

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

PHD IN AUTOMATION, ROBOTICS AND COMPUTER SCIENCE  
FOR INDUSTRY

DOCTORAT EN AUTOMÀTICA, ROBÒTICA I INFORMÀTICA INDUSTRIAL

## PhD Dissertation

# *Uncertainty in Postprandial Model Identification in type 1 Diabetes*

Author: Alejandro José Laguna Sanz

Directores: Jorge Bondía Company

Paolo Rossetti

DEPARTMENT OF CONTROL AND SYSTEMS ENGINEERING  
DEPARTAMENT D'ENGINYERIA DE SISTEMES I AUTOMÀTICA

VALENCIA, OCTOBER 8, 2013



## **Abstract**

Individual patient characterization is key in the creation of a closed loop system for glucose homeostasis. This identification of patients is currently performed by physicians using heuristic rules based on experimentation over periodical clinical visits to the patient.



# Contents

<b>I state of the art</b>	<b>1</b>
<b>1 Models for diabetes</b>	<b>2</b>
1.1 Insulin absorption models . . . . .	5
1.1.1 UVA model . . . . .	6
1.1.2 Cambridge model . . . . .	8
1.2 Glucose absorption models . . . . .	11
1.2.1 UVA model . . . . .	12
1.2.2 Cambridge model . . . . .	15
1.3 Endogenous models . . . . .	16
1.3.1 Bergman model . . . . .	17
1.3.2 Panunzi model . . . . .	18
1.3.3 UVA model . . . . .	20
1.3.4 Cambridge model . . . . .	23
1.4 Critical selection of models . . . . .	26
<b>2 Identification in Diabetes</b>	<b>35</b>

2.1	Patient Identification . . . . .	36
2.2	Experiment Design for Artificial Pancreas . . . . .	41
2.3	Uncertainty and Interval Identification in Diabetes . . . . .	46
<b>3</b>	<b>Parameter Estimation</b>	<b>51</b>
3.1	Optimization of Identifiability . . . . .	56
3.2	Interval Identification . . . . .	61
3.2.1	Error-bounded Estimation . . . . .	61
3.3	Optimization and software . . . . .	66
3.3.1	Scatter search for Matlab . . . . .	67
3.3.2	Covariance Matrix Adaptive Estimation . . . . .	67
3.3.3	$\epsilon$ - MOGA Evolutionary Algorithm . . . . .	68
<b>4</b>	<b>Conclusions</b>	<b>71</b>
<b>II</b>	<b>Enhancing CGM Experiments</b>	<b>76</b>
<b>5</b>	<b>Optimal Experiment Design</b>	<b>77</b>
5.1	Experiments designed with Bergman's model . . . . .	80
5.2	Experiments designed with modified Panunzi's model . . . . .	91
5.3	Discussion and clinical protocol . . . . .	101
<b>6</b>	<b>CGM Statistical Modeling and Validation</b>	<b>105</b>
6.1	Data and methodology . . . . .	106

6.2	CGM Modelling . . . . .	109
6.2.1	Analysis of delay . . . . .	111
6.2.2	Analysis of Stationarity . . . . .	112
6.2.3	Distribution fitting . . . . .	116
6.3	Validation . . . . .	119
<b>7</b>	<b>Conclusions</b>	<b>125</b>
<b>III Interval identification</b>		<b>128</b>
<b>8</b>	<b><i>In silico</i> identification</b>	<b>129</b>
8.1	Multiobjective Optimization Set-up . . . . .	130
8.2	Identification from reference glucose . . . . .	136
8.3	Identification from CGM . . . . .	136
8.4	Discussion . . . . .	146
<b>9</b>	<b>Pilot Experimental identification</b>	<b>149</b>
9.1	Optimization and index definition . . . . .	150
9.2	Identification from reference glucose . . . . .	158
9.3	Best-case permutation . . . . .	167
9.4	Discussion . . . . .	170
<b>10</b>	<b>Experimental Interval Identification</b>	<b>172</b>
10.1	Optimization settings and parameters . . . . .	172

10.2 Identification from reference glucose . . . . .	177
10.3 Identification from CGM . . . . .	188
10.4 Discussion . . . . .	198
<b>11 Conclusions</b>	<b>200</b>

# **Part I**

## **state of the art**

# Chapter 1

## Models for diabetes

One of the main problems for glucose control is the insufficient accuracy of existing mathematical models for describing the physiology of the glucoregulatory system. In this chapter the modeling context for the artificial pancreas will be reviewed, and the state of the art of mathematical models in literature will be described.

A mathematical model is characterized both by its objective and its structure. The main objective in diabetic patient models is of course to be able to reproduce the patient's metabolism from a clinical point of view. In the artificial pancreas context, models must be useful from the control point of view. According to the structure of a model, quoting the work of Walter and Pronzato [81], the main distinction to be made is whether it is *phenomenological* or *behavioral*. A phenomenological model is a model based on prior knowledge about a physical or, in the case of the artificial pancreas, physiological principles. This kind of processes are often called *knowledge-based* models as opposed to behavioral models, which merely approximate the observed behavior of the output without any prior knowledge of the process. Behavioral models are focused on data reproduction, and not at all in the process behind, while phenomenological models only use the data to adjust the parameters, while the structure is determined by the process itself.

Examples of phenomenological models are:

- Chemical reactions. Biological reactions
- Systems of force equilibrium.

- Models of deposit systems.
- Electromagnetic models. Electrical engines.

Examples of behavioral models are:

- ARX/ARMAX models.
- Polynomial models.
- Random models.

The behavioral models exposed do not have the purpose of imitating any experimental process, and are “all-purpose” models that have to be adjusted to any particular system. On the contrary, the phenomenological models listed are specific of the process they represent. Table 1.1 shows a comparison of the differences between both modeling paradigms. In the context of the artificial pancreas almost every model published is phenomenological, even the simplest ones, due to the vast knowledge of the physiology available from the physicians. Behavioral models have also been used to characterize diabetic patients behavior, mostly in proof-of-concept studies like that performed by [77].

	Phenomenological models	Behavioral models
Parameters	have a concrete meaning	have no concrete meaning
Simulation	long and difficult	quick and easy
Prior information	taken into account	neglected
Validity domain	large (if structure is correct)	restricted

Table 1.1: Phenomenological and behavioral models as seen by Walter and Pronzato [81]

Usually, phenomenological models tend to be complex and highly nonlinear. Linearizing a phenomenological model changes its aim and its nature. When a nonlinear phenomenological model is designed, its aim is usually a better understanding of the system through simulation. Reasons for linearizing are usually attempts to control, or design of better controllers, but this transformation neglects the prior information of the system and its complexity, so the linearized model results in a behavioral model, with a restricted validity domain and loss of the experimental meaning of the parameters. A tradeoff between model complexity and accuracy is also imposed by the individualization of the model to a specific patient. Excessively complex models may struggle when

fitting to an individual due to limitations of data available in the domiciliary context yielding to lack of identifiability of the model. Loss of identifiability hinders parameter interpretation of the identified model, which is an important goal for physiology-based phenomenological models of diabetes.

