Using these estimates, parameters in Equations (3.1c) and (3.1d) were identified. Then model fits were evaluated, and parameter re-estimation was conducted if necessary. With all parameters identified the model was re-

[22]Mathwork Inc., Natick, MA.

constructed introducing Equation (3.1e) to describe the relationship between plasma glucose and subcutaneous glucose levels. To asses model parameters, fits were calculated and interpreted by visual inspection.

3.1.4 Stochastic formulation

To accurately model the system and get a complete description of the data, the stochastic influence is essential to the model. The MVP model is therefore, on the basis of the theory from Chapter 2, written on stochastic state-space form as Equation (2.4):

Final model

State equations

$$\begin{bmatrix} dI_{sc} \\ dI_P \\ dI_{Eff} \\ dG_P \\ dG_{sc} \\ dD_1 \\ dD_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\tau_1} \cdot I_{sc}(t) + \frac{1}{\tau_1} \cdot \frac{ID(t)}{C_I} \\ -\frac{1}{\tau_2} \cdot I_P(t) + \frac{1}{\tau_2} \cdot I_{sc} \\ -p_2 \cdot I_{Eff}(t) + p_2 \cdot S_I \cdot I_P(t) \\ -(GEZI + I_{Eff}) \cdot G_P(t) + EGP + \frac{D_2}{\tau_{max}} + \frac{G_{iv}}{Vg} \\ -\frac{1}{\tau_3} \cdot G_{sc}(t) + \frac{1}{\tau_3} \cdot G_P(t) \\ A_g \cdot CHO(t) - \frac{D_1}{\tau_{max}} \\ \frac{D_1}{\tau_{max}} - \frac{D_2}{\tau_{max}} \end{bmatrix} dt + \boldsymbol{\sigma} \begin{bmatrix} d\omega_1 \\ d\omega_2 \\ d\omega_3 \\ d\omega_4 \\ d\omega_5 \\ d\omega_6 \\ d\omega_7 \end{bmatrix} \quad (3.2)$$

Inputs

$$ID(t)$$
$$CHO(t)$$
$$G_{iv}$$

Output

$$YSI = G(t) + sYSI$$
$$CGM = G_{sc}(t) + sCGM$$

where $sYSI$ and $sCGM$ are the measurements noises of the YSI device and CGM device, respectively.

The final model formulation, Equation 3.2, can model pharmacokinetics and pharmacodynamics in relation to glucose-insulin dynamics. It can be adjusted

to data not containing any meals by letting $D_1 = D_2 = 0$.

During the thesis different parameter estimates will be presented[23]. Two parameters are considered repeatedly throughout the thesis. This is the measurement noise of the observations: $sYSI$ for the measurement noise of the YSI device and $sCGM$ for the measurement noise of the CGM device. These two parameters give an indication of the physiological validity of the estimates.

3.2 Steil's model

The Steil model [54] is, like the MVP model, based on the Bergman Minimal Model. The difference lies in the compartments describing glucose dynamics. Steil's model has a more accurate physiological description of insulin's influence on plasma and subcutaneous glucose. It describes the movement of glucose between plasma and the subcutaneous tissue by diffusion based on gradients. Insulin has its effect in the subcutaneous tissue where it facilitates the transportation of subcutaneous glucose into the insulin dependent tissue. According to Pickup and Williams [45], this simplified description is widely accepted as a useful approximation to the actual dynamics.

Compared to the MVP model only two equations need to be altered to accommodate Steil's model. These two alterations are:

Alteration of Equation (3.2) to Steil's model:

$$\begin{bmatrix} dG_p \\ dG_{sc} \end{bmatrix} = \begin{bmatrix} -\left(GEZI + k_{21} \cdot \frac{V_1}{V_2}\right) \cdot G_p(t) + EGP + k_{12} \cdot G_{sc} \\ \left(k_{21} \cdot \frac{V_1}{V_2}\right) \cdot G_p - (k_{12} + k_{02} \cdot I_{Eff}) \cdot G_{sc} \end{bmatrix} dt + \boldsymbol{\sigma} \begin{bmatrix} d\omega_4 \\ d\omega_5 \end{bmatrix} \tag{3.3}$$

where k_{21} is the rate constant from the plasma to the subcutaneous tissue, k_{12} is the rate constant from the subcutaneous tissue to the plasma, k_{02} is the rate constant from the subcutaneous tissue into the cell, V_1 and V_2 are the volumes of plasma and subcutaneous tissue, respectively. The remaining parameters are as before. Effects of meal and glucose given intravenous can also be added.

A more correct physiological description of the system seems intuitively as a good idea. The closer the model is to the correct physiological description the

[23] The entire parameter estimate table will be found in Appendix A, and only relevant estimates are shown within the main text.

better. There are however some potential problems with the more accurate description of the physiology. These problems relate to the parameters, the volumes and oversimplification:

Parameters
An increase in complexity of the model will result in an increase in the number of parameters to be estimated. With a limited amount of data this may present a problem.

Volumes
V_1 and V_2 describe the volumes of plasma and subcutaneous tissue, respectively. These values are highly inter-dependable and very difficult to estimate based on physical attributes. In the model only the ratio $\frac{V_1}{V_2}$ is represented. A solution is to fix the values according to knowledge of the physical attributes.

Oversimplification
The physiological more correct model of the glucose dynamics is still a very simplified approximation of the true physiological relationships [53, 45, 43]. The actual dynamic is far more complicated and, probably, not completely known. It is, therefore a risk, that a slightly more complicated model does not give any improvement.

3.3 Summary

The fundamental concept in glucose-insulin dynamics is described by the Bergman Minimal Model. There are not many alternatives to this fundamental approach. This thesis aims to investigate the glucose-insulin dynamics. This is done in both the MVP model and the Steil model. Based on an initial analysis of both models, one will be selected for an in-depth investigation.

In the MVP model, insulin facilitates the transportation of glucose from the plasma to the subcutaneous tissue. This is a coarse-grained approximation to the real physiological process. In the subcutaneous tissue glucose is then transported the insulin dependent cells based on a rate constant. The advantage of the MVP model is that it is simple and at the same time physiologically meaningful.

In the Steil model, insulin facilitates the transportation of glucose from the subcutaneous tissue to the insulin dependent cell. The glucose transportation between plasma and subcutaneous tissue is governed by rate constants based

on gradients. This is physiologically more correct, but it is not obvious whether this more complex model offer any advantages. This will be examined in a later chapter.

Chapter 4

Introduction to analyses

This short chapter is an introduction to the following five chapters presenting the different models discovered and the analysis hereof. The results illustrate the iterative progression. Not all steps are presented or accounted for in detail, but it should be possible to track the progression.

Each iterative progression has, when modelled, undergone a residual analysis. This residual analysis is tedious and, in the early steps, not very informative. Therefore residual analysis is only presented in the main text to substantiate arguments. Parameter estimate tables are, were relevant, presented in Appendix A, and relevant estimates are emphasised in the main text. Correlation tables have been examined for every model - throughout these examinations the correlations in the off-diagonal for the presented models did not exceed 0.96, and the correlation tables are therefore not presented.

Throughout the chapters, figures of mainly the predictions of the YSI observations and CGM observations are presented. YSI observations are red, CGM observations are blue, prediction lines are black, and the 95%-confidence interval is grey. The green ribbon indicates the ideal range of glucose level determined as in [53]. The residuals, ACF, PACF and cummulative periodogram are also presented. When other figures are presented an elaboration is given.

Extensions and modifications of the model structure were implemented. Exten-

sions are presented in a framed box and are alterations of the system represented in Equation (3.2). At the end of each examined model, the likelihood, *degrees of freedom, DF* and AIC are presented to maintain a sound statistical examination as explained in Chapter 2.

4.1 Patient selection

The aim of this thesis was to develop a model suitable for low-exciting glucose-insulin dynamics. As explained in Chapter 1, one patient with a characteristic resting glucose dynamics, Figure 1.4, was selected for the entire investigation.

This patient was selected because he was deemed to have a typical glucose dynamic for a diabetes patient at rest. This patient was not given corrective carbohydrates. This made the patient attractive for an initial investigation since the infusion of corrective carbohydrates gives a non-physiological rise in glucose values. This would complicate the dynamics further. Having identified the insulin-glucose dynamics in a simple setting expansions can be made later.

4.2 Outline

The results are divided into four different chapters containing the following:

In Chapter 5 an investigation of the identifiability of the data is conducted. This investigation leads to an examination of CTSM-R tool and the calculations performed within this tool.

Chapter 6 investigates two different models:

- Model 1: The initial MVP model formulated as Equation (3.2).
- Model 2: A simplification of Model 1 correcting insulin kinetics.

Chapter 7 examines an alternative model:

- Model 3: The Steil model as formulated in Equation (3.3).

The Steil model is compared to the MVP model to see if a more complex model gives improved estimations. Thereby this chapter lays the foundation for how the system is extended.

Chapter 8 focuses on glucose dynamics and the CGM device. Four different models are presented:

- Model 4: An extension of Model 2 with a Bayesian approach.
- Model 5: An extension of Model 2 using *AutoRegressive, AR(p)*-processes.
  - Model 5a: MVP + Random Walk
  - Model 5b: MVP + AR(1)
- Model 6: An extension of Model 2 to include MAP and an autoregressive process.
  - Model 6a: MVP + MAP + Random Walk
  - Model 6b: Model 6a on the alternate CGM device
- Model 7: An extension of Model 6 to include two CGM devices.
  - Model 7a: MVP + MAP and both CGM devices
  - Model 7b: MVP + MAP + AR(1) and both CGM devices

Chapter 9 gives a summary and a discussion of all results obtained in this thesis. Analyses and explanations of significant discoveries and abnormalities will be given here.

Chapter 5

Identifiability

The first part of this chapter examines the identifiability of the data. Identifiability describes the prerequisites for constructing a model based on the available data and the parameters of the model. The investigation of identifiability lead to a discovery of a computational issue in the CTSM-R tool. The issue is related to the initial covariance matrix used in the calculation of the state prediction. The chapter describes this issue and a solution is presented.

The general theory in this chapter is based on Madsen [35] and Rudin [51]. The mathematical examination of the CTSM-R tool is based on Kristensen and Madsen [30].

Modelling short time series with little excitation has many challenges. One is the *identifiability* of the time series. Identifiability can be divided into two categories 1) *Structural identifiability* and 2) *Persistence of excitations.*

In general, any transfer-function model corresponds to a continuum of state-space models, so some structure on the state-space representation, to provide a unique relation between the parameters[24], must be imposed.

[24] For more details, refer to [35].

5.1 Structural identifiability

Structural identifiability can be defined as [35]:

Definition 5.1 (Structural identifiability) The parameter θ_i is said to be *structurally globally identifiable*, if for almost any $\boldsymbol{\theta}^*$

$$M(\boldsymbol{\theta}) = M(\boldsymbol{\theta}^*) \Rightarrow \theta_i = \theta_i^* \tag{5.1}$$

It is *structurally locally identifiable*, if for almost any $\boldsymbol{\theta}^*$ there exists a neighbourhood $v(\boldsymbol{\theta}^*)$ such that if $\boldsymbol{\theta} \in v(\boldsymbol{\theta}^*)$, then the above equation is true.

If all model parameters are globally identifiable, the model is said to be *structurally identifiable.* □

Remark 5.1 *In this thesis, $\boldsymbol{\theta} \in \mathbb{R}^n$. The concept of almost surely is measure theoretic in nature and formally expresses so [51, p. 27]:*

$$\lambda(\{\boldsymbol{\theta}^* \in \boldsymbol{\Theta} | M(\boldsymbol{\theta}) = M(\boldsymbol{\theta}^*) \wedge \theta_i \neq \theta_i^*\}) = 0$$

where $\boldsymbol{\Theta} \subseteq \mathbb{R}^n$ is chosen appropriately and λ represents Lebesgue measure. In Definition 5.1, it is assumed that $M : \mathbb{R}^n \to \mathbb{R}^m$ is Borel-measurable for some $m \in \mathbb{N}$.

Remark 5.2 *Think of M as a diffusion process, X_t. The definition above formalises the idea that if two models, in this case SDE models, attain the same values (same behaviour) then the parameter tuple is unique. Intuitively, this is the goal.*

If, for instance, the model at hand is

$$y_t = (p_1 + p_2)u_t + \epsilon_t$$

then it is only possible to estimate $(p_1 + p_2)$ unless physical knowledge, e.g. that p_1 and p_2 satisfy a relationship of the form $c_1 p_1 + c_2 p_2 = 0$, is at hand. With this physical knowledge, the model would be structurally identifiable.

This concept is relevant for both linear and non-linear models. For non-linear models two kinds of non-linearities are relevant; non-linear in the inputs and non-linear in the parameters[25]. Models that are non-linear in the parameters can have structurally locally identifiable parameters which are not the same as the structurally globally identifiable parameters [35].

[25] Refer to [35] for a detailed description.

5.2 Persistence of excitation

Persistence of excitations is a property of the time series [35]:

Definition 5.2 (Persistence of excitation) An input signal u satisfies the condition of *persistent excitations* of order n if the Yule-Walker matrix of the autocorrelations from lag 0 until lag $n-1$ is positive definite. □

An intuitive approach is an example were the following model is considered:

$$y_t = p_1 u_t + p_2 u_{t-1} + \epsilon_t$$

This model is structurally identifiable. If a constant (and trivial) time series given by $u_t = u_{t-1}$ is modelled, then the above equation is on the form

$$y_t = (p_1 + p_2) u_t + \epsilon_t$$

so even though the model is structurally identifiable, this input signal is not persistent exciting. Given trivial data, it is not possible to distinguish between models.

5.3 Transfer-functions and identifiability

For a transfer-functions model, a unique relation between the parameters of the model and its state-space representation must exist, to guarantee that it is structurally identifiable. From an engineering point-of-view, a continuous-time state-space model can be written on the form [35]:

$$\begin{aligned} \dot{\boldsymbol{X}}(t) &= \boldsymbol{A}\boldsymbol{X}(t) + \boldsymbol{B}\boldsymbol{U}(t) + \boldsymbol{e}_1(t) \\ \boldsymbol{Y}(t) &= \boldsymbol{C}\boldsymbol{X}(t) + \boldsymbol{D}\boldsymbol{U}(t) + \boldsymbol{e}_2(t) \end{aligned}$$

This may be written in innovation form[26] with p introduced as a difference operator. The transfer function model takes the form:

$$\boldsymbol{Y}(t) = \boldsymbol{G}(p)\boldsymbol{U}(t) + \boldsymbol{H}(p)\boldsymbol{e}(t)$$

where the transfer function is identified as:

$$\boldsymbol{G}(p) = \boldsymbol{C}(p\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D} \tag{5.2}$$

[26] See [35] for details.

It is now possible to observe the transfer matrix $\boldsymbol{G}(p)$ with elements of the form:

$$\boldsymbol{G}_{ij}(p) = \frac{b_u p^u + b_{u-1} p^{u-1} + \ldots + b_0}{a_v p^v + a_{v-1} p^{v-1} + \ldots + a_0} \tag{5.3}$$

where $a_v = 1$. When the system of equations is identified, the model is structurally identifiable if the system has a unique solution.

During initial investigations identifiability problems were found. The initial investigation was performed with a negligible diffusion term - following tradition this negligible value was set to $\exp(-12)$. Adjusting this negligible diffusion term between $\exp(-10)$ and $\exp(-18)$, gave different parameter estimates. This was unexpected since the diffusion term was believed to be negligible. A further investigation, where all but two parameters were fixed, revealed the problem. Figure 5.1 shows a dramatic change in the optimisation space when the diffusion term is altered[27].

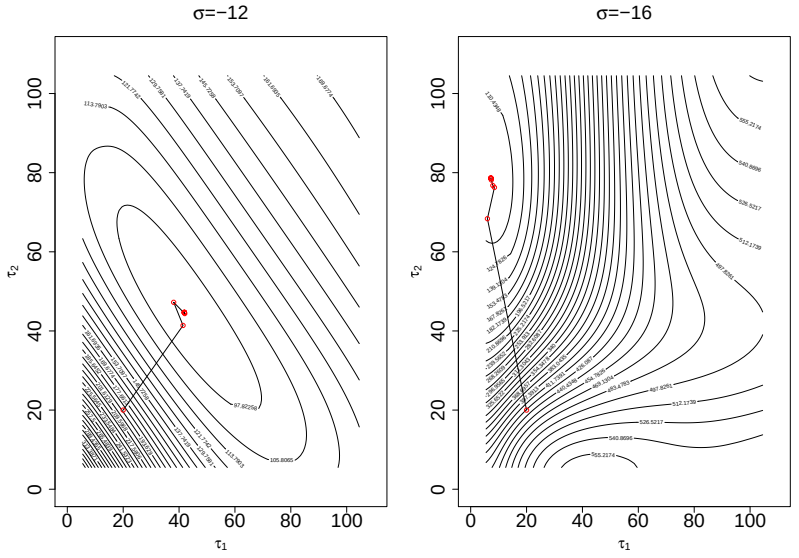


Figure 5.1: Contour plot visualising optimisation space of parameters, τ_1 and τ_2, with 2 different diffusion terms $\exp(-12)$ and $\exp(-16)$, respectively. Red circles represent each step in the optimisation procedure.

In the two-dimensional case it is straightforward to examine the theoretical structural identifiability by using transfer functions, here τ_1 and τ_2:

[27] While not being a proof, this result visualises the problem well. If the two-dimensional optimisation space is altered, it is expected that an n-dimensional space will likewise be altered.

The two-dimensional system can be written on matrix form as:

$$A = \begin{bmatrix} -\frac{1}{\tau_1} & 0 \\ \frac{1}{\tau_2} & -\frac{1}{\tau_2} \end{bmatrix}, B = \begin{bmatrix} \frac{1}{\tau_1} \\ 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 1 \end{bmatrix}, pI = \begin{bmatrix} p & 0 \\ 0 & p \end{bmatrix} \tag{5.4}$$

Using Equation (5.2) the following transfer-function is obtained with:

$$G(p) = \frac{1}{p^2 + p(\frac{1}{\tau_1} + \frac{1}{\tau_2}) + \frac{1}{\tau_1 \tau_2}} \tag{5.5}$$

This has to be compared with what is actually observable - Equation (5.3):

$$G_{ij} = \frac{1}{p^2 + a_1 p + a_0} \tag{5.6}$$

So the system is structurally identifiable if the following has one unique solution

$$\begin{aligned} a_1 &= \frac{1}{\tau_1} + \frac{1}{\tau_2} \\ a_0 &= \frac{1}{\tau_1 \tau_2} \end{aligned} \tag{5.7}$$

This clearly holds.

5.4 Investigation of the CTSM-R tool

This section examines the initial covariance matrix, $\boldsymbol{P}_0$, used in the calculation of the state prediction in the EKF. Though structural identifiability has been shown, the results are not consistent. A possible explanation may be the calculation of the initial covariance matrix:

$$\boldsymbol{P}_0 = P_s \int_{t_0}^{t_1} e^{\boldsymbol{A}s} \boldsymbol{\sigma}\boldsymbol{\sigma}^T \left(e^{\boldsymbol{A}s}\right)^T ds \tag{5.8}$$

where $P_s \geq 1$ is a pre-specified scaling factor, $\boldsymbol{A}$ is a non-linear function given by: $\boldsymbol{A} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}_t}|_{\boldsymbol{x}=\widehat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{u}=\boldsymbol{u}_k, t=t_k, \boldsymbol{\theta}}$, $\boldsymbol{\sigma}$ is a non-linear function given by $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{u}_k, t_k, \boldsymbol{\theta})$. $\boldsymbol{\sigma}$ is also the scaling of the diffusion terms in the stochastic state-space representation of the model.