Physiological models of the glucose-insulin system for type 1 diabetes involve three main sub-processes:

- **Insulin absorption model.** This model represents the insulin input to the organism. This involves insulin pharmacokinetics, diffusion through different tissues and natural insulin degradation. Insulin absorption depends on the type of insulin used for the therapy and the interface used for its delivery. Insulin is injected or infused in the subcutaneous tissue, delaying its appearance in plasma compared to insulin secretion by the pancreas. In case multiple daily injections are used, pharmacokinetics of both rapid-acting and long-acting insulin have to be considered. In case of insulin pump (artificial pancreas) only rapid-acting insulin is used.
- **Glucose absorption model.** Food glucose input is represented in this model, which is also called the gastrointestinal model. It involves the process of ingestion, digestion and absorption from the intestine into blood of glucose and other nutrients. The nutritional composition of the meal influences the process of digestion, affecting the flow of carbohydrates through the gut.
- **Glucoregulatory model.** The internal regulation of glucose is represented by this model. The transformation of glycogen to glucose by the liver (hepatic glucose production) and glucose uptake by peripheral tissues, the influence of different hormones in blood glucose, insulin independent consumption of glucose, and in summary, every effect that, in the organism, can affect the concentration in glucose. The models representing all these physiological processes and relations tend to be really complex, and it is really common to disregard some of the influences on the glucose concentration, so that the model becomes simpler and for other purposes than simulation.

The insulin absorption and gastrointestinal models are usually considered as the *input models* for the glucose-insulin system, for they characterize the two main exogenous (coming from outside the body) inputs into blood that influence the glucose concentration (figure 1.1). Several model reviews can be found in literature, being the most notorious Willinska's review [84] and Nucci and Cobelli review [61].

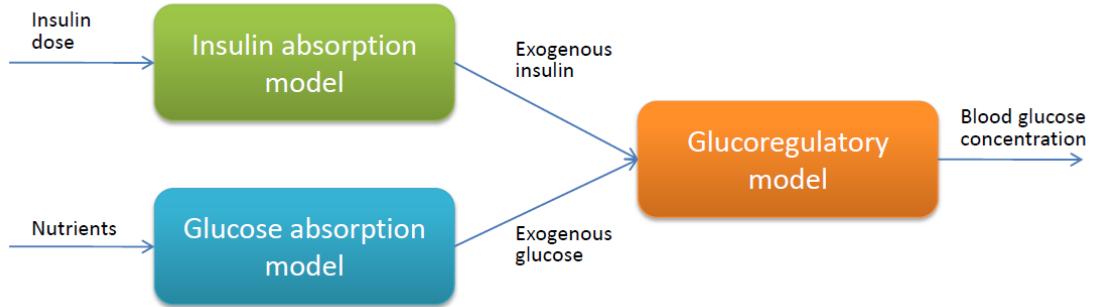


Figure 1.1: Glucose-insulin system and its sub-processes

Subcutaneous insulin injection or pump delivery is the main control action used to counteract the largest disturbance of the system, meal ingestion. Insulin pharmacokinetics has been studied for a long time, and the behavior of insulin analogues is well documented in literature. Complex glucose absorption models have been developed in the last years with increasing accuracy, but the physiological behavior of the digestive and intestinal absorption processes during a mixed meal ingestion, which is almost every meal a normal person has, is yet to be represented by a gastrointestinal model. The main difficulty in characterization of the gastrointestinal model is that glucose absorption is only measurable with tracer methods [2], but few studies have been done in this area so far. The only experiments done in this area were not performed using real mixed meals, but instead the patients ate marked jelly with eggs and bacon. The influence of the nutritional composition is very relevant in the final model output, and it has not been taken in consideration so far.

Several models for the systems described will be shown in the next sections, and later a critical review of the usefulness of these models will be performed in order to narrow down the scope of the research, not to forget that the last objective of this thesis is identification of postprandial models for control.

## 1.1 Insulin absorption models

Modeling of subcutaneous (s.c.) insulin diffusion and absorption dates back to the 80's. Kobayashi *et al.* published in 1983 a model based on a one-compartmental delay

differential equation for U40 Actrapid insulin [46] on type 1 and type 2 diabetic patients. Models proposed in literature have been increasing in complexity since then, like the two-compartmental model proposed by Kraegen *et al.* in 1984 [50], or the model considering two different insulin pathways proposed by [67] in 1995. Increasing in complexity, Mosekilde *et al.* proposed a model based on partial differential equations describing insulin dissociation, protein binding, diffusion and absorption [59]. Later, Trajanoski *et al.* simplified that model, and Tarín *et al.* extended it's use to new insulin formulations, specifically to the insulin analog glargine [78].

Works focused on the development of the artificial pancreas has studied in more detail the fast-acting insulin (lispro) used in insulin pumps. Wilinska *et al.* from the university of Cambridge compared eleven different models for insulin lispro kinetics on data from seven patients with type 1 diabetes [83], concluding that the best performance was achieved by a model which considered two insulin flows, one for the slow absorption and one for the fast absorption insulin. The group of Cambridge is one of the lead research teams for the artificial pancreas development, and they published a diabetic patient simulator in 2010 for *in silico* testing of controllers [82]. However, a simplified version of the insulin absorption model was included, comprising a two compartment single path absorption structure, probably due to identifiability problems. The group of the University of Virginia and the University of Padova also published a simulator (UVA simulator) in 2007 [54] with a very similar two compartment model. The models proposed by these two groups will now be displayed in more detail.

### 1.1.1 UVA model

The model proposed by Magni *et al.* in 2007 [54] has two compartments for the interstitial space and considers the elimination of insulin to happen entirely after the absorption to the plasma compartment. There are two variants of the model, depending on the complexity considered for the model of insulin distribution. The first and simpler model is the one shown in Figure 1.2. In this model, absorption takes place from both stages of the subcutaneous route.

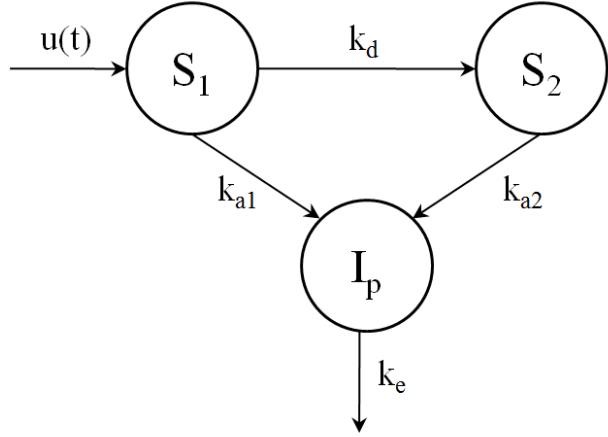


Figure 1.2: UVA model with one compartment for the insulin in plasma

The equations related to the model are:

$$\dot{S}_1(t) = -(k_{a1} + k_d)S_1(t) + u(t) \quad (1.1)$$

$$\dot{S}_2(t) = k_dS_1(t) - k_{a2}S_2(t) \quad (1.2)$$

$$\dot{I}_p(t) = k_{a1}S_1(t) + k_{a2}S_2(t) - k_eI_p(t) \quad (1.3)$$

Usually the subcutaneous part is used with a more complex distribution and elimination model, also proposed by the Virginia and Padova group. The new model is displayed in Figure 1.3. In this case, elimination of insulin takes places both by degradation in the plasma compartment and in the liver. The influence of the liver is displayed by a new compartment.

The corresponding system of equations related is very similar, but substituting equation 1.3 with the following two differential equations:

$$\dot{I}_l(t) = -(m_1 + m_3)I_l(t) + m_2I_p(t) \quad (1.4)$$

$$\dot{I}_p(t) = -(m_2 + m_4)I_p(t) + m_1I_l(t) + k_{a1}S_1(t) + k_{a2}S_2(t) \quad (1.5)$$

No published values of the parameters for T1DM patients have been found in literature. Insulin pharmacokinetics parameters have not been published either.

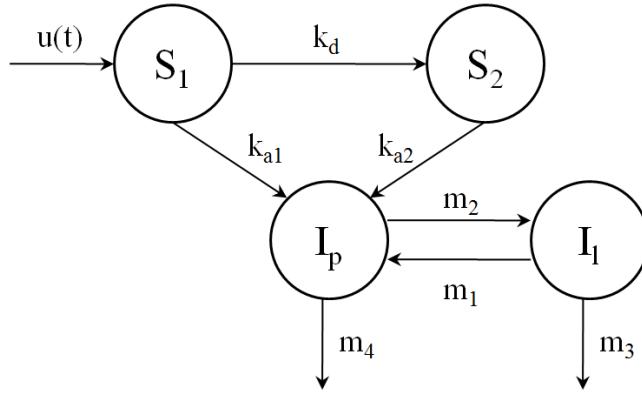


Figure 1.3: Cobelli model considering insulin in the liver and in plasma

Published values [18] for healthy and T2DM patients for the distribution and elimination part of the model are shown in Table 1.2. Parameter  $m_3$  is only constant for T1DM patients while for the published cases it has dynamic behavior depending on the endogenous insulin secretion. Nevertheless, this model is implemented in the University of Virginia Simulator [48], and nominal parameters for healthy and diabetic patients are used in glucose profiles simulated with it.

Parameter	Normal value	T2DM value	Units
$m_1$	0.19	0.379	$\text{min}^{-1}$
$m_2$	0.484	0.673	$\text{min}^{-1}$
$m_4$	0.194	0.269	$\text{min}^{-1}$

Table 1.2: Published values of the UVA s.c. insulin model

### 1.1.2 Cambridge model

In 2005 Willinska compared eleven subcutaneous models of increasing complexity [83]. Those models were evaluated for bolus-basal treatments and continuous infusion with insulin pumps with insulin *lispro*, a human insulin analog. The structure of the model with a better fit to experimental data is shown in Figure 1.4.

It is one of the most complex compartmental models in literature. The equations

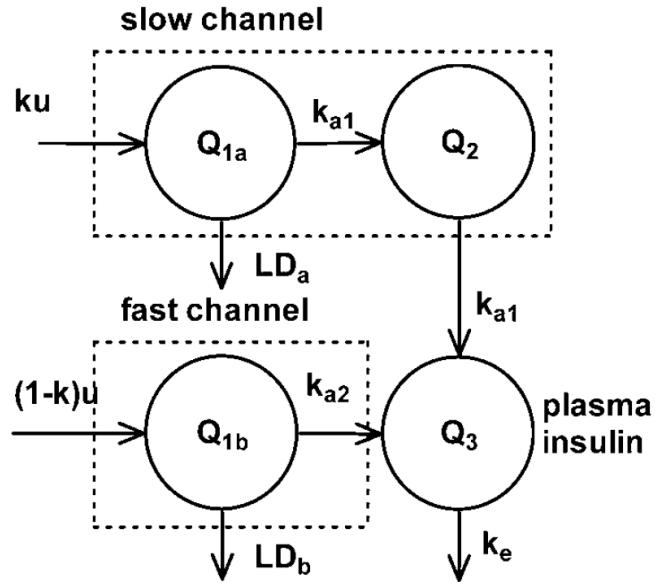


Figure 1.4: Willinska's model compartmental structure

related to the model are:

$$\dot{Q}_{1a}(t) = ku - k_{a1}Q_{1a} - LD_a \quad (1.6)$$

$$\dot{Q}_{1b}(t) = (1 - k)u - K_{a2}Q_{1b} - LD_b \quad (1.7)$$

$$\dot{Q}_2(t) = k_{a1}Q_{1a} - k_{a1}Q_2 \quad (1.8)$$

$$\dot{Q}_3(t) = k_{a1}Q_2 + k_{a2}Q_{1b} - k_eQ_3 \quad (1.9)$$

$$i(t) = \frac{Q_3}{V \cdot BW} \quad (1.10)$$

$$LD_a = \frac{V_{MAX,LD}Q_{1a}}{k_{M,LD} + Q_{1a}} \quad (1.11)$$

$$LD_b = \frac{V_{MAX,LD}Q_{1b}}{k_{M,LD} + Q_{1b}} \quad (1.12)$$

$$(1.13)$$

The most significant characteristic of this model is the existence of two channels of insulin, one of slow absorption and a fast absorption insulin channel, both with degradation of insulin ruled by a Michaelis-Menten kinetic equation. The insulin

concentration is directly calculated from the fourth compartment. Published parameters of the model are shown in Table 1.3.

Parameter	Published value	Units
$k_{a1}$	$1.12 \cdot 10^{-2}$	$\text{min}^{-1}$
$k_{a2}$	$2.1 \cdot 10^{-2}$	$\text{min}^{-1}$
$k_e$	$1.89 \cdot 10^{-2}$	$\text{min}^{-1}$
$k$	0.67	-
$V$	$56.45 \cdot 10^{-2}$	$\text{L kg}^{-1}$
$V_{MAX,LD}$	1.93	$\text{mU min}^{-1}$
$k_{M,LD}$	62.6	$\text{mU}$

Table 1.3: Nominal values of the parameters in Willinska's model

In the simulator published in 2010 [82] the Cambridge group implemented a much simpler linear compartmental model. The model structure is shown in Figure 1.5.

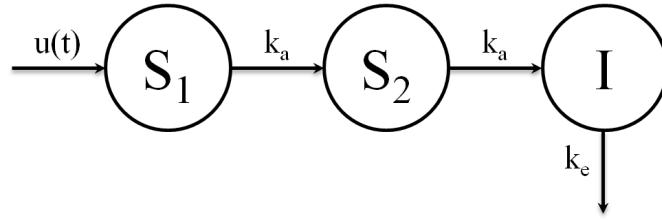


Figure 1.5: Cambridge simulator insulin absorption model structure

This model displays a simple flow of insulin through two compartments into the bloodstream, and it only considers elimination of insulin from blood. The flow constants between compartments is the same. The equations of this model are shown next:

$$\dot{S}_1(t) = u(t) - k_a S_1(t) \quad (1.14)$$

$$\dot{S}_2(t) = k_a S_1(t) - k_a S_2(t) \quad (1.15)$$

$$\dot{I}(t) = \frac{k_a S_2(t)}{V_i} - k_e I(t) \quad (1.16)$$

The insulin compartment  $I(t)$  is directly the insulin concentration. There are 3 parameters to this model:  $k_a$  as the insulin absorption rate,  $k_e$  is the elimination rate from the blood compartment, and  $V_i$  is the insulin distribution. Willinska *et al.* presented

probability distributions of these parameters calculated from a population of 18 type 1 diabetic patients. For sake of simplicity and uniformity only mean values will be shown in this manuscript in table 1.4, but the complete distributions can be found in [82].