The ODE-solution to the state prediction in the non-linear[28] [30, p. 17] case is given by:

$$\frac{d\widehat{\boldsymbol{x}}_{t|k}}{dt} = \boldsymbol{f}\left(\widehat{\boldsymbol{x}}_{t|k}, \boldsymbol{u}_t, t, \boldsymbol{\theta}\right), t \in [t_k, t_{k+1}[\tag{5.9}$$

$$\frac{\boldsymbol{P}_{t|k}}{dt} = \boldsymbol{A}\boldsymbol{P}_{t|k} + \boldsymbol{P}_{t|k}\boldsymbol{A}^T + \boldsymbol{\sigma}\boldsymbol{\sigma}^T t \in [t_k, t_{k+1}[\tag{5.10}$$

with $\boldsymbol{A}, \boldsymbol{\sigma}$ and $\boldsymbol{P}$ defined previously.

In an ideal ODE-solution the covariance matrix $\boldsymbol{P} = 0$. Solving Equation (5.10), with a small initial covariance matrix from Equation (5.8), may yield the first two terms to zero. This would mean that the state prediction for the covariance matrix is highly depending on the last term, the scaling of the diffusion term $\boldsymbol{\sigma}$. A large relative change, from $\boldsymbol{\sigma} = \exp(-12)$ to $\boldsymbol{\sigma} = \exp(-16)$, may be sufficient to influence state prediction, and would explain Figure 5.1.

The main problem using Equation (5.8) as initial covariance may be the domain of integration. It is not clear what a sufficiently large step between t_0 and t_1 should be. For long time series such as electrical market data, the step will be large and it will only have used a small subset of available data. In PK/PD modelling, time series are short. If a large subset of data is used to estimate the initial covariance matrix, $\boldsymbol{P_0}$, then the time until the EKF converges may be long. For shorter time series this gives unreliable results.

One way to accommodate this problem would be to set:

$$\boldsymbol{P}_0 = k \cdot \boldsymbol{I} \tag{5.11}$$

where k is the variance and $\boldsymbol{I}$ is the identity matrix.

Using Equation (5.11) the initialisation of the covariance matrix is always identical over the entire parameter space, as the scaling of the diffusion term, $\boldsymbol{\sigma}$, does not influence $\boldsymbol{P}_0$. An initial covariance matrix like Equation (5.11) is known to be erroneous due to the property of the identity matrix, so it presupposes that the EKF, using only few iterations, will have found a suitable structure of the covariance matrix $\boldsymbol{P}$.

Implementing Equation (5.11) and resolving the problem yields consistency. This is shown by the parameter estimates in Table 5.1.[29]

[28] In CTSM this implementation is based on the the algorithms of Hindmarsh [20]

[29] It is unfortunately not possible to make contour plots with the new initial covariance matrix in the tool CTSM-R.

Parameter	exp(−12)	exp(−16)
τ_1	6.73	6.73
τ_2	152.59	152.59

Table 5.1: Parameter estimates visualising an identical optimisation space for the parameters τ_1 and τ_2, with 2 different diffusion terms exp(−12) and exp(−16), respectively.

5.5 Summary of identifiability

Identifiability is always important when modelling time series. During the investigation of the identifiability, a theoretical computational issue in the CTSM-R tool was discovered. The short duration and lack of excitation of the recorded time series, revealed a problem with the calculation of the initial covariance matrix, $\boldsymbol{P}_0$. This problem does not appear when longer and more excited data is used. Before discovering that the initial covariance matrix caused the problem, many alternative theories where investigated. These theories will not be elaborated, but include ideas such as transforming rate constants to a logarithmic domain, to avoid the possibility for time constant equal to zero.

The next chapters will use the initial covariance matrix shown in Equation (5.11).

Chapter 6

Modelling using Medtronic Virtual Patient

Estimation using SDEs with negligible diffusion terms should yield results similar to [24], since SDEs with small fixed diffusion terms should resemble ODEs. Two different models are examined and extended in this chapter: Model 1 - the stochastic extension of the MVP model and Model 2 - a simplification of Model 1.

6.1 Model 1 - Original model structure

The predictions obtained with the SDEs having negligible diffusion terms, Figure 6.1, behaves as expected, showing ODE-like behaviour when predicting the YSI observations and CGM observations. It is clearly seen that the prediction of the YSI observation is partly higher and partly lower than the observed values. The insulin prediction, Figure C.1, is in concordance with the results presented in [24].

Figure B.1 shows the residual analysis pointing towards an AR(1)-process. However, it is necessary to consider the stochastic influence; the five different equation each have the possibility to estimate a stochastic diffusion term.

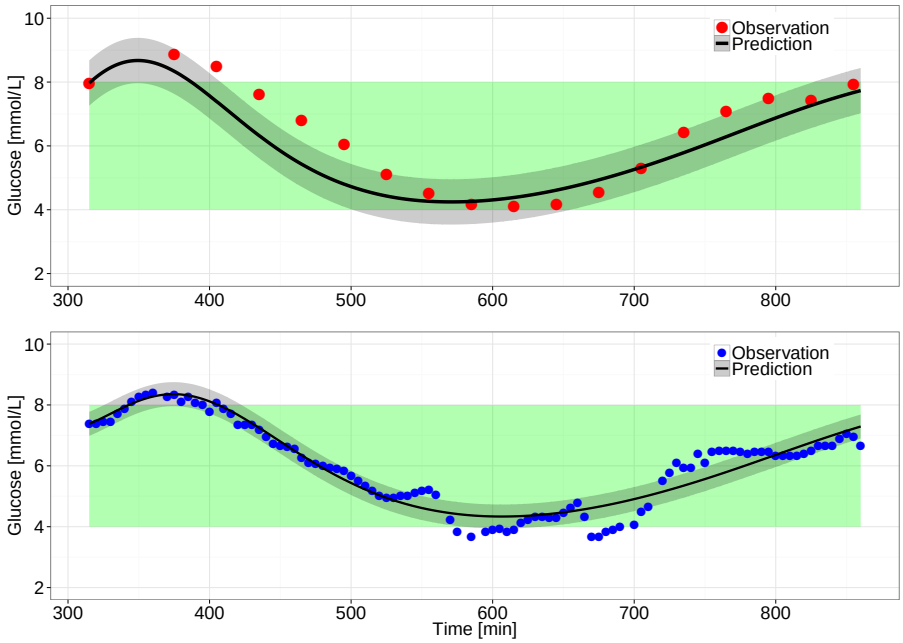


Figure 6.1: Model 1: Predictions using SDEs with negligible diffusion terms. ODE like structure is observed.
TOP PLOT) Prediction of YSI observations.
BOTTOM PLOT) Prediction of CGM observations.

The forward selection method presented in Figure 2.3 is applied. The selection process extending the system is shown in Figure 6.2.

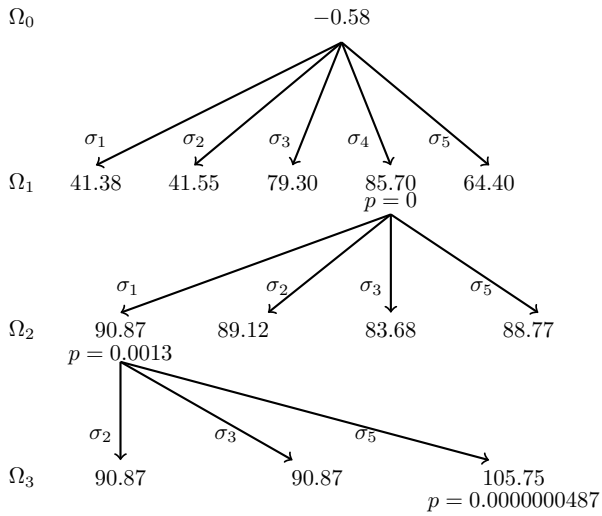


Figure 6.2: Forward selection of diffusion terms. Diffusion terms: σ_1, σ_2 and σ_5 were included in the order shown in the tree plot. Calculated loglikelihoods and p-value shown if $p < \alpha < 0.05$.

The model has an estimable diffusion term included in the state equations for: $\frac{dI_{sc}}{dt}$, $\frac{dG_p}{dt}$ and $\frac{dG_{sc}}{dt}$. Considering the glucose-insulin system presented in Equation (3.2) this seems physiologically reasonable since:

1. The state equations for $\frac{dI_{sc}}{dt}$ and $\frac{dI_P}{dt}$ relate to insulin dynamic. Insulin dynamic may be influenced by a coherence between the two compartments sufficiently described by one stochastic diffusion term.
2. The state equation for $\frac{dI_{EFF}}{dt}$, describes a, somewhat arbitrary, effect of insulin. All biological variation lies in either insulin dynamics or glucose dynamics, deeming this stochastic diffusion term redundant.
3. The state equations for $\frac{dG_P}{dt}$ and $\frac{dG_{sc}}{dt}$ both have a diffusion term. From a physiological view, there are several factors with considerable uncertainty linked to glucose dynamic and the CGM device. It is evident that the proposed dynamic is a simplification, and the behaviour of the CGM device is, to an extent, unknown.

6.1.1 Model 1 with estimation of σ_1, σ_4 and σ_5

Examining the predictions of the YSI observations and CGM observations in Figure 6.3, insulin prediction in Figure C.2, and the residual analysis in Figure B.2, it seems, that all requirements for model validation have been met.

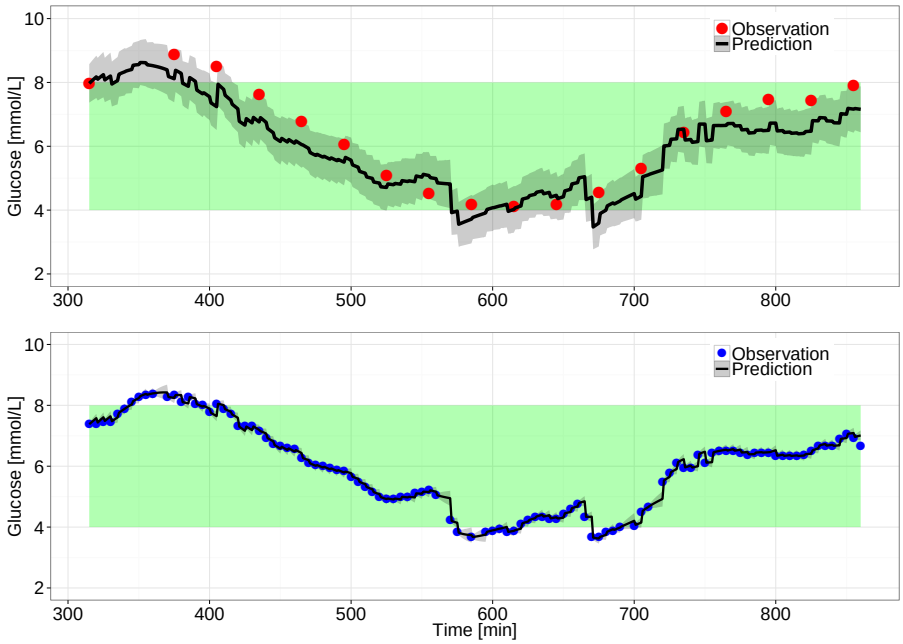


Figure 6.3: Model 1: Predictions using Equation (3.2) with estimation of diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$. Updates in the EKF are observed in both plots.
TOP PLOT) Prediction of YSI observations. Predictions seem to be off by a constant and having wide 95% confidence intervals.
BOTTOM PLOT) Prediction CGM observations. Predictions are perfect with a narrow 95% confidence interval.

This is nonetheless a flawed conclusion. The grey-box modelling approach allows a physical interpretation and it is important to consider the parameters and the information given in the estimates. Table A.1 presents the estimates.

The two main considerations in the physical interpretation of the parameter estimates is of the measurement noise $sYSI$ and $sCGM$.

The estimates indicate that the CGM device measurement noise, $sCGM$ =

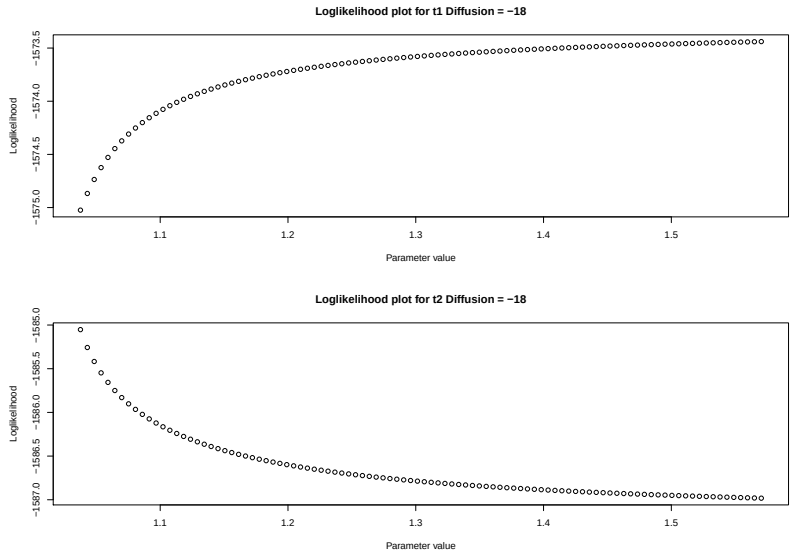


Figure 6.4: Visualisation of the calculated likelihood obtained with two rate constants describing insulin movement. An inverse relationship between τ_1 and τ_2 where $\tau = \ln \frac{1}{\phi}$ to avoid division with zero.
TOP PLOT) Loglikelihood of τ_1.
BOTTOM PLOT) Loglikelihood of τ_2.

$\exp(-16.11)$, is a precise measurement and the YSI measurement noise, $sYSI = \exp(-0.56)$, is less precise, resulting in a wide confidence interval of the YSI prediction and a narrow confidence interval for the CGM observation prediction.

The investigation in Chapter 5, revealed another deficiency. An inverse relationship between τ_1 and τ_2, describing insulin kinetics was found. It was believed that a division by zero in the estimations of the parameters caused the calulation of the likelihood to go towards zero and infinity, respectively. A transformation to $\tau = \ln \frac{1}{\phi}$ was tried to avoid division by zero. This did however not solve the problem as the visualisation in Figure 6.4 shows.

For Model 1 the likelihood, degrees of freedom and AIC is calculated to:

$$\begin{aligned} -\ln(L) &= 94.45 \\ DF &= 13 \\ AIC &= -98.01 \end{aligned}$$

6.2 Model 2 - Simplification of the MVP model

The inverse relationship, indicated previously, can be interpreted as: insulin flow through one compartment is very fast and through the other compartment very slow. A solution to this is to model the rate constants (τ_1 and τ_2) between the two compartments as one (τ_1):

Simplification of Equation (3.2):

$$\begin{bmatrix} dI_{sc} \\ dI_P \end{bmatrix} = \begin{bmatrix} -\frac{1}{\tau_1} \cdot I_{sc}(t) + \frac{1}{\tau_1} \cdot \frac{BA \cdot ID(t)}{C_I} \\ -\frac{1}{\tau_1} \cdot I_P(t) + \frac{1}{\tau_1} \cdot I_{sc} \end{bmatrix} dt + \boldsymbol{\sigma} \begin{bmatrix} d\omega_1 \\ d\omega_2 \end{bmatrix} \tag{6.1}$$

where $BA = 0.7$ is bioavailability of insulin, and remaining parameters have been defined previously.

From Equation 6.1 it is evident that the system is simplified by describing insulins route from the pump to the plasma by one time constant rather than two. Bioavailability allows for an adjustment of how much of the provided input actually is available[30]. The diffusion terms of the new model is again chosen by the forward selection approach. The inclusion tree is presented in Figure 6.5, and it is seen that the model structure is identical to the model structure obtained in Model 1[31].

The model with included diffusion terms gives the predictions shown in Figure 6.6 with the corresponding residuals shown in Figure B.3. If only the residuals are considered the identified model looks valid. But when the prediction and the parameter estimates are considered, there is still room for improvement. The parameter estimates are seen in Table A.2. It is noted that, as before, $sCGM = \exp(-16.14) \pm \exp(98.27)$[32] is low, reflected in the prediction of the CGM observations. The wider confidence interval on the YSI observation prediction comes from a larger measurement noise, $sYSI = \exp(-1.02) \pm \exp(0.40)$. Furthermore it seems that the YSI observation prediction is off by a constant. The estimation of the diffusion terms, $\sigma_4 = \exp(-2.03) \pm \exp(0.29)$ and $\sigma_5 = \exp(-2.68) \pm \exp(0.14)$ indicates that there is an undetected model structure in the pharmacodynamic part, namely in the state equations for $\frac{dG_p}{dt}$ and $\frac{dG_{sc}}{dt}$.

[30] $BA = 0.7$ making the last 30% are unavailable to the system.

[31] The predictions of the model using the SDE with negligible diffusion terms is very similar to the previous, so it is not presented.

[32] Insignificant in the estimation.

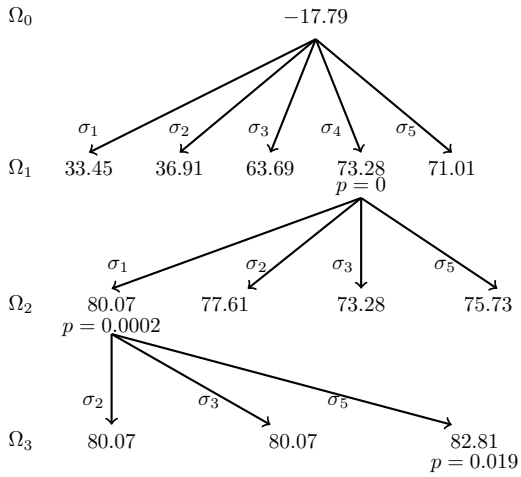


Figure 6.5: Forward selection of diffusion terms. Diffusion terms: σ_1, σ_4 and σ_5 were included in the order shown in the tree plot.

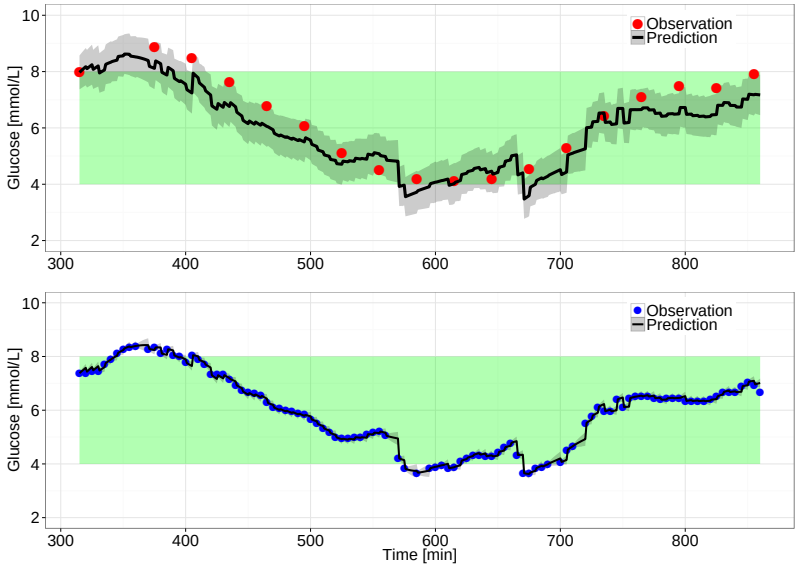


Figure 6.6: Model 2: Predictions using Equation (6.1) with estimation of diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$.
TOP PLOT) Prediction of YSI observations.
BOTTOM PLOT) Prediction CGM observations.

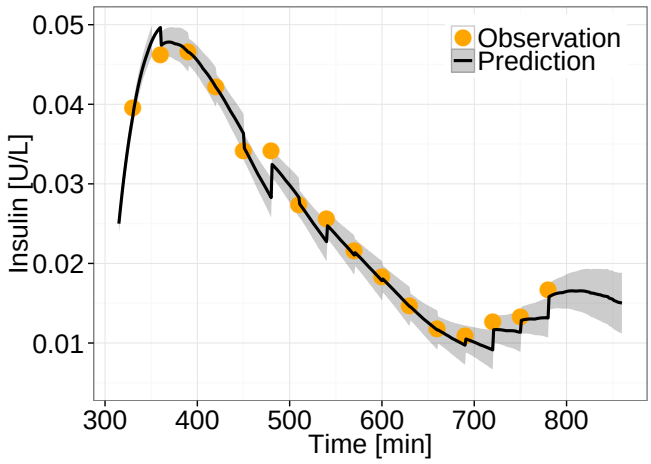


Figure 6.7: Model 2: Predictions using Equation (6.1) with estimation of diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$. Predictions of plasma insulin. The prediction behaves as expected, as it is adjusted when an observation is obtained.

Because this change in model structure relates to the pharmacokinetics it is important to ensure that the prediction of insulin still is acceptable. Insulin prediction is seen in Figure 6.7.

To validate insulin prediction, the residuals are examined. The residual examination has to be taken with some reservations. Only $n = 16$ observations are available for the entire time series. Furthermore it has not been possible to obtain a confidence indication of the measurements. But an indication of the validity of the insulin prediction is given.

The residual analysis is seen in Figure 6.8. ACF, PACF and the cumulative periodogram all indicate that the residuals are white noise. There is a slight tendency towards non-constant variance. However, it is also evident that the lack of observations has entailed large confidence bands. But all-in-all the prediction of plasma insulin seems to be valid, so the simplification of the model structure and introduction of bioavailability has made the model more robust without loss of predictability. Based on this examination insulin predictions will not be shown in Chapters 7 and 8.

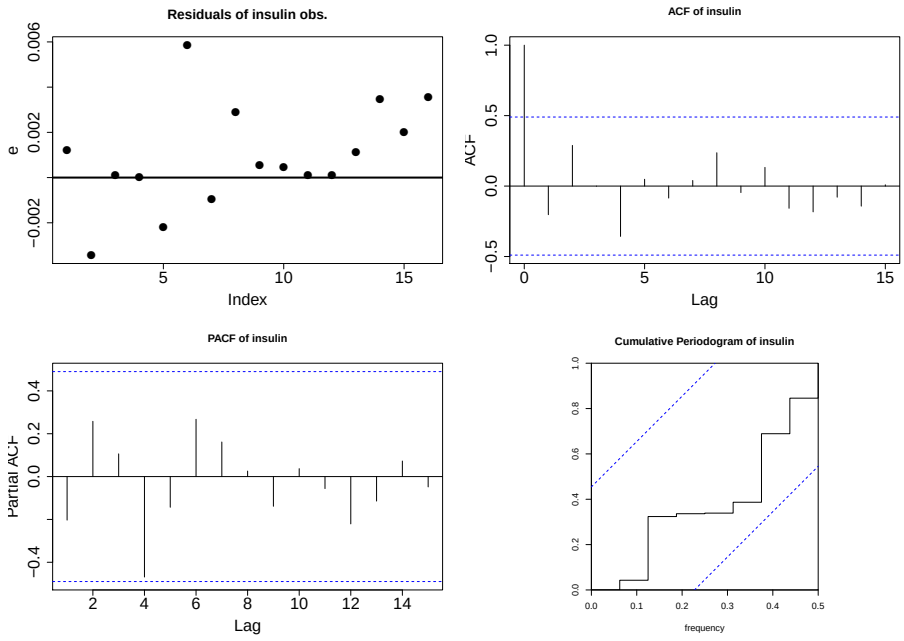


Figure 6.8: Model 2: Residual analysis of insulin prediction using Equation (6.1) with estimation of diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$. Resembles white noise.

The likelihood, degrees of freedom and AIC of the Model 2 is:

$$-\ln(L) = -92.73 \tag{6.2}$$
$$DF = 12 \tag{6.3}$$
$$AIC = -141.47 \tag{6.4}$$

As the two models in this chapter are not nested, AIC is considered for comparison. It is clear that AIC is lower, for Model 2, as $-141.47 < -98.01$, and therefore a statistical improvement has been shown.

6.3 Summary of initial findings

A suitable model structure has been identified, Equation (6.1). The grey-box modelling approach shows that further improvement of the model is possible. There are two possible improvements:

1. A Bayesian approach were MAP probability is applied on selected parameters.
2. Applying an Autoregressive process to the CGM device.

The following arguments are given for both approaches:

MAP

Applying MAP probability gives a possibility to include partial prior knowledge about parameters in the estimation. The prior knowledge is expressed by a prior probability density and the Bayesian approach is used for mixing the prior information with the information in the data [35]. The ability to include the prior standard deviations and the corresponding prior correlation matrix enables a physical interpretation of the magnitude of the estimations. Manufacture information about precision of devices enables adjustment of parameter estimates to the correct region of operation. This approach incorporates physical knowledge in the modelling. It furthermore gives a more correct estimation of the parameters, since any unexplained variation in the data will be reflected in the diffusion terms.

Autoregressive process

Applying an Ornstein-Uhlenbeck process, such as a Random Walk, enables the model to iteratively correct the prediction to the observation. The prediction of the CGM observation then essentially models the dynamic of the CGM device. The advantage is that the YSI observation will be

regarded as the most correct value, adjusting the predictions when a new observation $\boldsymbol{y}_j$ is obtained. There are however severe drawbacks:

1. Unexplained model structure is placed in the Random Walk.
2. The physiological considerations that can be made on the measurement equipment are neglected, minimising the influence of the grey-box modelling approach.

These two possible extension will be examined in Chapter 8. The next chapter will examine if Steil's model can be used as an alternative to the MVP model.

Chapter 7

Steil's model

Before an extension of the MVP model is made, an alternative model is investigated. Most literature focusing on glucose-insulin dynamics have an identical fundamental concept - Bergman's Minimal Model [14, 24, 33, 50]. Differentiation in approach is most often in the description of insulins effect. Steil's approach and the MVP model approach differ with regard to insulin dynamics. In this chapter focus is shifted to Steil's physiological more correct[33] approach.

7.1 Model 3 - Steil's model structure

The aim of this chapter is to give a short introduction to an alternative model. The details of the model are, therefore, kept at a minimum. A thorough investigation is out of the scope for this thesis, but it is believed that the achieved results will aid future work of PK/PD modelling. Only the original CGM device is used for modelling.

The diffusion term selection is shown in Figure 7.1. It is observed that only two diffusion terms, σ_1 and σ_4 are included. They belong to the state equations for

[33] The precise glucose-insulin dynamics is unknown, so more correct is relative to the extent of up-to-date knowledge.

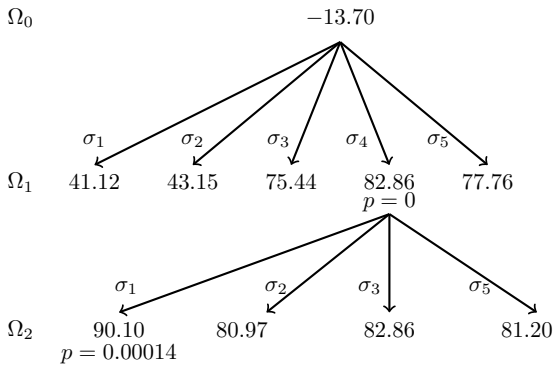


Figure 7.1: Forward selection for Steil model. Diffusion terms σ_1 and σ_4 are included in the order shown in the tree.

$\frac{dI_{sc}}{dt}$ and $\frac{dG_P}{dt}$, respectively, placing an estimable stochastic term on both the PK and the PD part of the model.

Having obtained relevant diffusion terms, the predictions and the residuals are presented in Figures 7.2 and 7.3, respectively. It is observed that by visual inspection the predictions are acceptable. Examining the residuals reveals that from a statistical point-of-view the goal has been reached. But when assessing the parameter estimates[34] from a grey-box modelling point of view, there is a large room for improvement as they are not physiologically interpretable.

For Model 3 the likelihood, degrees of freedom and AIC is calculated to:

$$-\ln(L) = 90.10 \tag{7.1}$$
$$DF = 13 \tag{7.2}$$
$$AIC = -128.20 \tag{7.3}$$

Model 3 and the Model 2 are compared using the AIC. Based on the AIC the most appropriate model is selected and the results are presented below:

$$AIC_{Model3} = -128.20$$
$$AIC_{Model2} = -141.47$$

So the Model 2 is chosen for further investigations because $-141.47 < -128.20$.

[34]The parameter estimates are not shown, since they do not contribute any relevant information.

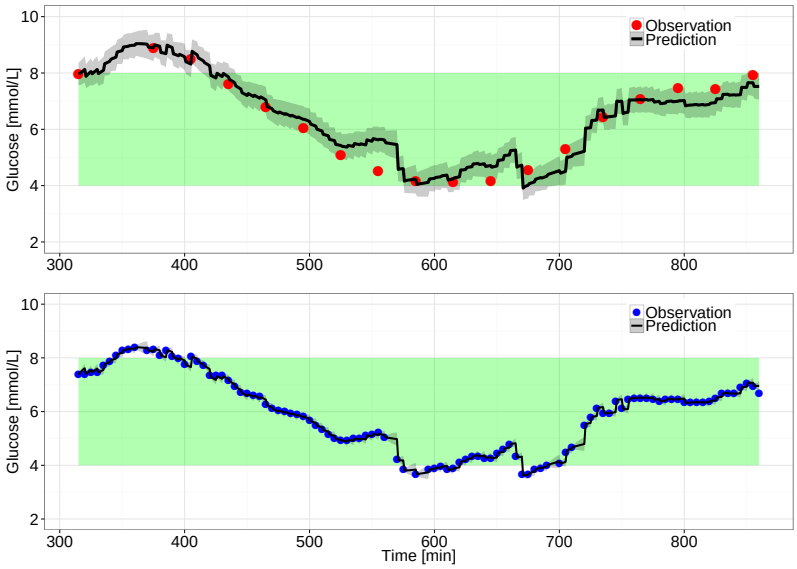


Figure 7.2: Model 3: Predictions using Equation (3.3) with estimation of diffusion terms $\boldsymbol{\sigma_1}$ and $\boldsymbol{\sigma_4}$.
TOP PLOT) Prediction of YSI observations.
BOTTOM PLOT) Prediction CGM observations. Predictions are perfect with a narrow 95% confidence interval.

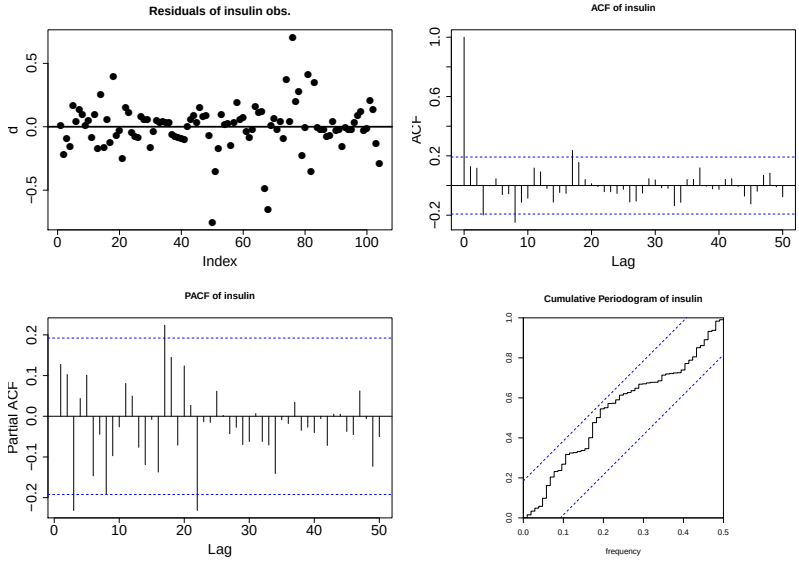


Figure 7.3: Model 3: Residual analysis of YSI and CGM observation predictions using Equation (3.3) with estimation of diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$. Resembles white noise.

7.2 Summary of Steil's model

An investigation of an alternative to the Medtronic Virtual Patient model has been conducted. The Steil model has a more complex and correct description of the effect of insulin. This more complex description introduces more parameters to the estimation. The introduction of these extra parameters do not give a significant improvement of the estimation of the glucose-insulin dynamics. Therefore the remainder of this thesis focuses on extending the MVP model.

Chapter 8

Extension of Medtronic Virtual Patient

This chapter explores the considerations given in Chapter 6. The first part uses the original CGM device. After identification of a suitable model, the alternate CGM device, observing the same time series, is included. Four different models, all an extension of Model 2, are examined in this chapter:

- Model 4: MVP + MAP
- Model 5:
  - Model 5a: MVP + Random Walk
  - Model 5b: MVP + AR(1)
- Model 6:
  - Model 6a: MVP + MAP + Random Walk
  - Model 6b: Model 6a on the alternate CGM device
- Model 7:
  - Model 7a: MVP + MAP and both CGM devices
  - Model 7b: MVP + MAP + AR(1) and both CGM devices

8.1 Model 4 - Bayesian approach

Multiple different MAPs can be included in the estimation. As the theory behind MAP is explained in Section 2.3, only the extension of that system is presented.

Extension of Equation (3.2) to include MAP:

Let $\boldsymbol{\Sigma_\theta}$ be the variance of the parameter set, $\boldsymbol{\theta}$. Adopting Equation (2.19), is done by specifying $\boldsymbol{\Sigma_\theta}$, as:

$$\boldsymbol{\Sigma_\theta} = \boldsymbol{\sigma_\theta} \boldsymbol{R_\theta} \boldsymbol{\sigma_\theta} \tag{8.1}$$

where $\boldsymbol{\sigma_\theta}$ is the diagonal matrix of standard deviations and $\boldsymbol{R_\theta}$ is the corresponding prior correlation matrix. Both specified only for the parameters that have a prior probability.

Specifying MAP is relevant on the YSI2300 and the CGM devices. The initial parameter values of the measurement noise, $sYSI$ and $sCGM$, correspond to the equipments theoretical expected variance. From the manufacturer precision charts, the variance can be found and transformed to the correct domain, using the *Coefficient of Variation, CV*:

$$CV = \frac{\sigma}{\mu}$$

The transformation to the estimation domain becomes:

$$\ln \sigma^2 = \ln \left((CV \cdot \mu)^2 \right) \tag{8.2}$$

where $\mu_{YSI} = \mu_{CGM} = 7$. $CV_{YSI} = 0.01$ [26], and $CV_{CGM} = 0.1$ set arbitrary, since no reliable literature is at hand. In the transformed space this gives estimates of $\boldsymbol{\Sigma}_{YSI} = -5.32852$ and $\boldsymbol{\Sigma}_{CGM} = -0.71133$.

Including MAP requires two important considerations:

Distribution in logarithmic scale

Applying MAP in CTSM-R tool is performed as described in Equation (8.1), where, based on previous studies, the prior information specifies the deviation from an expected value. This deviation is Gaussian distributed in the logarithmic parameter space. Intuitively this gives a problem in the interpretable parameter space since the

transformation to the interpretable parameter space will not be Gaussian distributed. It is assumed that this error is of such a small magnitude that it will not interfere with the result.

Uncertainty of the uncertainty

$\boldsymbol{\Sigma_\theta}$ is the uncertainty of a parameter $\boldsymbol{\theta}$. $\boldsymbol{\sigma_\theta}$ is the standard deviation of that uncertainty. A literature study has shown no useful indication of this standard deviation, so it is said that $\boldsymbol{\sigma}_{\boldsymbol{\theta},YSI} = \boldsymbol{\sigma}_{\boldsymbol{\theta},CGM} = 0.1$. This introduces an error to the model, but it is assessed to be the best solution.

The deviance, $\boldsymbol{\sigma_\theta}$, is calculated in the interpretable parameter space as the 5% deviation of $\boldsymbol{\Sigma_\theta}$. The value is then transformed and the difference between this transformed value and the transformed deviation of $\boldsymbol{\Sigma_\theta}$ is used as the prior $\boldsymbol{\sigma_\theta}$. The calculated values are:

$$\sigma_{YSI} = 0.1158 \tag{8.3}$$

$$\sigma_{CGM} = 0.048518 \tag{8.4}$$

Intuitively σ_{YSI} should be smaller than σ_{CGM}, but the logarithmic scale changes the intuitive relationship.

8.1.1 Model 4 - MVP + MAP

Figure 8.1 shows the result of implementing MAP[35] on $sYSI$ and $sCGM$. It is evident that the prediction of the YSI observations follows the actual YSI observations more precisely than Model 2. The prediction of the CGM observations is not as precise as before and the confidence interval is wider[36] substantiated by the parameter estimates in Table A.3.

It is seen that the estimation of the measurement noise is $sYSI = \exp(-5.33) \pm \exp(0.05)$ and $sCGM = \exp(-0.80) \pm \exp(0.05)$. Compared to Model 2 where $sYSI = \exp(-1.02) \pm \exp(0.40)$ and $sCGM = \exp(-16.14) \pm \exp(98.27)$[37], the measurement noise has decreased for the YSI device and increased for the CGM device as expected. It has stayed within the limits of what is statistically and physiologically reasonable.

[35] An examination of the quantiles in each domain could reveal this influence of the skewed logarithmic scale.

[36] $sCGM$ has increased.

[37] Both insignificant in the model.

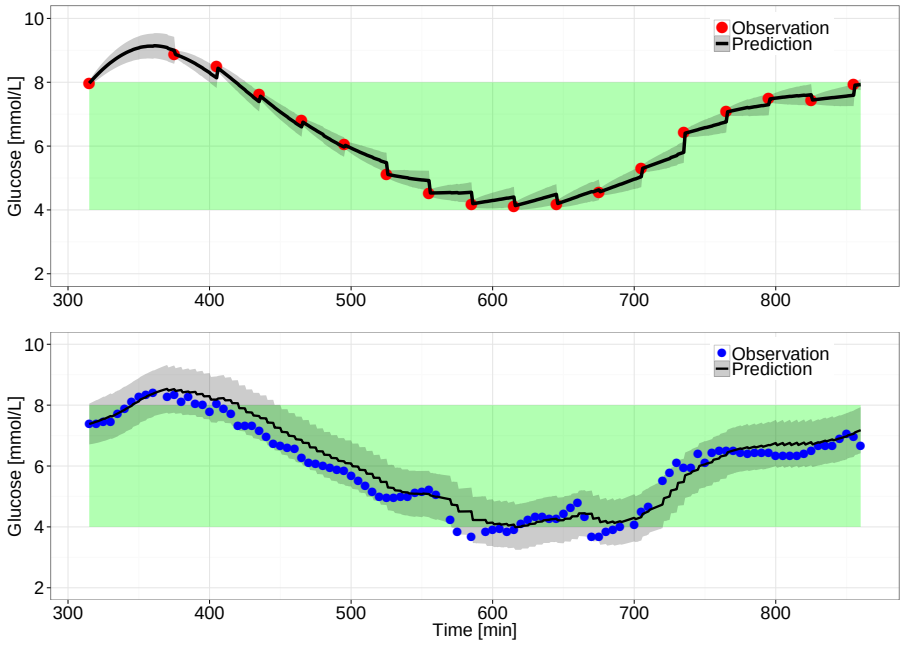


Figure 8.1: Model 4: Predictions using Equation (8.1) with estimation of diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$. Physiologically meaningful estimation is seen as 95% confidence interval is narrow on the YSI observation prediction and wide on the CGM observation prediction.
TOP PLOT) Prediction of YSI observations.
BOTTOM PLOT) Prediction CGM observations.