Parameter	Published value	Units
$k_a$	0.018	$\text{min}^{-1}$
$k_e$	0.14	$\text{min}^{-1}$
$V_i$	0.12	$\text{L kg}^{-1}$

Table 1.4: Mean values of the parameters in the Cambridge simulator model

## 1.2 Glucose absorption models

Glucose absorption models aim at characterizing the flux of exogenous glucose absorbed by the intestine under different circumstances taking under consideration the information on the meal intake. Modeling meal absorption has been proven difficult due to complex physiology of gastric emptying and intestinal absorption. The nature of the meal, its size and composition, as well as the speed of ingestion, and patient conditions, have influence on the rate of stomach emptying [57] and the final absorption rate of glucose. Large variability was reported by Klingensmith *et al.* [45] even within the same patient and meal. Furthermore, the rate of appearance of glucose in the bloodstream is not directly measurable (unlike blood insulin concentration in the insulin absorption models) which hardens the modeling endeavor.

One of the first models proposed in literature was an exponential model for gastric emptying presented in [39], where the emptying rate is dependent on the meal's volume and calorie density. After the contribution of Hunt *et al.* most of the models published base the dynamics of glucose appearance as a function of the carbohydrate content of the meal. In 1992 Lehman and Deutch [52] proposed a trapezoidal emptying of the stomach based on the carbohydrates ingested. In 2006 Fabietti *et al.* [22] proposed an input-output model based on the different absorption rates of mono- and polysaccharides. A model-less approach was proposed by Herrero *et al.* in [32], where the glucose absorption profile was extracted from a library comprehending several mixed meals and different patients. Also, deconvolution methods can be applied in order to infer the glucose profile as proposed in [33].

In the simulators published up to date, physiological compartmental models are

usually implemented. In the following lines we will describe in detail glucose absorption models from the UVA group and from the Cambridge group.

### 1.2.1 UVA model

Chiara Dalla Man published in 2006 a complete gastrointestinal model along with a critical review of the existing models in literature [17]. This model considers a two-compartment model for digestion and a simple single-compartmental model for the absorption in the gut. To overcome the issue of the measurability of glucose rate of appearance from the gut the UVA group used a triple tracer method for calculating glucose absorption, although in healthy subjects. The model follows the structure shown in Figure 1.6.

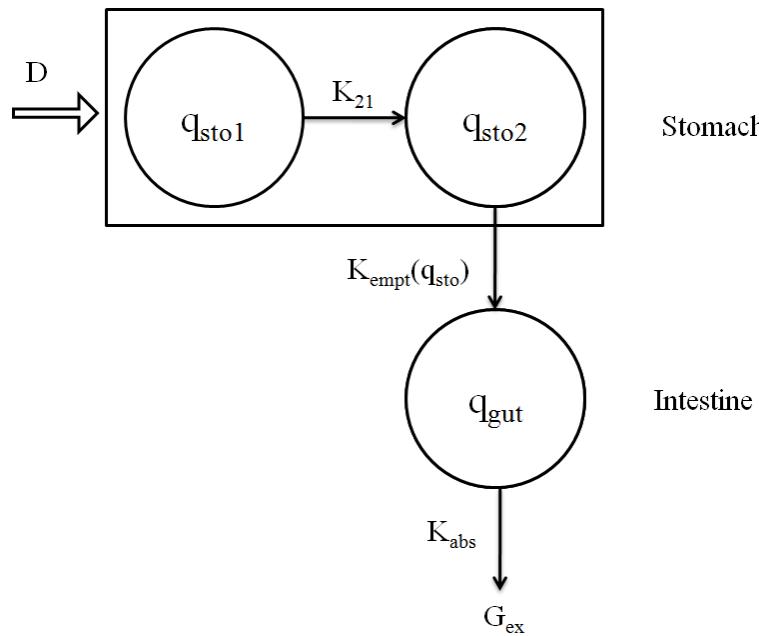


Figure 1.6: A two compartment model represents the stomach and a single compartment the intestine

The two compartments in the stomach part suppose two parts of the digestion of glucose before the gastric emptying. The emptying of the stomach is a non-linear function of the total amount of glucose in the stomach, as will be shown in the model

equations later. The compartmental model equations are:

$$\dot{q}_{sto1}(t) = -K_{21}q_{sto1}(t) + D\delta(t) \quad (1.17)$$

$$\dot{q}_{sto2}(t) = -K_{empt}(q_{sto})q_{sto2}(t) + K_{21}q_{sto1}(t) \quad (1.18)$$

$$\dot{q}_{gut}(t) = -K_{abs}q_{gut}(t) + K_{empt}(q_{sto})q_{sto2}(t) \quad (1.19)$$

$$G_{ex}(t) = fK_{abs}G_{gut}(t) \quad (1.20)$$

where  $\delta(t)$  is the Dirach delta, simulating an impulse input to the model. The rest of parameters added are flux constants, for the transfer of glucose through the system, except for the  $K_{empt}$  parameter, which is time-varying and defines the form of the gastric emptying. The equations describing the transfer rate describing the flow of glucose from the stomach to the intestine are:

$$K_{empt}(q_{sto}) = K_{min} + \frac{K_{max}-K_{min}}{2} \cdot \{\tanh[\alpha(q_{sto} - b \cdot D)] - \tanh[\beta(q_{sto} - c \cdot D)] + 2\} \quad (1.21)$$

$$q_{sto}(t) = q_{sto1}(t) + q_{sto2}(t) \quad (1.22)$$

$$\alpha = \frac{5}{2D(1-b)}; \quad \beta = \frac{5}{2Dc} \quad (1.23)$$

These equations give the gastric emptying a very characteristic shape. In Figure 1.7 the  $K_{empt}$  is plotted against the amount of glucose remaining in the stomach  $q_{sto}$ .

The parameters have been identified both for an Oral Glucose Tolerance Test (OGTT) and a mixed meal, although the mixed meal consisted only of traced jelly with eggs and bacon. Parameter  $K_{21}$  is forced to be equal to  $K_{max}$  for identifiability issues. The published values are shown in Table 1.5.

Parameter	Published value (OGTT)	Published value (meal)	Units
$K_{abs}$	0.205	0.071	$\text{min}^{-1}$
$K_{21}$	0.045	0.054	$\text{min}^{-1}$
$K_{max}$	0.045	0.054	$\text{min}^{-1}$
$K_{min}$	0.013	0.006	$\text{min}^{-1}$
$b$	0.85	0.69	-
$c$	0.25	0.17	-

Table 1.5: Nominal values of the parameters in Dalla Man model

Some work has been done within this research group on identifiability of mixed meals

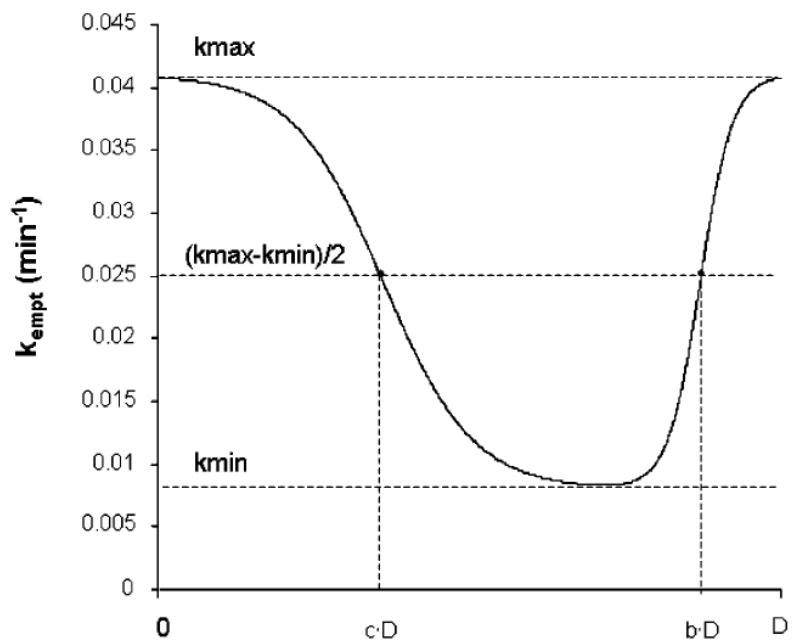


Figure 1.7: Gastric emptying rate versus glucose remaining in the stomach [17]

with the Dalla Man model [1], showing a dependence of the identification results on the size of the meal considered.

### 1.2.2 Cambridge model

Hovorka *et al.* proposed a simple model for glucose absorption in his paper of 2004 [35], where the gastrointestinal system is modeled by two identical compartments with the same transfer rate. Later, the model was refined in a new paper [82] considering the transfer rate  $t_{max}$  as a time-varying parameter. The equations of the model are:

$$\dot{G}_1(t) = -\frac{G_1(t)}{t_{max}} + Bio \cdot D(t) \quad (1.24)$$

$$\dot{G}_2(t) = \frac{G_1(t)}{t_{max}} - \frac{G_2(t)}{t_{max}} \quad (1.25)$$

$$G_{ex} = \frac{G_2}{t_{max}} \quad (1.26)$$

where:

- $D(t)$  is the amount of carbohydrates ingested in grams. The meal in this model is considered as a pulse input.
- $Bio$  is the effectiveness of the absorption of the carbohydrates ingested i.e. is the portion of the carbohydrates that have been eaten that will go into the circulatory system.
- $t_{max}$  is the maximum absorption time of the carbohydrates. This parameter regulates the transfer speed between the compartments. It is a bounded parameter following:

$$t_{max} = \begin{cases} t_{max\ ceil} & \text{if } G_{ex} > G_{ex\ ceil} \\ t_{max} & \text{otherwise} \end{cases} \quad (1.27)$$

where  $t_{max\ ceil} = \frac{G_2}{G_{ex\ ceil}}$  and  $G_{ex\ ceil}$  is the maximum glucose flux from the gut.

The nominal values of the parameters are shown in Table 1.6.

This model has the weakness of not considering the different compositions of mixed meals, as other models do. That is not the aim of this model though, but to rapidly provide

Parameter	Published value	Units
$t_{max}$	40	min
$Bio$	0.8	-
$G_{ex\ ceil}$	[0.02, 0.035]	mmol kg <sup>-1</sup> min <sup>-1</sup>

Table 1.6: Nominal values of the parameters in Hovorka model

a coherent input for the glucoregulatory system in the simulation environment.

### 1.3 Endogenous models

The endogenous model is the part of the glucose-insulin model that describes the different regulatory metabolic pathways of blood glucose concentration. Given that this thesis is focusing in the identification of T1DM people, the author decided that the dynamics and equations describing the secretion of insulin were no relevant to be shown for those models that describe it.

Usually the endogenous model has two sub-models, one for the insulin distribution and elimination system (if not considered in the insulin absorption model), and the other for the glucose metabolism, transformation and elimination. This second sub-model has to consider the influence of the liver and the kidneys on the blood glucose production and elimination, as well as the peripheral intake by muscles and adipose tissue, and any other influences that may affect glucose concentration.

Two different groups of models will be described next. The first two models (Bergman and Panunzi) are considered as *minimal* models, and are very simple models that do not attempt at simulating the whole metabolic system of glucose regulation. Minimal models are widely used for research either in simulation studies or for control, but lack the accuracy of more complex models. The other two models reviewed next are physiology based models developed by the UVA group and the Cambridge group, and are much more complex than the previous ones.

### 1.3.1 Bergman model

The first model to be reviewed is going to be the (probably) most used and best known among all the endogenous models used in diabetes. It was proposed by Bergman *et al.* in 1981 [4]. It is considered a minimal model because it only describes the influence of insulin on blood glucose concentration, and it does not consider many other phenomena, or it considers them in a very simplified way. The justification for this simplicity is that the objective of this model was, initially, to simulate the response of the IntraVenous Glucose Tolerance Test (IVGTT), which has a very simple behavior.

Bergman minimal model considers that insulin actions on glucose are delayed, and that delay is represented by a new compartment of *remote* insulin. The scheme of the model is shown in Figure 1.8.

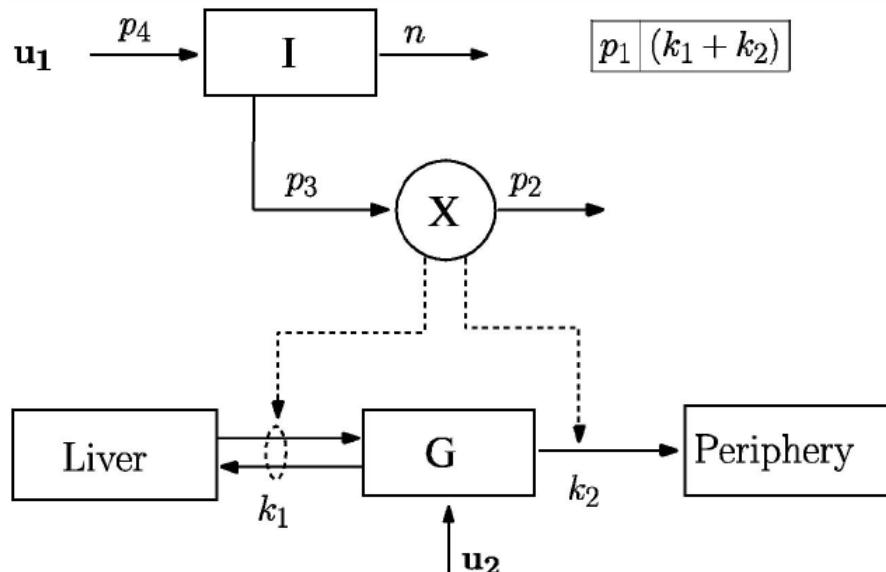


Figure 1.8: Bergman minimal model of insulin and glucose dynamics, adapted from Bergman and colleagues [4]

The equations that describe the model are:

$$\dot{I}(t) = -nI(t) + p_4u_1(t) \quad I(0) = I_b = \frac{p_4}{n}u_{1b} \quad (1.28)$$

$$\dot{X}(t) = -p_2X(t) + p_3[I(t) - I_b] \quad X(0) = 0 \quad (1.29)$$

$$\dot{G}(t) = -p_1G(t) - X(t)G(t) + p_1G_b + \frac{u_2(t)}{Vol_G} \quad G(0) = G_b \quad (1.30)$$

Depending on the model of insulin absorption used, equation 1.28 can be substituted by the corresponding equation of the model chosen.  $X(t)$  is the remote insulin compartment,  $G(t)$  is the blood glucose concentration,  $u_1(t)$  is the insulin flow coming from the insulin system,  $u_2(t)$  is the exogenous flow of glucose coming either intravenously or from the gastrointestinal system. Published values for the model parameters are shown in Table 1.7.