Examining the residuals, Figure B.4, it is revealed that an implementation of an AR(1)-process [34] is appropriate for the CGM device, because of the structure in the residuals [34]. Supporting this solution are the magnitudes of $\sigma_4 = \exp(-2.76)$ and $\sigma_5 = \exp(-2.26)$, both diffusion terms associated with the devices. It has to be noted that there is not constant variance. An AR(1)-process extension may improve this.

The likelihood, degrees of freedom and AIC are calculated for Model 4:

$$\begin{aligned} -\ln(L) &= -7.466 \\ DF &= 12 \\ AIC &= 58.93 \end{aligned} \tag{8.5}$$

When comparing with Model 2, Wilk's Likelihood ratio test cannot be used

since the models are not nested. Considering AIC, it does not seem as there is an improvement compared to Model 2, as $-141.47 < 58.93$. But here it is important to consider the achieved physiological improvement.

8.2 Model 5 - Autoregressive modelling

This section will focus on autoregressive modelling. An understanding of the functionality of the CGM device hopefully gives an improvement in the general understanding of glucose-insulin dynamics.

This section will first present an investigation using the original CGM device. The alternate CGM device will be introduced to investigate the possibility of a more accurate modelling.

Extension of Equation (3.2) to a continuous AR(1)-process [57]: The Ornstein-Uhlenbeck process is a continuous AR(1)-process that can be written as:

$$dk_t = \phi\left(\mu - k_t\right)dt + \sigma d\omega_t \tag{8.6}$$

where $\phi > 0$ is the rate of decay and $\mu > 0$ is the limit of the decay.

The Ornstein-Uhlenbeck process allows for a stochastic relaxation process. This relaxation does not necessarily go towards zero when $t \to \infty$, so further structure will be introduced to the system.

A simple Ornstein-Uhlenbeck process with an exponential decay towards zero is expected to work. Equation (8.6) simplifies to:

$$dk_t = -\phi \cdot k_t\, dt + \sigma d\omega_t \tag{8.7}$$

where $\phi > 0$.

No extra structure is introduced into the system, as the solution to the integral will be exponentially decaying towards zero as $t \to \infty$.

To obtain an indication of ϕ and μ, a prediction using Random Walk, can be made. The prediction of the Random Walk, $\hat{k}$, can be examined by a non-parametric spline representation plotted against time. A special case of the Ornstein-Uhlenbeck process is the Random Walk described below:

Extension of Equation (3.2) to include Random Walk:

$$dk_t = \sigma d\omega_t \tag{8.8}$$

with output

$$CGM = G_{sc}(t) + k_t + sCGM$$

where k_t is a Random Walk, and all remaining parameters have been defined previously.

8.2.1 Model 5a: MVP + Random Walk

Implementing Equation (8.8) gives the predictions shown in Figure 8.2. The prediction of the YSI observations improve, as the predictions between observations overall reflect the dynamic. With a new observation, $\boldsymbol{y}_j$, the model prediction is adjusted, sometimes with a steep transition. In the middle half of the time series from 500 min to 700 min steep transitions are observed. These transitions will be discussed in Section 9.3.2. The confidence interval of the YSI observation prediction also reflects the nature of the Random Walk. Around an YSI observation the confidence interval is narrow, and as time progresses it expands until a new observation is obtained.

From the parameter estimates, Table A.4, it is evident that the measurement noise are estimated very low, $sYSI = \exp(-17.15)$ and $sCGM = \exp(-16.92)$[38]. Since they are both insignificant, from a statistical point-of-view they can be removed. From a grey-box modelling point-of-view, the two parameters should have an influence. The reason for their insignificance is the Random Walk, and its ability to adjust according to new observations. The residual analysis, Figure B.5, indicates that a Random Walk makes the model statistically sensible.

The approach to model the dynamics of the CGM device with a Random Walk works as a concept. But the approach is flawed, as the improvement is not an improvement of the model or better description of physiology, but rather a purely statistical improvement.

[38] Both insignificant.

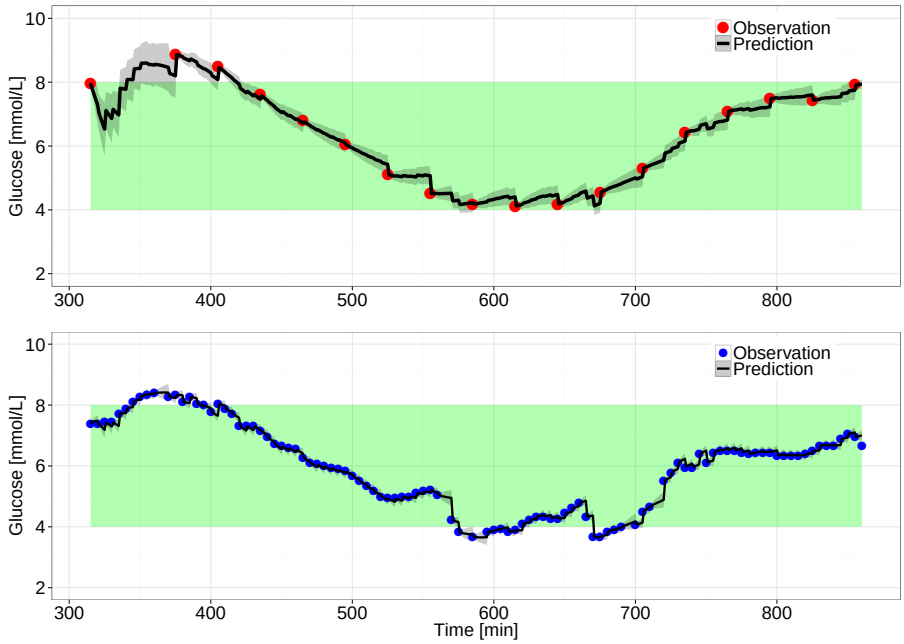


Figure 8.2: Model 5a: Predictions using Equation (6.1) and Random Walk with estimation of diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$. As expected the nature of the Random Walk corrects the prediction to match each new observation.
TOP PLOT) Prediction of YSI observations.
BOTTOM PLOT) Prediction CGM observations.

For Model 5a the likelihood, degrees of freedom and AIC are calculated:

$$\begin{aligned} -\ln(L) &= 102.02 \\ DF &= 14 \\ AIC &= -156.04 \end{aligned} \tag{8.9}$$

Using Wilk's likelihood ratio test to compare with Model 2 in Chapter 6, gives $p = 0$. Including Random Walk significantly improves the model[39]. However, using a grey-box approach the model estimates do not make sense.

8.2.2 Model 5b: MVP + AR(1)-process

As all residual analyses up to this point has indicated an AR(1)-process extension, this is investigated. Initially the simple Ornstein-Uhlenbeck process, Equation (8.7), is used. It is expected that this will give a significant improvement as predictions are adjusted according to the previous value of the observations. Figure 8.3 show the prediction. The predictions are, as expected, good - the initial variation of the prediction of the YSI observation is a result of the removed observation. The residuals shown in Figure B.6 are close to white noise, but indicate that an AR(2)-process may improve the estimation, as there is still indication of structure in the ACF and the PACF.

The estimates of the measurement noise: $sYSI = \exp(-16.44)$ and $sCGM = \exp(-16.39)$ indicate that there is a physiological problem with the model[40]. The measurement noise is unrealistically small, when compared to the theoretical values obtained previously in the Bayesian approach.

Remark 8.1 *As the residuals of the AR(1)-process indicated that an extension to an AR(2)-process is needed. This was investigated, but the extension did not yield any new results, and thus it is not presented.*

A pure implementation of an AR(1)-process gave a significant improvement to the model estimation. The next section will investigate the combined implementation of MAP probability and an AR(1)-process.

[39] It is meaningless to compare the model with an estimation using MAP as the estimates are not physiologically comparable.

[40] The remaining estimates are not shown as they are not informative.

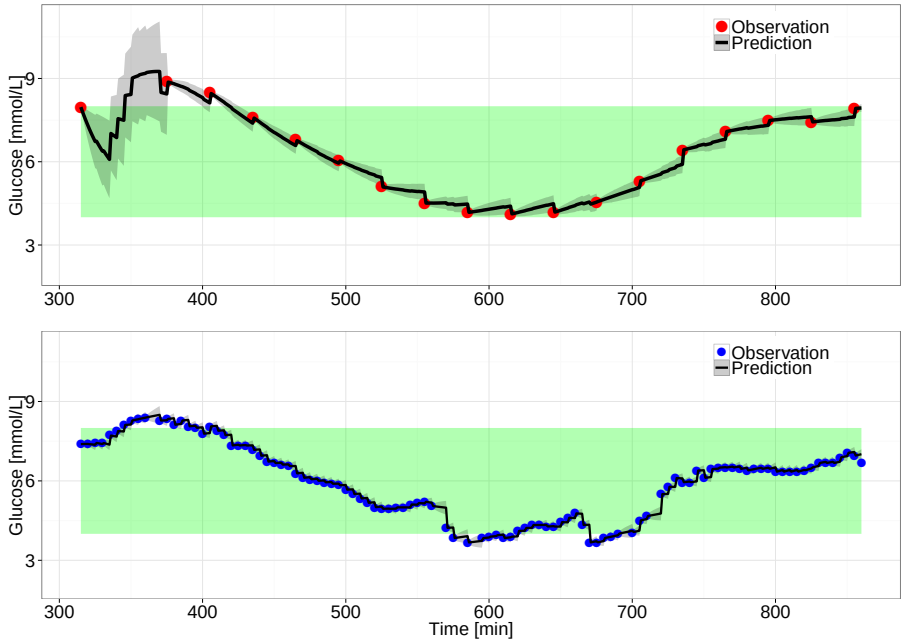


Figure 8.3: Model 5b: Predictions using Equation (6.1) and AR(1)-process with estimation of diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$. As expected the AR(1)-process adjusts the prediction to according to the previous observation.
TOP PLOT) Prediction of YSI observations.
BOTTOM PLOT) Prediction CGM observations.

For Model 5b the likelihood, degrees of freedom and AIC are calculated:

$$\begin{aligned} -\ln(L) &= 93.09 \\ DF &= 15 \\ AIC &= -136.18 \end{aligned} \tag{8.10}$$

Remark 8.2 *Comparing with Model 5b, it seems strange that the likelihood decreases when extending the model from a Random Walk to an AR(1)-process. Theoretically the estimate should at least be equally good as the nested model. This is discussed in Section 9.3.2.*

8.3 Model 6 - Combining Bayesian and autoregressive modelling

A simple Ornstein-Uhlenbeck process did not give the expected improvement. Therefore, non-parametric modelling is performed, where a Random Walk is combined with MAP probability. The aim of the non-parametric modelling is to detect an overall tendency in the predictions. Based on the observed tendencies an Ornstein-Uhlenbeck process describing the behaviour may be constructed.

8.3.1 Model 6a: MVP+ MAP + Random Walk

Having one time series observing G_{sc} makes it difficult to distinguish structure and noise in the non-parametric plot. For the original CGM device the non-parametric smoothed model is shown in Figure 8.4, where the Random Walk $\hat{k}$ is plotted against time and the output from the CGM device, respectively. The top plot does not resemble noise, but no general tendency is seen. The bottom plot shows, with a pragmatic approach, an exponential decay towards $\mu = -0.1$ making it reasonable to apply Equation (8.7). The residual analysis, Figure B.7, strongly implies that an implementation of some AR(1)-process is reasonable, because of the structure in the residuals [34]. The prediction of the YSI observations and CGM observations can be seen in Figure C.3.

For Model 6a the likelihood, degrees of freedom and AIC are calculated to:

$$\begin{aligned} -\ln(L) &= 2.12 \\ DF &= 14 \\ AIC &= 43.75 \end{aligned} \tag{8.11}$$

It only makes sense to compare with models using MAP[41]. Using Wilk's likelihood ratio test to compare with Model 4 gives, $p = 0.000069$ so a statistical improvement is obtained expanding from MAP to MAP combined with Random Walk.

Remark 8.3 *An implementation of an AR(1) process as Equation (8.7) was investigated. This only gave a visual improvement of the prediction. No statistical improvement was seen. ϕ was estimated around $\phi = 2$, giving a rate of decay at $\frac{1}{\phi} = \frac{1}{2}$*min*, on a time series having a resolution of* 1 min. *This investigation is not shown here.*

[41] Comparing to models not including MAP is meaningless as the models are not physiological comparable.

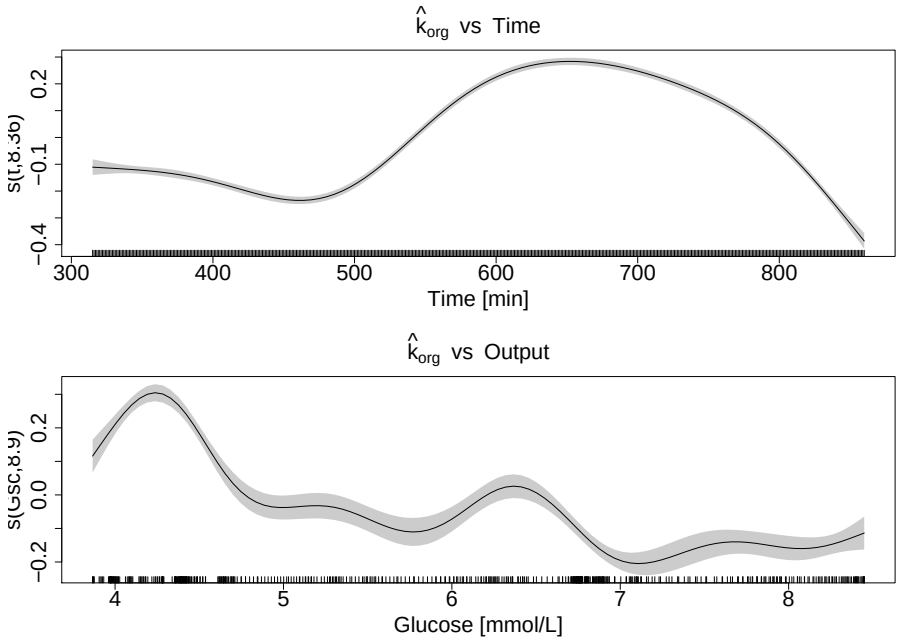


Figure 8.4: Model 6a: Non-parametic spline plot using Equation (6.1) with MAP and Random Walk and the original CGM device.
TOP PLOT) Non-parametric smoothed plots of $\hat{k}$ plotted against time.
BOTTOM PLOT) Non-parametric smoothed plots of $\hat{k}$ plotted against output.

Until now the investigation has been based on one patient and one CGM device. To get a better understanding of the general dynamics of CGM devices, the tendency of the devices can be examined different ways:

Investigation of a different CGM device

The alternate CGM device observed the same experiment on a different location of the body. This presents two possibilities:

1. An estimation using the alternate device to investigate concordance between the two devices.
2. A model estimation using both devices. In this case the devices observe the same state.

Investigation of several patients

Using the same device on many different patients would be the optimal

solution. By use of non-parametric modelling, device tendencies are separated from noise, allowing identification of the proper AR-process. This is left for future work.

8.3.2 Model 6b: Using the alternate device

In the following the alternate CGM device is investigated, see Figure 1.4. The same tedious diffusion term examination has been performed, and yielded the same result[42] as in Section 6.2. Following this examination there were, as before, indications that an AR(1)-process was needed.

The non-parametric representation of the prediction is shown in Figure 8.5, where $\hat{k}$ plotted against time and the output from the CGM device[43].

From the the top plot it is not possible to identify a distinct dynamic. The variation does not resemble white noise, but it is difficult to identify an actual effect. The bottom plot, with a pragmatic approach, resembles an exponential increase towards $\mu = 0.1$. It however seems strange that the dynamics is so different from that in Figure 8.4.

For Model 6b the likelihood, degrees of freedom and AIC are calculated to:

$$\begin{aligned} -\ln(L) &= -28.00 \\ DF &= 14 \\ AIC &= 104.00 \end{aligned} \tag{8.12}$$

Comparing Model 6b with Model 6a with AIC, it is evident that the using the alternate device does not improve the estimations, as $43.75 < 104.00$. This is not surprising considering the dynamics of the alternate CGM device.

Remark 8.4 *The residuals indicated, also in this case, the need to implement an AR(1)-process. As before only a visual improvement of the prediction is observed. No statistical improvement was seen.* ϕ *was estimated around* $\phi = 2$*, giving a rate of decay at* $\frac{1}{\phi} = \frac{1}{2}\text{min}$*, on a time series having a resolution of* 1 min*. This investigation is not shown here.*

[42] The investigation is not presented.

[43] The predictions of the YSI observations and CGM observations is seen in Figure C.4, the residuals in Figure B.8.

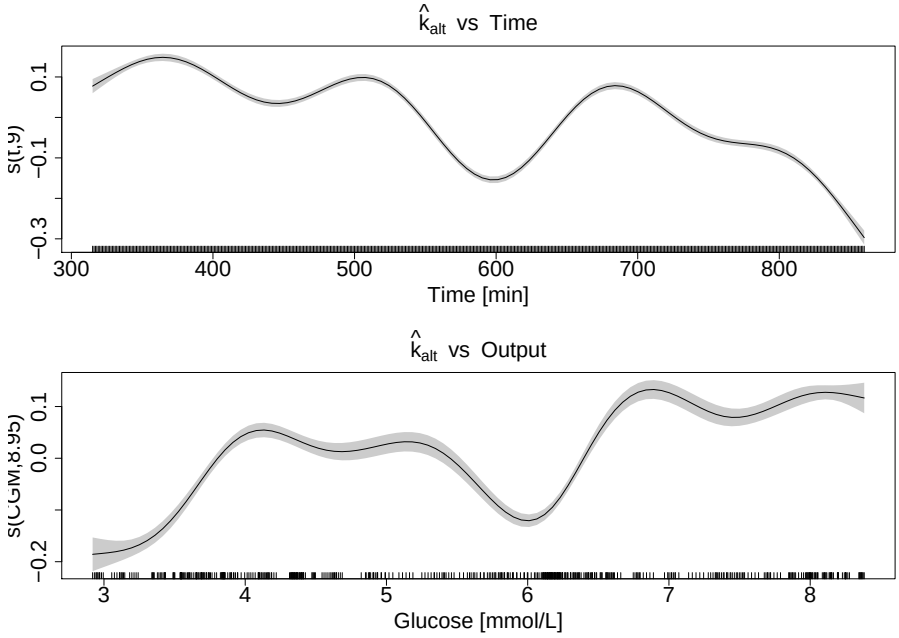


Figure 8.5: Model 6b: Non-parametic spline plot using Equation (6.1) with MAP and Random Walk and the alternate CGM device.
TOP PLOT) Non-parametric smoothed plots of $\hat{k}$ plotted against time.
BOTTOM PLOT) Non-parametric smoothed plots of $\hat{k}$ plotted against output.

8.4 Model 7 - MAP and an AR-process using two CGM devices

A combined investigation of the two CGM devices monitoring the same state is tried.