Parameter	Published value	Units
$p_1$	0.035	$\text{min}^{-1}$
$p_2$	0.05	$\text{min}^{-1}$
$p_3$	0.000028	$\text{ml}/\mu U \cdot \text{min}^2$
$p_4$	0.098	$\text{ml}^{-1}$
$n$	0.142	$\text{min}^{-1}$
$Vol_G$	117	dl

Table 1.7: Nominal values of the parameters in Bergman model [75]

Bergman model has been used for more than 20 years in diabetes research due to its identifiability and controllability properties, but it is far of being the perfect model. Nevertheless, Bergman model will be analyzed in detail in the next chapters of this thesis.

### 1.3.2 Panunzi model

Simona Panunzi and her group published a study in 2007 comparing some of the characteristics of Bergman model and new features of a proposed model for the IVGTT scenario, with a delayed insulin secretion rate [63]. The new model proposed surpassed the rest in simulated experiments and in identifiability properties, but it was only tested

in healthy patients. The equations of the model are:

$$\dot{G}(t) = -K_{xgl}I(t)G(t) + \frac{T_{gh}}{V_g} \quad (1.31)$$

$$\dot{I}(t) = -K_{xi}I(t) + \frac{T_{ig\ max}}{V_i} \frac{\left(\frac{G(t-\tau_g)}{G^*}\right)^\gamma}{1 + \left(\frac{G(t-\tau_g)}{G^*}\right)^\gamma} \quad (1.32)$$

This model includes delayed differential equations for the insulin production sub-model, but when simulating type 1 diabetic patients, whom do not have endogenous insulin secretion, the model becomes much more simple. The parameters involved in the previous model are:

- $G_b$  is the basal glucose concentration.
- $I_b$  is the basal plasma insulin.
- $K_{xgl}$  is the insulin sensitivity. It represents the insulin-dependent glucose uptake by tissues per unit of insulin concentration.
- $T_{gh}$  represents the balance between the hepatic outtake of glucose and the insulin-independent glucose intake, including the one of the liver.
- $V_g$  is the apparent glucose distribution volume.
- $K_{xi}$  is the disappearance rate of insulin.
- $G^*$  is the glycemia at which the insulin secretion rate is half of its maximum.
- $T_{ig\ max}$  is the maximum rate of insulin release.
- $V_i$  is the apparent insulin distribution volume.
- $\tau_g$  represents the apparent delay with which the pancreas changes insulin release in response to a variation in blood glucose.
- $\gamma$  is the progressivity with which the pancreas reacts to circulating glucose concentrations.

The values published for these parameters, only for healthy patients, are shown in Table 1.8.

Parameter	Published value	Units
$V_g$	0.152	$\text{L kg}^{-1}$
$\tau_g$	19.271	min
$K_{xgl}$	$1.43 \times 10^{-4}$	$\text{min}^{-1} \text{ pM}^{-1}$
$K_{xi}$	0.101	$\text{min}^{-1}$
$\gamma$	2.464	-

Table 1.8: Nominal values of the parameters in Panunzi model

Panunzi model is also used later in this thesis because of its simplicity, but some variations were made in chapter 1.4 due to the focus of this model on healthy patients and the IVGTT scenario, and in order to extend its use to diabetic patients.

### 1.3.3 UVA model

The core of the UVA model was presented in [18], and it is one of the most important models in diabetes research. It is a very complex model almost exclusively based in physiological knowledge of the glucose metabolism. It is usually combined with the Dalla Man gastrointestinal model and the insulin pharmacokinetics model of the same group (in fact all the models were developed together) in a large mathematical model of the glucose metabolism. Magni et al. [54] used the complete UVA model to control an *in silico* diabetic patient. It is also the model implemented in the UVA simulator [48] which was accepted by the FDA as substitute of animal trials in the context of a protocol application of University of Virginia and Padova for a controller validation clinical trial. The core structure is pretty simple, as shown in Figure 1.9:

The model equations are:

$$\dot{G}_p(t) = EGP(t) + Ra(t) - U_{ii}(t) - E(t) - k_1G_p(t) + k_2G_t(t) \quad (1.33)$$

$$\dot{G}_t(t) = -U_{id}(t) + k_1G_p(t) - k_2G_t(t) \quad (1.34)$$

$$G(t) = G_p(t)/V_G \quad (1.35)$$

In the previous equations there are many inputs and outputs to the glucose compartments that need description. It must be noted that so far there are no insulin related equations; insulin influences the flow of glucose coming in or out of the different compartments. The meaning of these variables is explained next:

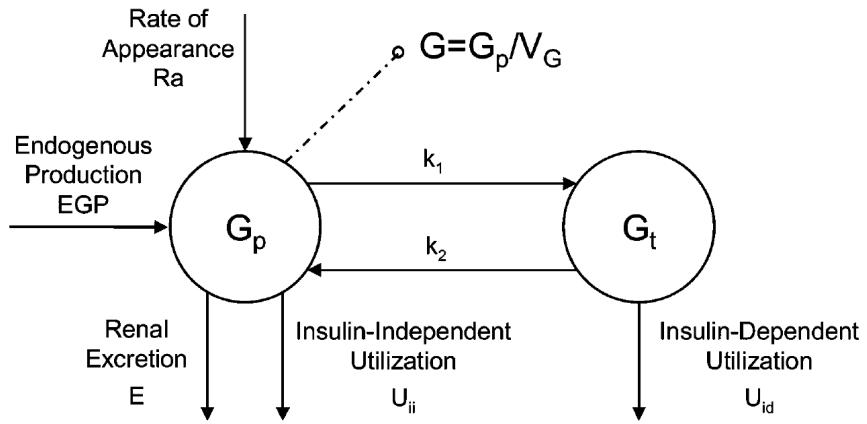


Figure 1.9: Cobelli model core is composed of two compartments of glucose, one for blood and one for the tissues interstitial fluid

- $Ra$  is the exogenous flux of glucose coming from the gut.
- $U_{ii}$  is the utilization of glucose that is non dependent on insulin. It is usually considered constant and equal to  $F_{cns}$ .
- $U_{id}$  is the utilization that depends on the insulin concentration, and it follows the following set of equations:

$$\dot{X}(t) = -p_{2U}X(t) + p_{2U}[I(t) - I_b] \quad (1.36)$$

$$V_m(t) = V_{m0} + V_{mx}X(t) \quad (1.37)$$

$$U_{id}(t) = \frac{V_m(t)G_t(t)}{K_m + G_t(t)} \quad (1.38)$$

where  $X(t)$  is the remote insulin,  $I(t)$  is the plasma insulin,  $I_b$  is the basal insulin and  $V_m(t)$  is the transfer rate for the Michaelis-Menten equation shown in equation 1.38.