Extension of Equation (8.6) to two CGM devices:

Estimating the model using two different CGM devices presents a challenge as they are observing the same state. If the new CGM device is added as just an observation, the prediction for the two CGM devices will be identical. One way to avoid this is to introduce calibration parameters as proposed by Breton and Kovatchev in [8]:

$$y_1 = \alpha_1 x_1 + \beta_1 \tag{8.13}$$
$$y_2 = \alpha_2 x_2 + \beta_2 \tag{8.14}$$

where α and β are parameters allowing the model to distinguish the predictions of the two CGM devices.

8.4.1 Model 7a - MVP + MAP and both CGM devices

The usual tedious investigation of diffusion terms, parameter settings and boundaries has been performed[44]. Enforcing MAP on all devices gave the following likelihood, degrees of freedom and AIC for Model 7a[45]:

$$\begin{aligned} -\ln(L) &= -106.21 \\ DF &= 15 \\ AIC &= 262.41 \end{aligned} \tag{8.15}$$

A non-parametric spline representation of the predictions $\hat{k}_{org}$ and $\hat{k}_{alt}$ from the Random Walk is seen in Figure 8.6. Observing the non-parametric rendition no easily identifiable dynamic can be seen. However, plotting $\hat{k}_{org}$ and $\hat{k}_{alt}$ against the output of each respectable CGM device, Figure 8.7, a vague pattern emerges. With a very pragmatic approach an exponential increase towards $\mu = 0$ is seen.

[44] These results are not presented, as they are essentially identical with what has been observed earlier.

[45] It is not sensible to compare values with previous models, so this is to be used to detect improvements when expanding the current model.

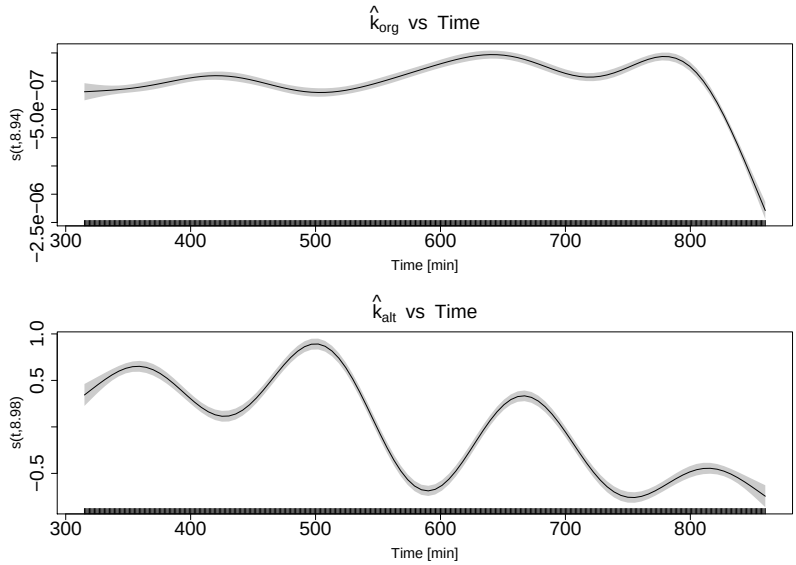


Figure 8.6: Model 7a: Non-parametic spline plot using Equation (8.13) with MAP and Random Walk on both CGM devices.
TOP PLOT) Non-parametric smoothed plots of $\hat{k}_{org}$ plotted against time for the orginial CGM device.
BOTTOM PLOT) Non-parametric smoothed plots of $\hat{k}_{alt}$ plotted against time for alternate CGM device.

This together with the residuals, Figures B.9 and B.10, indicate that an AR(1)-process will improve the model estimation.

8.4.2 Model 7b - MAP + AR(1) and both CGM devices

Figure 8.8 shows the final result in this chapter, which is an implementation of Equation (8.7) on both devices[46]. It is evident that the prediction follows the observations closer, especially evident in the bottom plot. This is, undoubtedly a result of the AR(1)-process, allowing the model to adjust according to the last observation.

Combining the two CGM devices gives the following inferential statistics for

[46]The prediction of the YSI observations is not shown since it is similar to previous predictions.

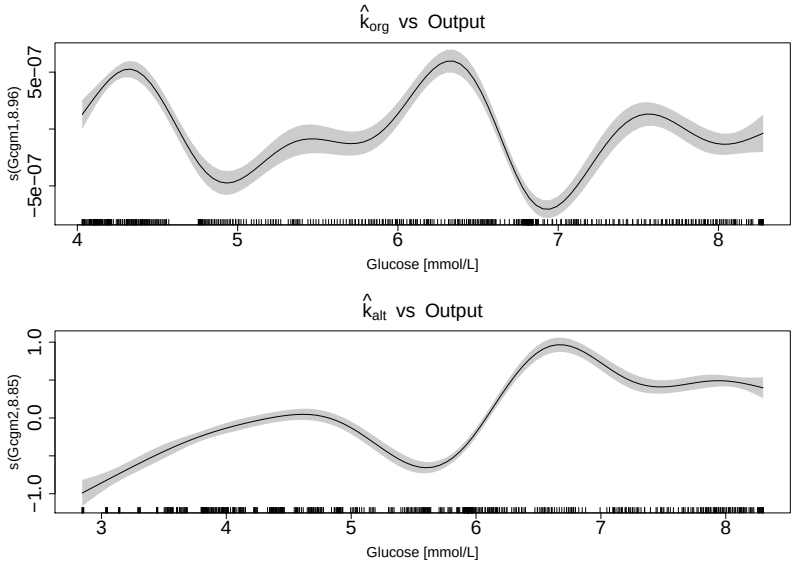


Figure 8.7: Model 7a: Non-parametic spline plot using Equation (8.13) with MAP and Random Walk on both CGM devices.
TOP PLOT) Non-parametric smoothed plots of $\hat{k}_{org}$ plotted against the output of the orginial CGM device.
BOTTOM PLOT) Non-parametric smoothed plots of $\hat{k}_{alt}$ plotted against the output of alternate CGM device.

Model 7b:

$$\begin{aligned} -\ln(L) &= -91.82 \\ DF &= 19 \\ AIC &= 245.65 \end{aligned} \tag{8.16}$$

Comparing Model 7a and Model 7b with Wilk's likelihood ratio test gives, $p = 0.00000866$. A significant improvement is found when extending to an AR(1)-process on each of the CGM devices.

Considering the parameter estimates Table A.5 it is evident that, from a grey-box modelling perspective, estimates have improved. The estimates of measurement noise are[47]: $sCGM1 = \exp(-0.815) = 0.44$, $sCGM2 = \exp(-0.786) = 0.46$ and $sYSI = \exp(-5.329) = 0.0048$.

[47] All significant, with Std. Error shown in Table A.5

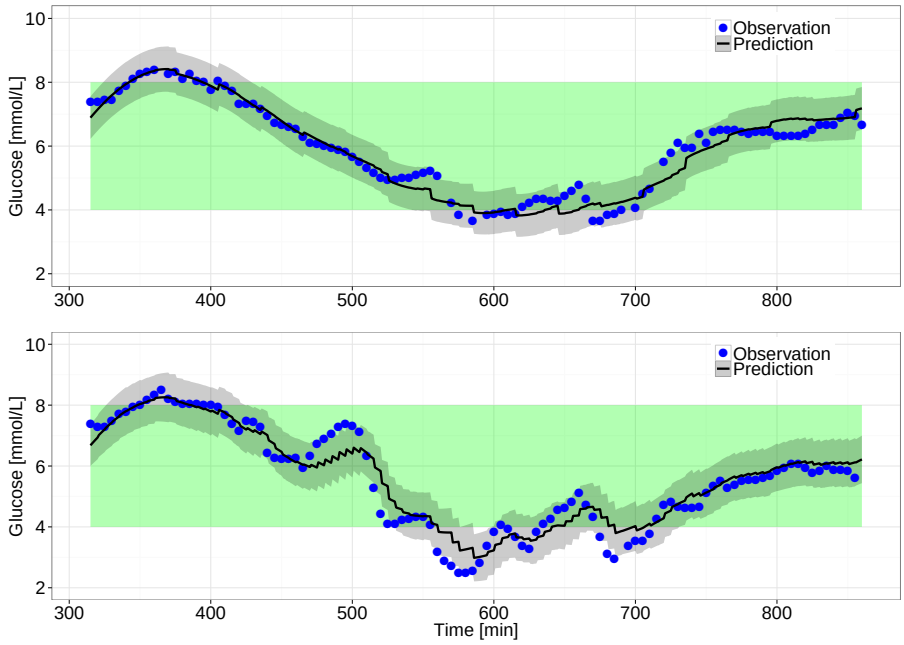


Figure 8.8: Model 7b: Predictions using Equation (8.13) and AR(1)-process with estimation of diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$.
TOP PLOT) Prediction of the original CGM observations.
BOTTOM PLOT) Predictions of the alternate CGM observations.

Table 8.1 shows a comparison of the parameter estimates of Model 7b with the estimates from the MVP model [24]. For the MVP model the range of the estimates, from lowest to highest, are shown. For Model 7b the estimate is shown[48]. Blue estimates indicate that they fall within the range of the MVP model estimates.

All estimates fall within the range of the MVP model. Variations in MVP model estimates are most likely patient specific, but this thesis has shown that extending to SDEs maintains reasonable[49] estimates. Estimates of the time constant, τ_1 and τ_3, are physiologically meaningful and in concordance with available literature. As information about deficiencies is available in the diffusion term, $\sigma_4 = \exp(-2.748) = 0.064$ must be emphasised being considerably larger than the remaining four state diffusion terms. σ_4 relates to G_P, indicating that further improvement must start here.

[48] A statistical examination of deviations between the two models is not performed, as the inferential data from the stochastic model is only based on one patient.
[49] Compared to current literature.

θ	Estimates	
	MVP	Final
C	$0.54 - 2.01$	1.2^*
τ_1	$41 - 131$	65.71
τ_3	10^*	9.58
p_2	0.0096 - 0.0102	0.0098
EGP	$0.033 - 0.19$	0.05
$GEZI$	$7.58 \cdot 10^{-8}$ - $1.00 \cdot 10^{-3}$	$3.05 \cdot 10^{-7}$
S_i	$0.0964 - 1.73$	0.38

Table 8.1: Model 7b: Comparison of estimates from the MVP and the final model obtained in this thesis. Blue values indicate that estimates obtained in this thesis is within the range of the estimated values in the MVP. Asterisk (*) denotes fixed value. Parameters have been transformed so they have the same units as presented in Chapter 3.

Remark 8.5 *The residual analysis is seen in Figures B.11 and B.12. The residual analysis strongly indicates that further improvements can be obtained. Comparing with the residuals having a Random Walk (B.9 and B.10) there is not a large visual difference. An extension to an AR(2)-process on both CGM devices was investigated but did not give a significant improvement, nor a change in the residual analysis.*

An overview of the model progression is seen in Table 8.2. Here models are grouped as they are nested. The full parameter set should give the best statistical description. It is evident that this is not the case in the nested models without MAP. This will be discussed in Section 9.3.2. The table illustrates that with a purely statistical approach Model 5a would be chosen as the best model. Using physiological considerations it is however evident that Model 7b models the data physiologically more correct.

8.5 Summary of extensions

This chapter has investigated different extension of the Medtronic Virtual Patient model. Using a Bayesian approach MAP probabilities were implemented in the model, restricting parameter estimates to a physiologically interpretable range. The residual analysis of the predictions revealed that there is room for improvement.

	Model	DF	$-\ln(L)$	AIC
Nested models without MAP				
$\Omega_{full,1}$	Model 5b	15	93.09	-136.18
$\Omega_{full-1,1}$	Model 5a	14	102.02	-156.04
$\Omega_{full-2,1}$	Model 2	12	-92.73	-141.47
Nested models with MAP				
$\Omega_{full,2}$	Model 6a	14	2.12	43.75
$\Omega_{full-1,2}$	Model 4	12	-7.46	58.93
Nested models with two CGM devices and MAP				
$\Omega_{full,3}$	Model 7b	19	-91.85	245.65
$\Omega_{full-1,3}$	Model 7a	15	-106.21	262.41

Table 8.2: Overview of likelihoods and AIC values for the original glucose monitor grouped as nested. Ω_{full} has the largest parameter set, thus making the below models subsets, within the respective model structure. Illustrates the problem between physiological modelling and statistical modelling.

Shifting focus to autoregressive modelling, improvements were sought using Ornstein-Uhlenbeck processes on the measurement devices. This yielded, not surprisingly, an improvement in the residual analysis, as the previous value of the subcutaneous glucose plasma observation was given an influence in the prediction. The parameter estimations were however not physiologically meaningful.

To obtain physiologically meaningful and statistical sound estimates a combination of MAP and Ornstein-Uhlenbeck processes was examined. Using non-parametric modelling it was hoped to identify parameters for the Ornstein-Uhlenbeck process. Indications were given that an AR(1)-process will aid this estimation. Furthermore, a combination of two CGM devices was shown to give a significant improvement in the estimations, also when introducing an AR(1)-process. An improvement not obtained when using just one CGM device.

Chapter 9

Discussion

This thesis covers PK/PD modelling using stochastic differential equations on state-space form to investigate glucose-insulin dynamic in T1D-patients. Results and perspectives will be discussed in this chapter leading to the conclusion in Chapter 11.

The initial investigation revealed a problem with the identifiability of the data. The deficiency manifested when estimating SDEs having negligible diffusion terms - values varying between $\exp(-12)$ to $\exp(-18)$ gave different estimates. Initially, lack of excitation was blamed. A thorough investigation of structural identifiability and persistence of excitation was conducted, revealing an identifiable system. Proving persistence of excitation in large systems is difficult, so the system was simplified to two estimable parameters. This enabled a visualisation of the problem - the parameter space changed as the diffusion term changed.

Theoretical identifiability was established using Equation (5.7). This discovery lead to an examination of the initial covariance matrix, $\boldsymbol{P}_0$. For longer time series, Equation (5.8) is not problematic as there is an abundance of data. As presented in Chapter 5, the diffusion term $\boldsymbol{\sigma}$ may influence the state prediction in Equation (5.10), causing problems. The problem was solved by using an initial covariance matrix of the form:

$$\boldsymbol{P}_0 = k \cdot \boldsymbol{I}$$

producing the first major, and quite unexpected, result of this thesis.

The focus in this thesis was an investigation of the Medtronic Virtual Patient model, and it was therefore natural to formulate the MVP model on state-space form with stochastic differential equations. The MVP model can cope with exogenous meal inputs, but this is not investigated.

9.1 Model 2

With the available data a simplification of the PK-part was necessary to obtain a sufficient model description. Insulins movement from pump to plasma was described by one rate constant, τ_1[50]. The simplified model structure entailed an investigation of the predictability of plasma insulin as the change in structure relates to this part[51]. As seen in Figure 6.8, white noise is obtained. However, with only 16 observations, there are reservations. It is difficult to say whether there is constant variance. The measurement noise of the insulin measurement, $sins = -13.41$, is unreasonably low, compared to the measurement noise of both the YSI device and CGM device. But as no device information was available, it is difficult to account for the reliability of this estimate. The predictions seem reasonable, so it was assumed not to influence the general modelling.

9.2 Model 3

The Steil model, having a more complex (and correct) description of insulin effect, was examined. It was included to examine whether a complex model structure would describe data variation better. A more detailed structure than that of the MVP model, was hoped to capture the dynamics better.

It has however been shown that more complicated model structure does not give a statistical difference. A more thorough investigation of the Steil model has not been conducted. Indications point towards that the MVP model statistically described data as well as Steil's model. Emphasis was, therefore, put on the MVP model.

[50] For alternative suggestions for insulin PK consult [62].

[51] Comparing Figure 6.7 with Figures C.1 and C.2 a visual inspection reveals that the prediction seems feasible.

9.3 Extensions of the MVP model

The severe indications of a deficiency in the PD-part of the model system[52] resulted in an analysis of the measurement equipment. Two different approaches have been incorporated and tested: a MAP probability approach and an autoregressive approach.

9.3.1 Model 4

Using MAP probability, information about devices are incorporated in the model estimation. This gives physiologically reliable estimations of the relevant parameters and thus reliable estimates of the remaining parameters. This improvement does not come for free:

1. Equation (8.1) gives a Gaussian distribution in the estimation space, and, because this is a logarithmic scale, the transformation to the parameter space is skewed.
2. Furthermore, because the uncertainty of the uncertainty is not a subject well-described in the literature, a pragmatic approach quantifying these magnitudes is taken. It is assumed that the advantages outweigh the disadvantages and uncertainties created - no indications to the opposite have been seen.

From Figure 8.1 it is evident that the prediction of both the YSI observations and CGM observations are more reasonable. Since the YSI observations are considered gold standard they are considered very reliable.

The measurement noise for the CGM device has increased and the prediction does not follow the observations exactly. Steep transitions in the CGM observations are not followed by a direct transition in the prediction. This is substantiated by the parameter estimates that are in a physiologically interpretable range. As plasma insulin is also measured using a device it may seem strange that MAP is not enforced on the measurement noise of insulin: *sins*. MAP was not enforced, since no information about the measurement device was available. In this study it was assumed not to influence the modelling. For future studies MAP should be enforced on all measurement devices.

[52] Estimates of $sYSI$, $sCGM$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$

9.3.2 Model 5

An autoregressive approach to accommodate non-physiological parameter estimates was taken. A special case of the Ornstein-Uhlenbeck process, the Random Walk, was implemented. Unsurprisingly the predictions of the CGM observations became perfect, as they were adjusted with each new observation $\boldsymbol{y}_j$. These adjustments gave steep, unphysiological transitions in the predictions. Furthermore a Random Walk makes k-step prediction and simulation very unreliable, emphasising the need of alternative methods. The residuals indicate that an AR(1)-process seemed reasonable. An Ornstein-Uhlenbeck process, Equation (8.7), gave visually nice predictions and a residual analysis indicating white noise. But parameter estimates and measurement noise, $sYSI$ and $sCGM$, are unphysiological.

In Table 8.2, it seems strange that the implementation of an AR(1)-process does not significantly improve the model compared to the implemented Random Walk. Theoretically this should not happen. It is expected that models, when expanded, are at least as good as the preceding model. This irregularity may be caused by an overparametrisation of the model or an issue with the implementation and estimation of the AR(1)-process. Due to time constraints this issue was not investigated further, but future studies should be aware of the problem.

9.3.3 Model 6

To ensure both statistically and physiologically correct models, MAP and autoregressive processes were combined. Indications of the structure of the Ornstein-Uhlenbeck processes were obtained with non-parametric analysis of the prediction $\hat{k}$ when implementing a Random Walk. Analysing one CGM device on one patient did not give a reliable indication of residual structure. Extending with the alternate CGM device measuring the same patient at a different location, did not clarify what was white noise and what was device influence.

9.3.4 Model 7

To utilise all available information, both CGM devices were used in the estimation, observing the same compartment (G_{sc}). Expanding to an AR(1)-process gave a statistical improvement[53]. Using multiple devices parameter estimates,

[53] Multiple devices have been shown to increase validity [10].

Model	DF	$-\ln(L)$	AIC
Model 1	13	94.45	-98.01
Model 2	12	-92.73	-141.47
Model 3	13	90.10	-128.20
Model 4	12	-7.46	58.93
Model 5a	14	102.02	-156.04
Model 5b	15	93.09	-136.18
Model 6a	14	2.12	43.75
Model 6b	14	-28.00	104.00
Model 7a	15	-106.21	262.41
Model 7b	19	-91.82	245.65

Table 9.1: Overview of all models presented in this thesis in the order of presentation.

comparable estimates with those found in [24], were obtained. In this thesis estimates are based on a single patient[54], but as Table 8.1 indicates, results are promising.

9.4 Physiological modelling

The state-space method applied in this thesis has allowed a physiological interpretation of the models, opposed to a purely statistical interpretation. Table 9.1 shows the statistical results obtained in chronological order. It is evident, that the statistical best model is not the physiological most correct model, justifying the state-space approach that has been used in this thesis. From a purely statistical point Model 5a is the best model. But if the physiological considerations are included then Model 7b is most reasonable. Here the parameter estimates are sensible, even though there is still room for statistical improvements.

Throughout the presented residual analyses there have been indications that the variance is not constant. Statistically this is a problem but it was assessed that the physiological interpretability was more important. The deficiency has not been prioritised as this thesis is based on grey-box modelling of just one patient.

[54] More patients should be studied in order to obtain reliable statistics.

Chapter 10

Future work

Using stochastic differential equations on state-space form to model patients at rest has given good results and ideas for future work. Unfortunately a limited time period confines the extent of the research. This chapter gives theoretical and practical suggestions for improved modelling of glucose-insulin dynamics.

Using the tool CTSM-R an examination of the initial covariance matrix used to calculate the state prediction in the EKF was conducted. This examination resulted in a change of the calculation method from $\boldsymbol{P}_0 = P_s \int_{t_0}^{t_1} e^{\boldsymbol{A}s} \boldsymbol{\sigma}\boldsymbol{\sigma}^T \left(e^{\boldsymbol{A}s}\right)^T ds$ to $\boldsymbol{P}_0 = k \cdot \boldsymbol{I}$. A substantial theoretical and practical comparative study, should be conducted to clarify consequences and implications on different types of data when using $\boldsymbol{P}_0 = k \cdot \boldsymbol{I}$.

Simplifying insulin kinetics by modelling the flow through two compartments with only one time constant was assumed to be the best solution handling the inverse relationship between two compartments. Expanding on the ideas used in [62] a segregated four compartment model, with two compartments acting as a slow channel and one compartment acting as a fast channel delivering insulin to the plasma compartment could be tried. This compartment structure is more elaborate, and may improve modelling if longer time series become available.

Using the MAP probability approach, assumptions based on literature about device reliability, were made. To fully model device variation a study examining

device properties should be conducted. Several experimental designs should be considered:

1. Using the same device on different patients.
2. Using the same device series on a large variety of patients.

The latter option is available in the DiaCon study giving an excellent point-of-departure for further device investigation.

Quite a bit of work has been used adjusting parameter estimates to physiological correct region of operation. To ease this adjustment a large-scale population study using for example *Population Stochastic Modelling, PSM*[55] [27] should be conducted on all available patients. This way, valuable insight of intra-individual variability could be gained. Using said insight individually tailored treatment comes one step closer.

When expanding to several patients, non-parametric spline modelling should be used to analyse tendencies on a wide range of different CGM devices. Thereby valuable information regarding dynamics of devices can be obtained. Many non-parametric spline models enables a tendency evaluation of dynamics, indicating structure that would otherwise be undetected. Using non-parametric splines combined with multiple CGM devices indicated improvements. This discovery opens exciting possibilities for extension of the system.

Combining identical devices has already been performed, but the idea can be extended to studies with a variety of different CGM devices. As explained in [58] the detection technology can be classified according to the transduction mechanisms used for glucose sensing. Currently, three major areas exist: 1) electrochemical[56] 2) optical 3) piezoelectric. CGM devices can furthermore be divided into three levels of penetration: 1) invasive[57] 2) minimally invasive 3) non-invasive[58].

This thesis has established a reliable PK model system. Focus can therefore now be placed on extensions to the system by including new observation equations of CGM devices using other transduction mechanisms. Theoretically there is no limit to the number of CGM devices that could be placed on a patient:

[55] A package available for R.
[56] As the devices used in this thesis.
[57] As the devices used in this thesis.
[58] For a full schematic overview of degree of penetration and transduction mechanism, please consult [58, p. 1542].

Extension of Equation (3.2) to j CGM devices:

This would be an addition to the output vector $\boldsymbol{y}_j$:

$$\begin{aligned} y_1 &= \alpha_1 x_1 + \beta_1 \\ y_2 &= \alpha_2 x_2 + \beta_2 \\ &\vdots \\ y_{j-1} &= \alpha_{j-1} x_{j-1} + \beta_{j-1} \\ y_j &= \alpha_j x_j + \beta_j \end{aligned} \tag{10.1}$$

The increase of parameters, α_j and β_j, will be met by the drastic increase in data points available for modelling. Each new CGM device will add substantial data compared to the two extra parameters.

Since sensor lag is a well-known problem, a combination of different penetration levels should be studied. This could be combined with a mixture of CGM devices with different transduction mechanisms. Using different transducter mechanisms with different penetration-levels will give a better understanding of glucose movement between tissues. To achieve the maximum amount of information, considerations towards optimal *design of experiments* should be taken on the basis of [13] and [42].

Optical sensors seem ideal for this type of study because monitors on all three kinds of penetration level exist. If optical methods are implemented, sensor selectivity becomes a critical issue [58]. A study examining this issue should be conducted before a larger study is commenced.

Advantages of a multiple sensor approach is not well-investigated but reports describing the advantages exists. Harman-Boehm [19], using non-invasive glucose monitors, increased signal-to-noise ratio in estimations. Castle [10] has found an improvement in accuracy when combining two sensors. They also claim that calibration error, sensor drift and lag can be further minimised by adding extra sensors. Indications of this are also seen in this thesis.

Finally a large effort should be invested in further examination of the current model. Using the model on the same patient a different day or predicting on different patients are just among the possibilities. Furthermore, k-step prediction could be explored. When a solid model has been build simulation studies could be conducted.

Chapter 11

Conclusion

A model using stochastic differential equations on state-space form to describe glucose-insulin dynamics, in resting patients, has been developed, analysed and tested. Parameter estimation was conducted using a Maximum Likelihood approach and the Extended Kalman Filter.

For data analysis, the tool CTSM-R was used, allowing modelling of stochastic continuous time series. A computational deficiency was found in the calculation of the initial covariance matrix. An alternative initial covariance matrix was implemented and used to obtain consistent estimates.

The MVP model was successfully extended to a system of stochastic differential equations on state-space form. Based on initial investigations, the PK part describing insulin kinetics was simplified to accommodate available data, giving a significant improvement in model estimation.

Examining an alternative, more complex model, revealed that increased complexity does not give statistically better models. More detailed physiological knowledge about glucose-insulin dynamics is necessary, when increasing the complexity of the model structure.

One patient was selected for model testing. As no standard solution exists, expansions were divided into three separate parts:

1. A Bayesian approach using MAP probability.
2. An autoregressive approach using Ornstein-Uhlenbeck processes.
3. A combination of the Bayesian and autoregressive approach.

The MAP approach gave improvements from a physiological perspective. Parameters were estimated in areas making physiological sense. Statistical analysis of the residuals indicated that improvements could be achieved implementing autoregressive processes.

Expanding using a purely autoregressive approach, gave - as expected - excellent statistical improvements to parameter estimates. This statistical improvement comes at the cost of physiological interpretability. It is concluded that a purely autoregressive approach cannot stand alone as the interpretability of the parameters is important.

A combined approach was proposed - using MAP probability and Ornstein-Uhlenbeck processes were relevant. It is concluded that physiologically interpretable estimates can be obtained with this combination. By use of non-parametric spline modelling indications about residual structure was obtained. Using different CGM devices to observe the compartment one at a time, gave contradicting results. Combining multiple CGM devices gave statistical improvement, and furthermore similar residual structure. This thesis concludes that an extension to several CGM devices will support the identification of glucose-insulin dynamics. An interesting aspect is the use of different measurement techniques and penetration levels.

Overall the goal set in Chapter 1 is met: the use of stochastic differential equations on state-space form, on patients at rest, is an excellent and flexible method to model glucose-insulin dynamics.

The perspective improvement in life quality of Type 1 Diabetes patients motivates further studies and investigations using SDEs to control a Closed-Loop artificial pancreas.

The results and methods in this thesis are promising and has lead to suggestions for future work presented in the previous chapter.

Appendix A

Model estimation

This appendix presents parameter estimates relevant to the thesis, in the sense that they aid the interpretation of the model estimates.

Table A.1: Model 1: Parameter estimates.

	Estimate	Std. Error	t value	Pr(>|t|)	dF/dPar	dPen/dPar
Isc0	0.100	NA	NA	NA	NA	NA
Ip0	0.025	NA	NA	NA	NA	NA
Ieff0	0.000	NA	NA	NA	NA	NA
Gp0	7.970	NA	NA	NA	NA	NA
Gsc0	7.380	NA	NA	NA	NA	NA
C	1.200	NA	NA	NA	NA	NA
EGP	-2.546	0.577	-4.411	0.000	0.000	0.000
GEZI	-5.450	1.212	-4.496	0.000	0.000	0.000
p2	-4.792	0.364	-13.18	0.000	0.000	0.000
sCGMl	-9.499	1.343	-7.072	0.000	0.000	0.000
Si	-0.790	0.466	-1.697	0.092	0.000	0.000
sig1	-6.863	0.300	-22,91	0.000	0.000	0.000
sig2	-20,00	NA	NA	NA	NA	NA
sig3	-20,00	NA	NA	NA	NA	NA
sig4	-2.077	0.295	-7.051	0.000	0.000	0.000
sig5	-2.669	0.141	-18,87	0.000	0.000	0.000
sins	-13.38	0.554	-24,16	0.000	0.000	0.001
sysi	-1.015	0.360	-2.818	0.006	0.000	0.000
τ_1	50.456	7.168	7.039	0.000	0.000	0.000
τ_2	64.211	6.761	9.497	0.000	0.000	0.000
τ_3	14.891	6.336	2.350	0.020	0.000	0.000
scaleVariance	-17.43	6.611	-2.636	0.009	0.000	0.095

Table A.2: Model 2: Parameter estimates.

	Estimate	Std. Error	t value	Pr(> |t|)	dF/dPar	dPen/dPar
Isc0	0.10	NA	NA	NA	NA	NA
Ip0	0.00	NA	NA	NA	NA	NA
Ieff0	0.00	NA	NA	NA	NA	NA
Gp0	7.97	NA	NA	NA	NA	NA
Gsc0	7.38	NA	NA	NA	NA	NA
BA	0.70	NA	NA	NA	NA	NA
C	1.20	NA	NA	NA	NA	NA
EGP	-2.37	0.54	-4.40	0.00	0.00	0.00
GEZI	-12.29	85.33	-0.14	0.89	0.00	0.00
Si	-0.28	0.57	-0.50	0.62	0.00	0.00
p2	-5.19	0.43	-11.97	0.00	0.00	0.00
sCGMl	-16.14	98.27	-0.16	0.87	0.00	0.00
sig1	-6.80	0.29	-23.13	0.00	0.00	0.00
sig2	-20.00	NA	NA	NA	NA	NA
sig3	-20.00	NA	NA	NA	NA	NA
sig4	-2.03	0.29	-6.97	0.00	0.00	0.00
sig5	-2.68	0.14	-19.27	0.00	0.00	0.00
sins	-13.41	0.54	-25.05	0.00	0.00	0.00
sysi	-1.02	0.40	-2.57	0.01	0.00	0.00
τ_1	73.09	11.03	6.63	0.00	0.00	0.00
τ_3	14.75	6.38	2.31	0.02	0.00	0.00
scaleVariance	-8.09	0.78	-10.36	0.00	0.00	0.00

Table A.3: Model 4: Parameter estimates including MAP on the YSI device and CGM device.

	Estimate	Std. Error	t value	Pr(>|t|)	dF/dPar	dPen/dPar
Isc0	0.10	NA	NA	NA	NA	NA
Ip0	0.03	NA	NA	NA	NA	NA
Ieff0	0.00	NA	NA	NA	NA	NA
Gp0	7.97	NA	NA	NA	NA	NA
Gsc0	7.38	NA	NA	NA	NA	NA
BA	0.70	NA	NA	NA	NA	NA
C	1.20	NA	NA	NA	NA	NA
EGP	-2.99	0.20	-14.77	0.00	0.00	0.00
GEZI	-15.79	120.60	-0.13	0.90	0.00	0.00
Si	-0.98	0.22	-4.53	0.00	0.00	0.00
p2	-4.63	0.17	-27.26	0.00	0.00	0.00
sCGMl	-0.80	0.05	-16.95	0.00	0.00	0.00
sig1	-6.76	0.28	-23.93	0.00	0.00	0.00
sig2	-20.00	NA	NA	NA	NA	NA
sig3	-20.00	NA	NA	NA	NA	NA
sig4	-2.76	0.18	-15.44	0.00	0.00	0.00
sig5	-2.26	0.18	-12.77	0.00	0.00	0.00
sins	-13.32	0.57	-23.54	0.00	0.00	0.00
sysi	-5.33	0.11	-49.20	0.00	0.00	0.00
τ_1	65.57	6.57	9.98	0.00	0.00	0.00
τ_3	42.65	18.76	2.27	0.03	0.00	0.00
scaleVariance	-17.62	3.59	-4.90	0.00	0.00	0.22

Table A.4: Model 5a: Parameter estiamtes with Random Walk.

	Estimate	Std. Error	t value	Pr(> |t|)	dF/dPar	dPen/dPar
Isc0	0.10	NA	NA	NA	NA	NA
Ip0	0.01	NA	NA	NA	NA	NA
Ieff0	0.00	NA	NA	NA	NA	NA
Gp0	7.97	NA	NA	NA	NA	NA
Gsc0	7.38	NA	NA	NA	NA	NA
k10	0.01	0.01	0.76	0.45	0.00	0.00
BA	0.70	NA	NA	NA	NA	NA
C	1.20	NA	NA	NA	NA	NA
EGP	-3.01	0.33	-9.19	0.00	0.00	0.00
GEZI	-15.60	85.67	-0.18	0.86	0.00	0.00
Si	-1.00	0.33	-3.01	0.00	0.00	0.00
p2	-4.57	0.32	-14.51	0.00	0.00	0.00
sCGMl	-16.82	53.00	-0.32	0.75	0.00	0.00
sig1	-6.78	0.30	-22.68	0.00	0.00	0.00
sig2	-20.00	NA	NA	NA	NA	NA
sig3	-20.00	NA	NA	NA	NA	NA
sig4	-2.68	0.21	-12.75	0.00	0.00	0.00
sig5	-8.63	34.90	-0.25	0.81	0.00	0.00
sig6	-2.61	0.09	-29.71	0.00	0.00	0.00
sins	-13.32	0.62	-21.44	0.00	0.00	0.00
sysi	-17.15	45.78	-0.38	0.71	0.00	0.00
τ_1	68.71	11.13	6.17	0.00	0.00	0.00
τ_3	13.57	6.07	2.24	0.03	0.00	0.00
scaleVariance	-9.53	0.89	-10.69	0.00	0.00	0.00

Table A.5: Model 7b: Parameter estimates with AR(1) process and two CGM devices.

Estimate	Std. Error	t value	Pr(>|t|)	dF/dPar	dPen/dPar	
Isc0	0.100	NA	NA	NA	NA	NA
Ip0	0.025	NA	NA	NA	NA	NA
Ieff0	0.000	NA	NA	NA	NA	NA
Gp0	7.970	NA	NA	NA	NA	NA
Gsc0	7.380	NA	NA	NA	NA	NA
x10	0.100	NA	NA	NA	NA	NA
x20	0.100	NA	NA	NA	NA	NA
BA	0.700	NA	NA	NA	NA	NA
C	1.200	NA	NA	NA	NA	NA
EGP	-2.976	0.196	-15,183	0.000	0.000	0.000
GEZI	-15,741	119,026	-0.132	0.895	0.000	0.002
Si	-0.964	0.209	-4.611	0.000	0.000	0.000
a	-3.256	3.417	-0.953	0.342	0.000	0.000
b	-4.195	0.796	-5.269	0.000	0.000	0.000
p2	-4.631	0.162	-28,534	0.000	0.000	0.000
q	0.920	0.011	83.952	0.000	0.000	0.000
r	0.891	0.037	23.882	0.000	0.000	0.000
sCGMl	-0.815	0.047	-17,511	0.000	0.000	0.000
sCGMr	-0.786	0.048	-16,537	0.000	0.000	0.000
sig1	-6.752	0.282	-23,933	0.000	0.000	0.000
sig2	-20	NA	NA	NA	NA	NA
sig3	-20	NA	NA	NA	NA	NA
sig4	-2.748	0.175	-15,705	0.000	0.000	0.000
sig5	-8.529	269,041	-0.032	0.975	0.000	0.000
sig6	-8.786	34.290	-0.256	0.798	0.000	0.000
sig7	-2.229	0.261	-8.542	0.000	0.000	0.000
sins	-13.227	0.533	-24,815	0.000	0.000	0.001
sysi	-5.329	0.115	-46,394	0.000	0.000	0.000
τ_1	65.710	6.561	10.015	0.000	0.000	0.000
τ_3	9.588	3.524	2.721	0.007	0.0007 0.000	
scaleVariance	-17.616	3.291	-5.352	0.000	0.000	0.215

Appendix B

Residual plots

This appendix presents the residual analysis relevant for the thesis, in the sense that they aid the interpretation of the model estimates.

The residuals shown in this appendix are primarily with models having diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$ included. If another combination is used, this is stated in the caption of the respective model.

Figure B.1: Model 1: with negligible diffusion terms

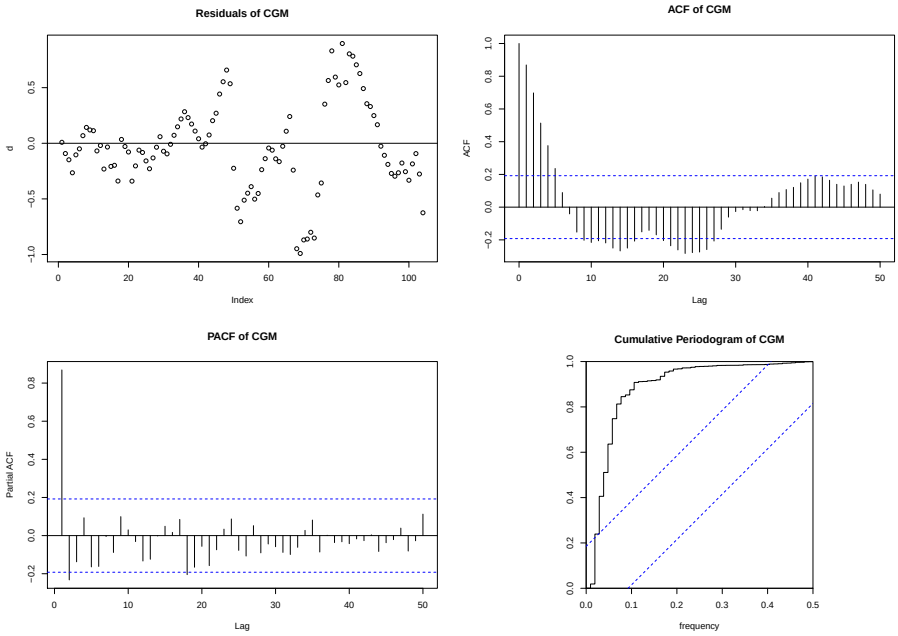


Figure B.2: Model 1.