- $E(t)$  represents the renal excretion, which occurs if plasma glucose exceeds a certain threshold. It is modeled as follows:

$$E(t) = \begin{cases} k_{e1}[G_p(t) - k_{e2}] & \text{if } G_p(t) > k_{e2} \\ 0 & \text{otherwise} \end{cases} \quad (1.39)$$

where  $k_{e1}$  is the glomerular filtration rate and  $k_{e2}$  is the renal threshold of glucose.

- $EGP(t)$  is the Endogenous Glucose Production, and it depends on a delayed insulin signal as follows:

$$\dot{I}_1(t) = -k_i[I_1(t) - I(t)] \quad (1.40)$$

$$\dot{I}_d(t) = -k_i[I_d(t) - I_1(t)] \quad (1.41)$$

$$EGP(t) = \max\{0, k_{p1} - k_{p2}G_p(t) - k_{p3}I_d(t)\} \quad (1.42)$$

where  $I(t)$  is the insulin concentration in plasma.

The published parameters for healthy and type 2 diabetic patients (no type 1 diabetic parameters have been published) are those in Table 1.9.

Parameter	Healthy	Type 2 diabetes	Units
$V_G$	1.88	1.49	$\text{dL kg}^{-1}$
$k_1$	0.065	0.042	$\text{min}^{-1}$
$k_2$	0.079	0.071	$\text{min}^{-1}$
$k_{p1}$	2.70	3.09	$\text{mg kg}^{-1} \text{min}^{-1}$
$k_{p2}$	0.0021	0.0007	$\text{min}^{-1}$
$k_{p3}$	0.009	0.005	$\text{mg kg}^{-1} \text{min}^{-1} \text{per pmol L}^{-1}$
$k_i$	0.0079	0.0066	$\text{min}^{-1}$
$F_{cns}$	1	1	$\text{mg kg}^{-1} \text{min}^{-1}$
$V_{m0}$	2.5	4.65	$\text{mg kg}^{-1} \text{min}^{-1}$
$V_{mx}$	0.047	0.034	$\text{mg kg}^{-1} \text{min}^{-1} \text{per pmol L}^{-1}$
$K_{m0}$	225.59	466.21	$\text{mg kg}^{-1}$
$p_{2U}$	0.0331	0.0840	$\text{min}^{-1}$
$k_{e1}$	0.0005	0.0007	$\text{min}^{-1}$
$k_{e2}$	339	269	$\text{mg kg}^{-1}$

Table 1.9: Nominal values of the parameters in Cobelli model

This model has been used in several experimental settings for closed-loop controller testing. Kovatchev *et al.* published a study on 20 T1DM patients using an MPC controller tuned with the UVA simulator [47] [12]. Improvement was reported in hypoglycemic events occurrence and in amount of time spent in euglycemic range. Currently, the UVA group is moving towards performing home control experiments with using the controllers designed with this model.

### 1.3.4 Cambridge model

The Cambridge group proposed an endogenous model with split focus on each insulin action on different phenomena and its final effect on blood glucose [37]. The model adds a new compartment for every action of insulin, and there are three considered events:

- Insulin increases the flow of glucose from blood to the tissues.
- Insulin increases the glucose uptake by muscles and adipose tissue.
- Insulin inhibits production of glucose in the liver.

These three influences are reflected in the model as virtual compartments. The relation between actual insulin in plasma, every virtual compartment representing insulin actions and the two compartments for glucose is shown in Figure 1.10.  $x_1$ ,  $x_2$  and  $x_3$  represent the insulin actions,  $Q_1$  is the glucose mass in the accessible compartment, and  $Q_2$  is the glucose present in the non-accessible compartment.

The equations that represent this model are:

$$\dot{x}_1(t) = -k_{a1}x_1(t) + k_{b1}I(t) \quad x_1(0) = 0 \quad (1.43)$$

$$\dot{x}_2(t) = -k_{a2}x_2(t) + k_{b2}I(t) \quad x_2(0) = 0 \quad (1.44)$$

$$\dot{x}_3(t) = -k_{a3}x_3(t) + k_{b3}I(t) \quad x_3(0) = 0 \quad (1.45)$$

$$\dot{Q}_1(t) = - \left[ \frac{F_{01}^c}{V_G G(t)} + x_1(t) \right] Q_1(t) + k_{12}Q_2(t) - F_R + EGP + G_{ex}(t) \quad Q_1(0) = Q_{1,0} \quad (1.46)$$

$$\dot{Q}_2(t) = x_1(t)Q_1(t) - [k_{12} + x_2(t)]Q_2(t) \quad Q_2(0) = Q_{2,0} \quad (1.47)$$

$$G(t) = Q_1(t)/V_G \quad (1.48)$$

Equation 1.46 has several terms to be defined:

- $EGP$  stands for Endogenous Glucose Production, which is the flux of glucose coming from the liver. It is defined as:

$$EGP = \begin{cases} EGP_0[1 + x_3(t)] & \text{if } EGP \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (1.49)$$