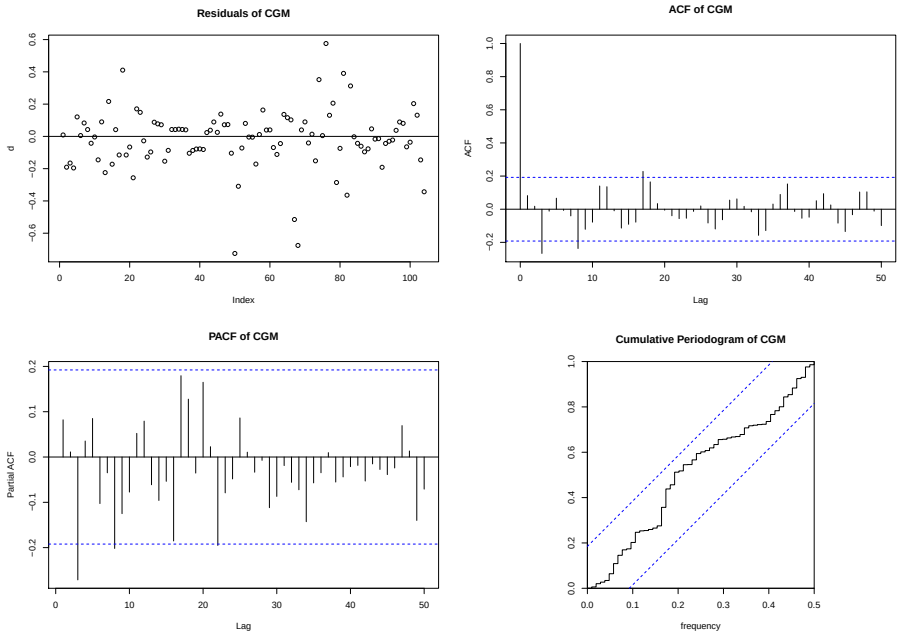


Figure B.3: Model 2.

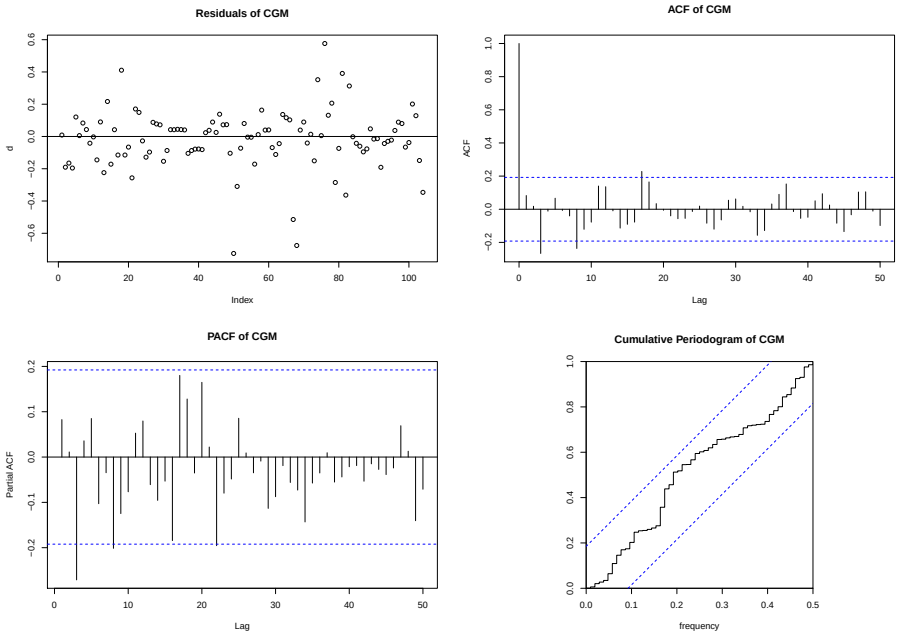


Figure B.4: Model 4: including MAP on YSI device and CGM device.

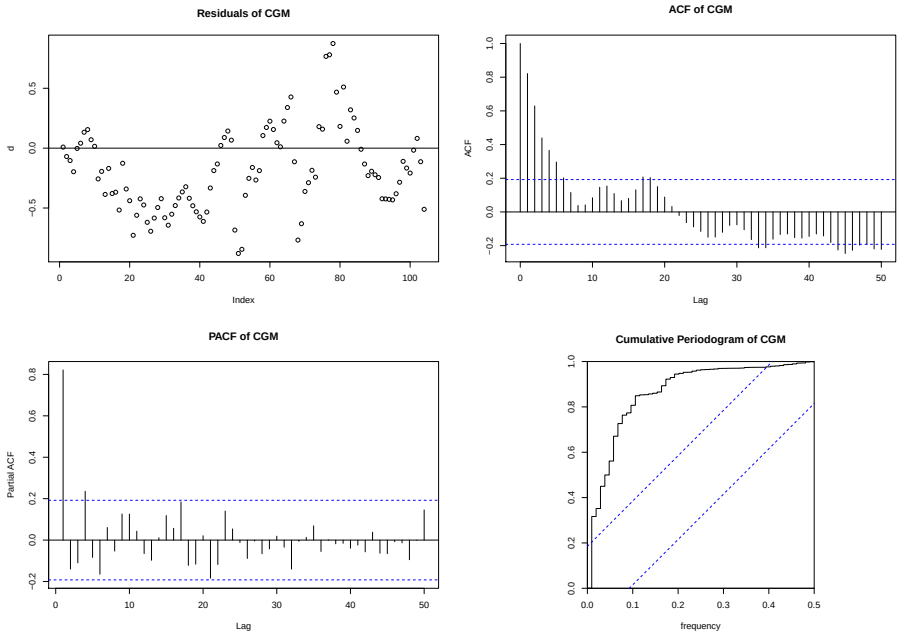


Figure B.5: Model 5a: with Random Walk.

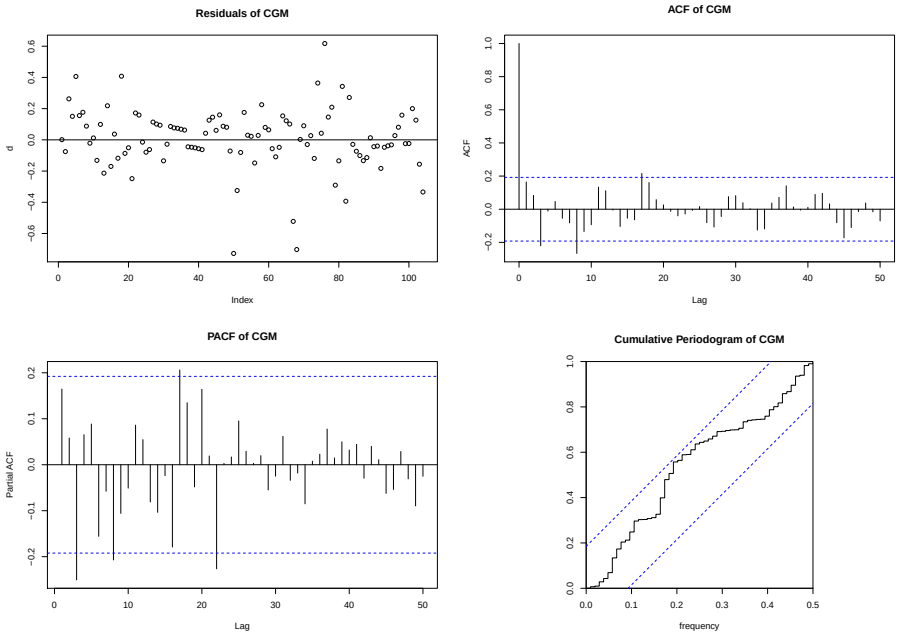


Figure B.6: Model 5b: with AR(1)-process.

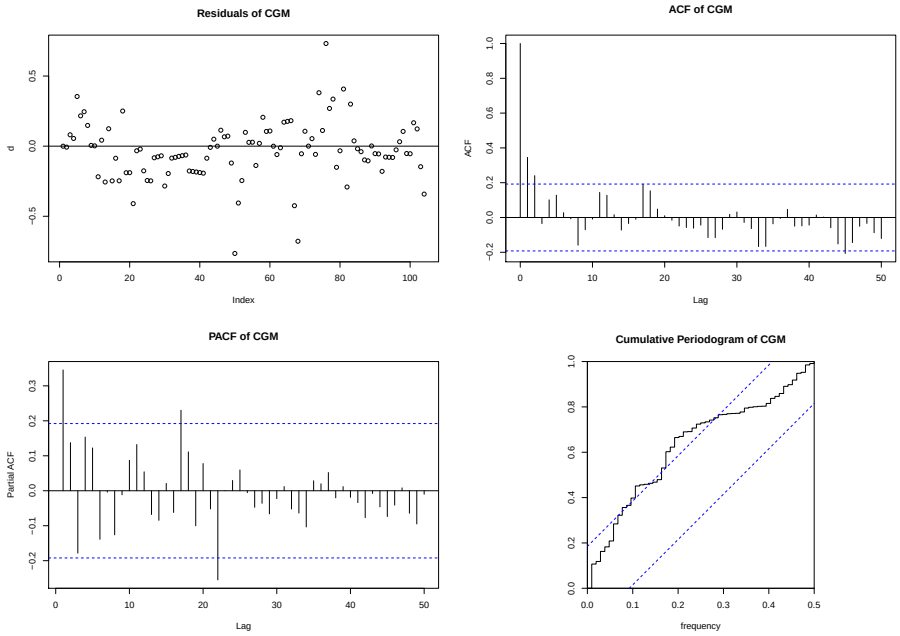


Figure B.7: Model 6a: MAP and Random Walk using original CGM device.

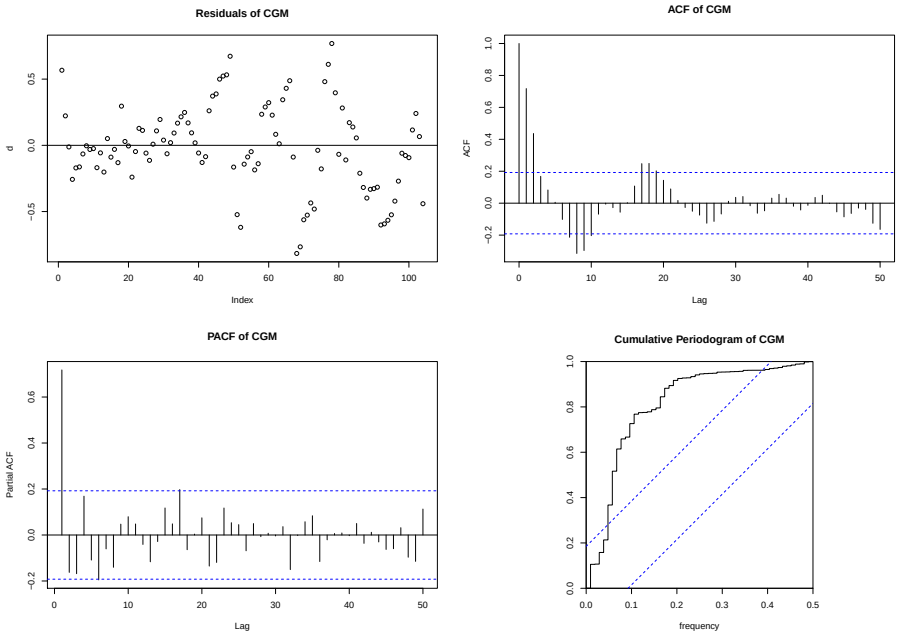


Figure B.8: Model 6b: MAP and Random Walk using alternate CGM device.

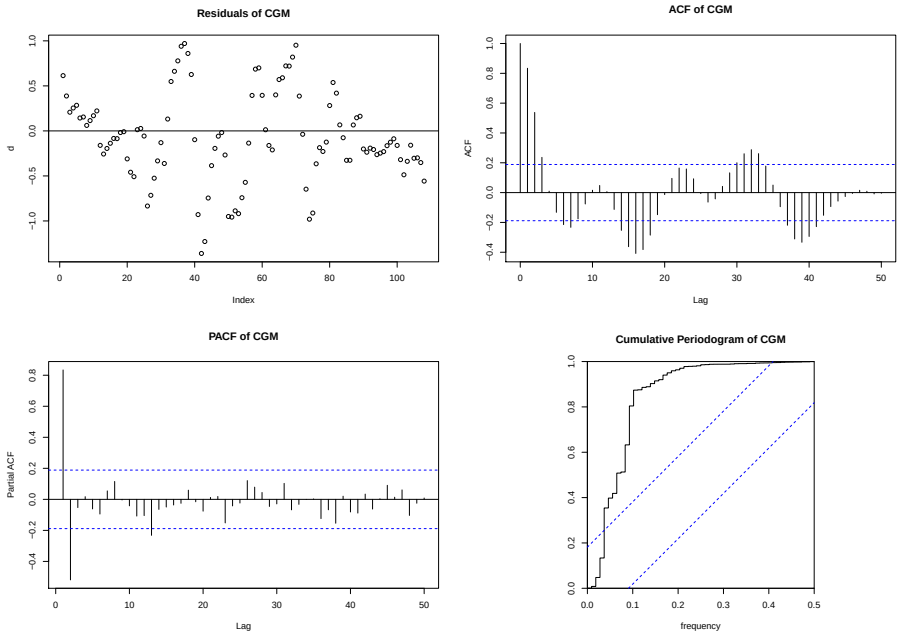


Figure B.9: Model 7a: MAP and Random Walk using two CGM devices. Residual analysis for original CGM device.

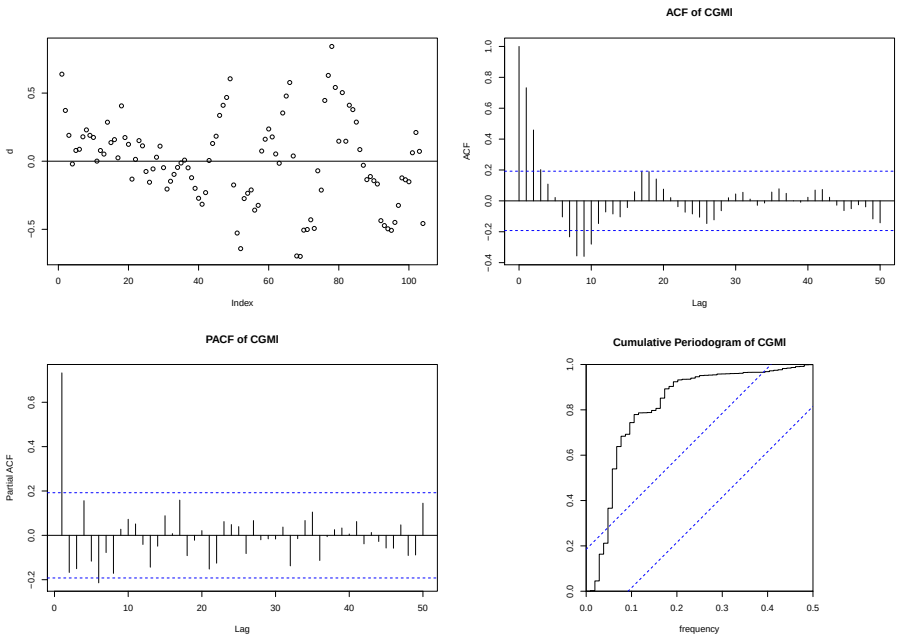


Figure B.10: Model 7a: MAP and Random Walk using two CGM devices. Residual analysis for alternate CGM device.

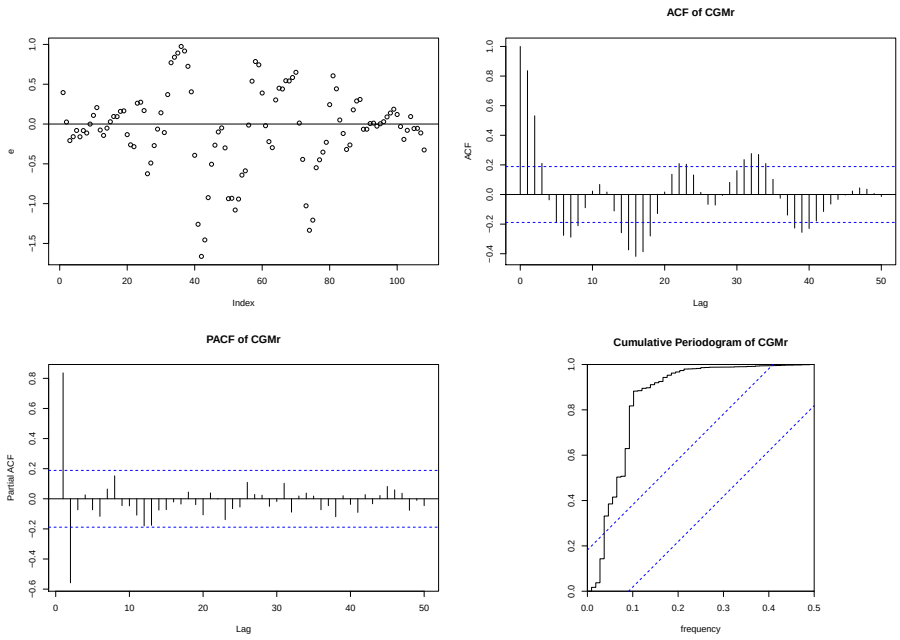


Figure B.11: Model 7b: MAP and AR(1)-process using two CGM devices. Residual analysis for original CGM device.

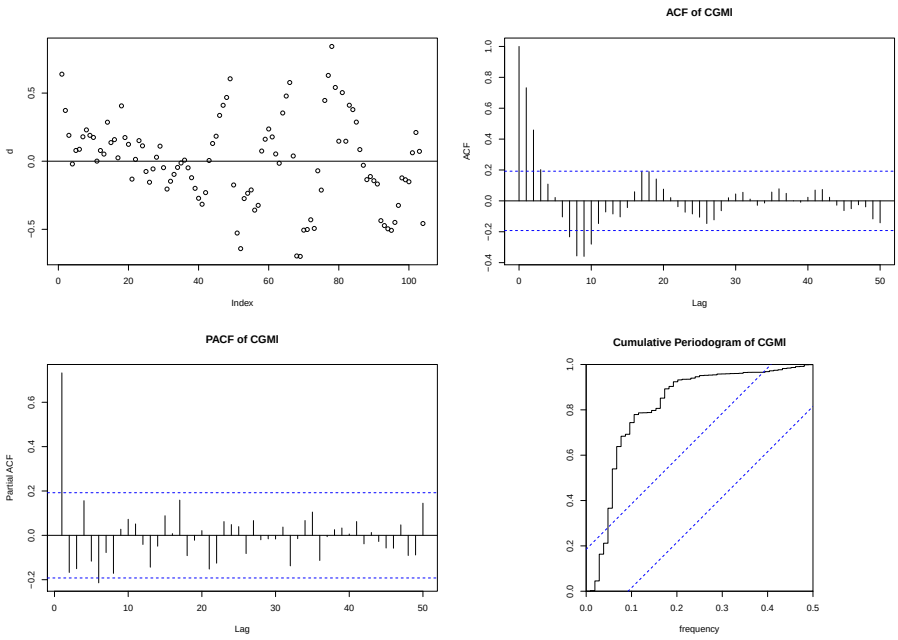
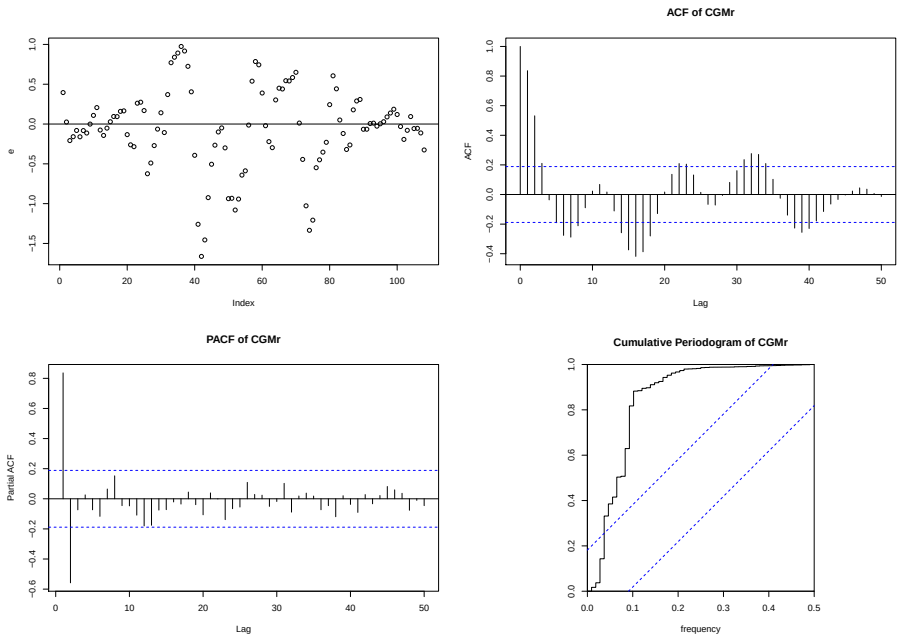


Figure B.12: Model 7b: MAP and AR(1)-process using two CGM device. Residual analysis for alternate CGM devices.



Appendix C

Prediction plots

This appendix presents predictions relevant for the thesis, in the sense that they aid the interpretation of the model estimates.

The predictions shown in this appendix are primarily with models having diffusion terms $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_4}$ and $\boldsymbol{\sigma_5}$ included. If another combination is used, this is stated in the caption of the respective model.

Figure C.1: Model 1 with negligible diffusion terms: Prediction of plasma insulin

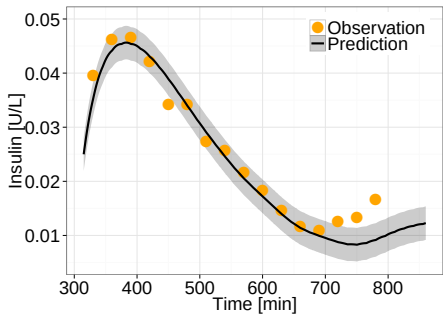


Figure C.2: Model 1: Prediction of plasma insulin.

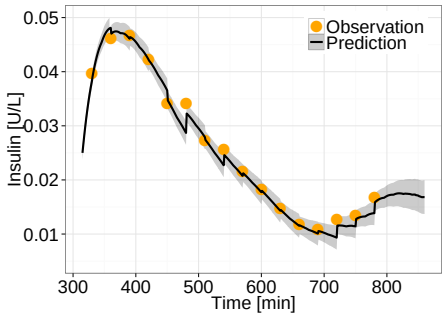


Figure C.3: Model 6a: with MAP and Random Walk on original CGM device. Prediction of YSI observations and CGM observations.

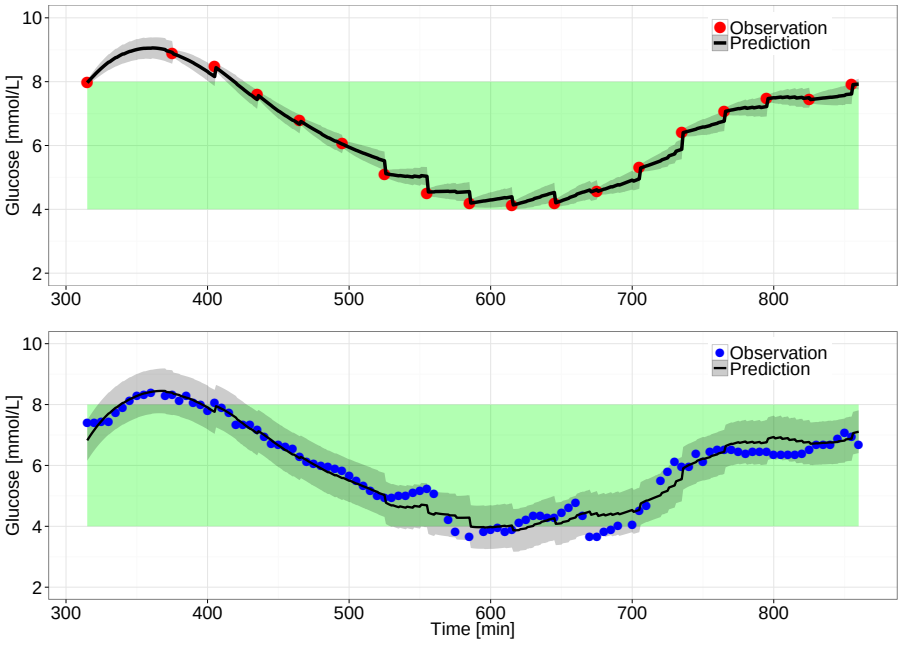
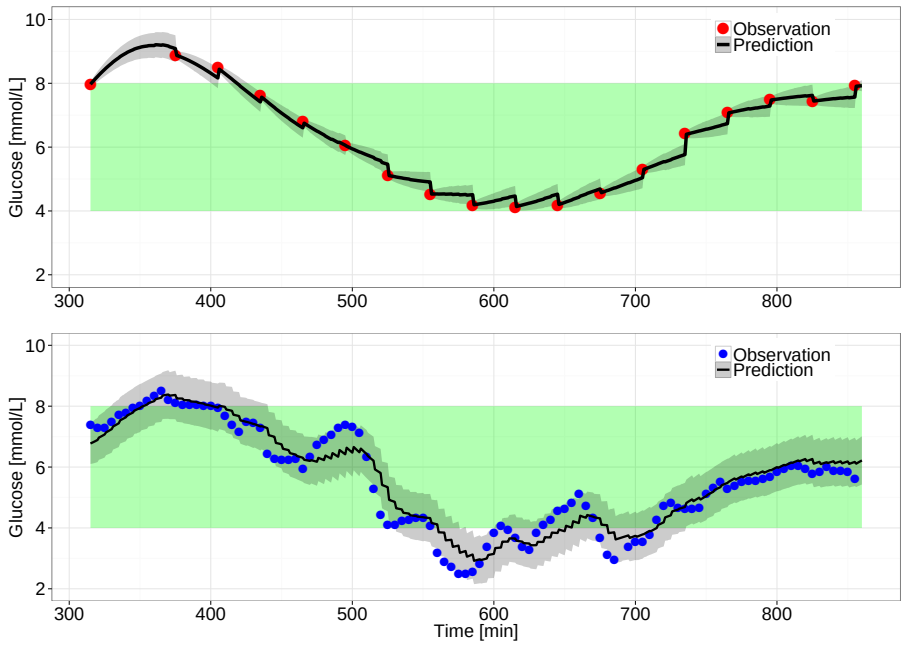


Figure C.4: Model 6b: with MAP and Random Walk on alternate CGM device. Prediction of YSI observations and CGM observations.



Bibliography

[1] A.M. Albisser, B.S. Leibel, T.G. Ewart, Z. Davidovac, C.K. Botz, and W. Zingg. An artificial endocrine pancreas. *Diabetes*, 23(5):389–396, May 1974.

[2] P. Bacher and H. Madsen. Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings*, 43(7):1511 – 1522, 2011.

[3] B.W. Bequette. Continuous glucose monitoring: real-time algorithms for calibration, filtering, and alarms. *J Diabetes Sci Technol*, 4(2):404–418, Mar 2010.

[4] R.N. Bergman, L.S. Phillips, and C. Cobelli. Physiologic evaluation of factors controlling glucose tolerance in man: measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose. *J Clin Invest*, 68(6):1456–1467, Dec 1981.

[5] C. Bernard. Lecons sur les phénomènes de la vie communs aux animaux et aux végétaux. *Librairie J.-B. Baillière et fils*, 1878.

[6] T. Bjørk. *Arbitrage theory in continuous time*. Oxford university press, Oxford, 1998.

[7] A. Breinholt, J.K. Møller, H. Madsen, and Mikkelsen P.S. A formal statistical approach to representing uncertainty in rainfall runoff modelling with focus on residual analysis and probabilistic output evaluation a " distinguishing simulation and prediction. *Journal of Hydrology*, 472,473(0):36 – 52, 2012.

[8] M. Breton and B. Kovatchev. Analysis, modeling, and simulation of the accuracy of continuous glucose sensors. *J Diabetes Sci Technol*, 2(5):853–862, Sep 2008.

[9] P.J. Brockwell and R.A. Davis. *Time Series: Theory and Methods*. Springer Series in Statistics. Springer, 2009.

[10] J.R Castle and W.K. Ward. Amperometric glucose sensors: Sources of error and potential benefit of redundancy. *Journal of Diabetes Science and Technology*, 4, 2010.

[11] C. Cobelli and A. Mari. Control of diabetes with artificial systems for insulin delivery algorithm independent limitations revealed by a modeling study. *Biomedical Engineering, IEEE Transactions on*, BME-32(10):840–845, Oct.

[12] A.K Duun-Henriksen, S. Schmidt, R.M. Røge, , J.B. Møller, K. Nørgaard, J.B. Jørgensen, and H. Madsen. Model identification using stochastich differential equation grey-box models in diabetes. *J Diabetes Sci Technol*, 7, March 2013.

[13] V. Fedorov. *Theory of Optimal Experiments*. Academic Press, 1972.

[14] M.E. Fisher. A semiclosed-loop algorithm for the control of blood glucose levels in diabetics. *Biomedical Engineering, IEEE Transactions on*, 38(1):57 –61, jan. 1991.

[15] R.A. Fisher. On the mathematical foundations of theoretical statistics. *Philosopical Transactions of the Royal Society of London, Series A*, 222:309–368, 1922.

[16] J. Gabrielsson and D. Weiner. *Pharmacokinetic / Pharmacodynamic Data Analysis: Concepts and Applications*. Apotekarsocieteten, 2000.

[17] P. Giabardo, M. Zugno, P. Pinson, and H. Madsen. Feedback, competition and stochasticity in a day ahead electricity market. *Energy Economics*, 32(2):292 – 301, 2010.

[18] V. Gontis, J. Ruseckas, and A. Kononovius. A long-range memory stochastic model of the return in financial markets. *Physica A: Statistical Mechanics and its Applications*, 389(1):100 – 106, 2010.

[19] I. Harman-Boehm, A. Gal, A.M. Raykhman, E. Naidis, and Y. Mayzel. Noninvasive glucose monitoring: increasing accuracy by combination of multi-technology and multi-sensors. *J Diabetes Sci Technol*, 4(3):583–595, May 2010.

[20] A.C. Hindmarsh. ODEPACK, a systematized collection of ODE solvers , r. s. stepleman et al. (eds.), North-Holland, amsterdam, (vol. 1 of), pp. 55-64. *IMACS Transactions on Scientific Computation*, 1:55–64, 1983.

[21] R. Hovorka, V. Canonico, L.J. Chassin, U. Haueter, M. Massi-Benedetti, M.O. Federici, T.R. Pieber, H.C. Schaller, L. Schaupp, T. Vering, and M.E. Wilinska. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiol Meas*, 25(4):905–920, Aug 2004.

[22] Møller J.K., H. Madsen, and Carstensen J. Parameter estimation in a simple stochastic differential equation for phytoplankton modelling. *Ecological Modelling*, 222(11):1793 – 1799, 2011.

[23] R. Juhl, N. R. Kristensen, P. Bacher, J. Kloppenborg, and H. Madsen. *CTSM-R User Guide*. Technical University of Denmark, 2.3 edition, November 2012.

[24] S.S. Kanderian, S. Weinzimer, G. Voskanyan, and G.M. Steil. Identification of intraday metabolic profiles during closed-loop glucose control in individuals with type 1 diabetes. *J Diabetes Sci Technol*, 3(5):1047–1057, Sep 2009.

[25] D.B. Keenan, J.J. Mastrototaro, G. Voskanyan, and G.M. Steil. Delays in minimally invasive continuous glucose monitoring devices: a review of current technology. *J Diabetes Sci Technol*, 3(5):1207–1214, Sep 2009.

[26] A.I.. Khan, Y. Vasquez, J. Gray, F.H. Wians Jr, and M.H. Kroll. The variability of results between point-of-care testing glucose meters and the central laboratory analyzer. *Archives of Pathology & Laboratory Medicine*, 130(10):1527–1532, 2013/02/26 2006.

[27] S. Klim, S.B. Mortensen, N.R. Kristensen, R.V. Overgaard, and H. Madsen. Population stochastic modelling (psm)" an r package for mixed-effects models based on stochastic differential equations. *Computer Methods and Programs in Biomedicine*, 94(3):279 – 289, 2009.

[28] B.P. Kovatchev, M. Breton, C.D. Man, and C. Cobelli. In silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes. *J Diabetes Sci Technol*, 3(1):44–55, Jan 2009.

[29] N.R. Kristensen, Madsen H., and S.B. Jørgensen. A method for systematic improvement of stochastic grey-box models. *Computers & Chemical Engineering*, 28(8):1431 – 1449, 2004.

[30] N.R. Kristensen and H. Madsen. *Continuous Time Stochastic Modelling - Mathematics Guide*. Technical University of Denmark, 2.3 edition, December 2003.

[31] E. Kulcu, J.A. Tamada, G. Reach, R.O. Potts, and M.J. Lesho. Physiological differences between interstitial glucose and blood glucose measured in human subjects. *Diabetes Care*, 26(8):2405–2409, 2003.

[32] C. Lodi, P. Bacher, J. Cipriano, and H. Madsen. Modelling the heat dynamics of a monitored test reference environment for building integrated photovoltaic systems using stochastic differential equations. *Energy and Buildings*, 50(0):273 – 281, 2012.

[33] S.M. Lynch and B.W. Bequette. Model predictive control of blood glucose in type i diabetics using subcutaneous glucose measurements. In *American Control Conference, 2002. Proceedings of the 2002*, volume 5, pages 4039 – 4043 vol.5, 2002.

[34] H. Madsen. *Time Series Analysis*. Chapman & Hall/CRC Texts in Statistical Science Series. Chapman & Hall/CRC, 2008.

[35] H. Madsen and J. Holst. Modelling non-linear and non-stationary time series. Unpublished book about the modelling of non-linear and non-stationary time series, 2006.

[36] H. Madsen and P. Thyregod. *Introduction to General and Generalized Linear Models*. Chapman & Hall/CRC Texts in Statistical Science Series. CRC PressINC, 2011.

[37] J.B. Møller, R.V. Overgaard, H. Madsen, T. Hansen, Pedersen O., and S.H. Ingwersen. Predictive performance for population models using stochastic differential equations applied on data from an oral glucose tolerance test. *Journal of Pharmacokinetics and Pharmacodynamics*, 37(1):85–98, 2010.

[38] J.K. Møller, Kirsten Riber Bergmann, Lasse Engbo Christiansen, and Henrik Madsen. Development of a restricted state space stochastic differential equation model for bacterial growth in rich media. *Journal of Theoretical Biology*, 305(0):78 – 87, 2012.

[39] H.A. Nielsen and Madsen H. A generalization of some classical time series tools. *Computational Statistics & Data Analysis*, 37(1):13–31, 2001.

[40] B Øksendal. *Stochastic Differential Equations: An Introduction with Applications (Universitext)*. Springer, 6th edition, September 2010.

[41] R.S. Parker, F.J. Doyle, and N.A. Peppas. The intravenous route to blood glucose control. *Engineering in Medicine and Biology Magazine, IEEE*, 20(1):65–73, Jan/Feb.

[42] A. Pázman. *Foundations of optimum experimental design*. D. Reidel Boston, 1986.

[43] O.B. Pedersen and Beck-Nielsen et al. *Medicinsk Kompendium (II)*. Nyt Nordisk Forlag Arnold Busck, 17th edition, 2009.

[44] E.F. Pfeiffer, C. Thum, and A.H. Clemens. The artificial beta cell–a continuous control of blood sugar by external regulation of insulin infusion (glucose controlled insulin infusion system). *Horm Metab Res*, 6(5):339–342, Sep 1974.

[45] J.C. Pickup and G. Williams. *Textbook of Diabetes: 2 Volume Set*. Textbook of Diabetes. John Wiley & Sons, 2002.

[46] F.Z. Poleto, C.D. Paulino, G. Molenberghs, and J.M. Singer. Inferential implications of over-parametrization: A case study in incomplete categorical data. *International Statistical Review*, 79(1):92–113, 2011.

[47] D. Pollard. *Convergence of stochastic processes*. Springer Series in Statistics. Springer-Verlag, New York, 1984.

[48] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0.

[49] C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning*. Adaptive Computation And Machine Learning. Mit Press, 2006.

[50] M. Rowland and T.N. Tozer. *Clinical Pharmacokinetics: Concepts and Applications*. Lea & Febinger Book. Lea & Febiger, 1995.

[51] W. Rudin. *Real and complex analysis*. McGRAW-HILL, 1987.

[52] S. Schmidt, D. Boiroux, A.K. Duun-Henriksen, L. Frøssing, O. Skyggebjerg, J.B. Jørgensen, N.K. Poulsen, H. Madsen, S. Madsbad, and K. Nørgaard. Model-based closed-loop glucose control in type 1 diabetes – the diacon experience. *Submitted*, 2013.

[53] R.R. Seeley, T.D. Stephens, and P. Tate. *Anatomy and Physiology*. McGraw-Hill, 2004.

[54] G.M. Steil, K. Rebrin, F. Hariri, S. Jinagonda, S. Tadros, C. Darwin, and M.F. Saad. Interstitial fluid glucose dynamics during insulin-induced hypoglycaemia. *Diabetologia*, 48:1833–1840, 2005.

[55] G.M. Steil, K. Rebrin, J. Mastrototaro, B. Bernaba, and M.F. Saad. Determination of plasma glucose during rapid glucose excursions with a subcutaneous glucose sensor. *Diabetes Technol Ther*, 5(1):27–31, 2003.

[56] C.W. Tornoe, R.V. Overgaard, H. Agerso, H.A. Nielsen, H. Madsen, and N.E. Jonsson. Stochastic differential equations in nonmem: implementation, application, and comparison with ordinary differential equations. *Pharm Res*, 22(8):1247–1258, Aug 2005.

[57] G. E. Uhlenbeck and L. S. Ornstein. On the theory of the brownian motion. *Phys. Rev.*, 36:823–841, Sep 1930.

[58] S. Vaddiraju, D.J. Burgess, I. Tomazos, F.C. Jain, and F. Papadimitrakopoulos. Technologies for continuous glucose monitoring: current problems and future promises. *J Diabetes Sci Technol*, 4(6):1540–1562, Nov 2010.

[59] G. Welch and G. Bishop. An introduction to the kalman filter. *Department of Computer Science, University of North Carolina*, 2006.

[60] S. Wild, G. Roglic, A.s Green, R. Sicree, and H. King. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care*, 27(5):1047–1053, May 2004.

[61] M.E. Wilinska, L.J. Chassin, C.L. Acerini, J.M. Allen, D.B. Dunger, and R. Hovorka. Simulation environment to evaluate closed-loop insulin delivery systems in type 1 diabetes. *J Diabetes Sci Technol*, 4(1):132–144, Jan 2010.

[62] M.E. Wilinska, L.J. Chassin, H.C. Schaller, L. Schaupp, T.R. Pieber, and R. Hovorka. Insulin kinetics in type-1 diabetes: continuous and bolus delivery of rapid acting insulin. *Biomedical Engineering, IEEE Transactions on*, 52(1):3–12, Jan.

[63] K. Zierler. Whole body glucose metabolism. *Am J Physiol*, 276(3 Pt 1):409–426, Mar 1999.