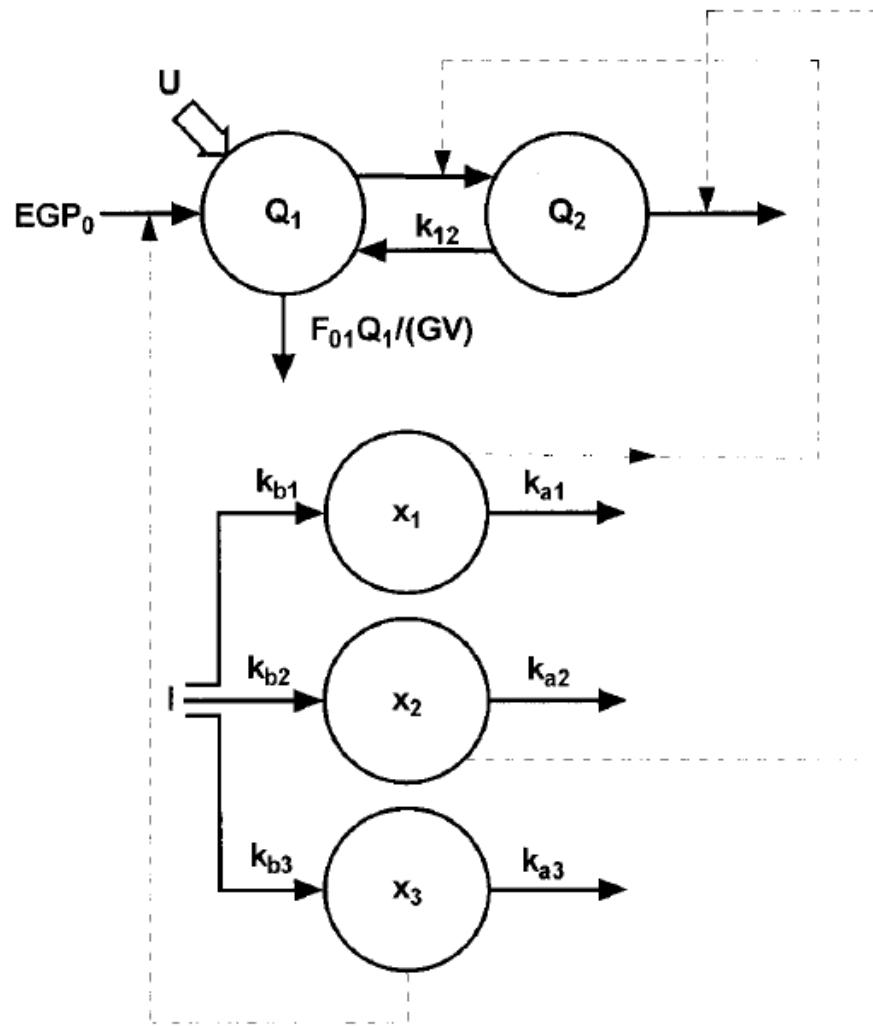


Figure 1.10: Hovorka endogenous model structure arranged in compartments

- $F_{01}^c$  is the insulin-independent glucose flux, and it is defined as:

$$F_{01}^c = \frac{F_{01}^s G}{G + 1.0} \text{ where } F_{01}^s = \frac{F_{01}}{0.85} \quad (1.50)$$

- $F_R$  is the renal glucose clearance above the glucose threshold of  $R\_thr$ , and it is defined as:

$$F_R = \begin{cases} R\_cl(G - R\_thr)V_G & \text{if } G \geq R\_thr \\ 0 & \text{otherwise} \end{cases} \quad (1.51)$$

Where  $R\_cl$  is the renal clearance.

The model has many parameters to be tuned, specially in the part of insulin actions, where there are two parameters for each action corresponding to the input and output flows of the compartment. Usually these parameters are reformulated into the so called *insulin sensitivities*, due to their physiological meaning since they correspond to the glucose decrement per unit of insulin given. The reformulation is then:

- $S_{IT} = \frac{k_{b1}}{k_{a1}}$  where  $S_{IT}$  is the insulin sensitivity to the transport of glucose.
- $S_{ID} = \frac{k_{b2}}{k_{a2}}$  where  $S_{ID}$  is the insulin sensitivity to the distribution of glucose.
- $S_{IE} = \frac{k_{b3}}{k_{a3}}$  where  $S_{IE}$  is the insulin sensitivity to the endogenous glucose production.

After this transformation, equations (1.43), (1.44) and (1.45) result in:

$$\dot{x}_1(t) = -k_{a1}x_1(t) + S_{IT}k_{a1}I(t) \quad x_1(0) = 0 \quad (1.52)$$

$$\dot{x}_2(t) = -k_{a2}x_2(t) + S_{ID}k_{a2}I(t) \quad x_2(0) = 0 \quad (1.53)$$

$$\dot{x}_3(t) = -k_{a3}x_3(t) + S_{IE}k_{a3}I(t) \quad x_3(0) = 0 \quad (1.54)$$

The published values of all the parameters are those in Table 1.10. The parameters shown in here are mean values of the several sets of parameters published.

The Cambridge model has been used both for simulation and control purposes, in many different scenarios, from critical patients [36] to overnight experiments [35] with successful results, and recently it has been implemented in a complete mathematical patients simulator [82]. This simulator was also used for patient prediction and controller tuning in a recent closed-loop experiment performed both in adolescents [34] and in

Parameter	Published value	Units
$k_{12}$	0.066	$\text{min}^{-1}$
$V_G$	0.16	$\text{L kg}^{-1}$
$EGP_0$	0.0161	$\text{mmol kg}^{-1} \text{ min}^{-1}$
$F_{01}$	0.0097	$\text{mmol kg}^{-1} \text{ min}^{-1}$
$k_e$	0.138	$\text{min}^{-1}$
$V_i$	0.12	$\text{L kg}^{-1}$
$k_{a1}$	0.006	$\text{min}^{-1}$
$k_{a2}$	0.06	$\text{min}^{-1}$
$k_{a3}$	0.03	$\text{min}^{-1}$
$S_{IT}$	$51.2 \times 10^{-4}$	$\text{mU L}^{-1} \text{ min}^{-1}$
$S_{ID}$	$8.2 \times 10^{-4}$	$\text{mU L}^{-1} \text{ min}^{-1}$
$S_{IE}$	$520 \times 10^{-4}$	$\text{mU L}^{-1} \text{ min}^{-1}$

Table 1.10: Nominal values of the parameters in Hovorka model

adults [38] in an overnight controlled environment. Currently, the Cambridge group is planning and performing domiciliary studies under similar premises.

## 1.4 Critical selection of models

The identification problem in diabetes has been proven really difficult even for the simplest models. The straight solution to this problem is to develop new, more efficient models for the glucoregulatory system, but that is not the main objective of this thesis. Instead, a critical review of literature models is going to be done in the following lines.

When trying to choose an endogenous model the quest is really complex. UVA's model is a clear outstanding model in literature and is likely to be chosen for simulation and control purposes because of it's FDA acceptance. However, Willinska *et al.* published a review of the identifiability of models for simulation [84], stating that UVA's model has the problem of having too many parameters to be identified from clinical data, which ensures identifiability problems with the model structure.

Minimal models (not used for simulation) such Bergman's or Panunzi's model skip the problem of overpopulation of parameters, but obviously describe with less accuracy glucose behavior. Bergman's model has been strongly criticized in late years because of

its simplicity and the fact that it does not fit correct glucose behavior against insulin. Quon *et al* [68] proved in 1994 with a series of experiments involving the Biostator device that Bergman model underestimates insulin action on glucose removal from blood, and it overrates the effect of glucose concentration in its own disappearance (*glucose effectiveness*).

In 1999, Regitnig *et al.* [70] confirmed the overestimation of *glucose effectiveness*. They also proved that all minimal models are inevitably wrong if they consist of one single compartment for glucose, without considering the interstitial dynamics or other situations. Considering that a minimal model has to stay simple, let us take a closer look to the equations related to the glucose compartment in minimal models, like the Bergman model. Considering that there is no exogenous input of glucose, equation (1.30) stands:

$$\dot{G}(t) = -p_1 G(t) - X(t)G(t) + p_1 G_b \quad (1.55)$$

The part of the equation  $-p_1 G(t)$  is the term of the Bergman model that represents the glucose effectiveness, and it depends directly on the parameter  $p_1$ , which is patient-dependent. The term  $p_1 G_b$  sets the equilibrium point of the model to the basal glucose if insulin is at basal level and there is no glucose input. The term  $X(t)G(t)$  is the insulin related term, applying insulin action through a delay compartment. Looking at the equation of the glucose compartment of another minimal model, the Panunzi model:

$$\dot{G}(t) = -K_{xgl} I(t)G(t) + \frac{T_{gh}}{V_g} \quad (1.56)$$

In this case, the term related to glucose effectiveness does not exist. The term related to insulin input,  $-K_{xgl} I(t)G(t)$ , is applied without a delay because Panunzi's model was designed for healthy patients and the delay is considered in the endogenous secretion of insulin. The term  $\frac{T_{gh}}{V_g}$  is similar to the equilibrium point term in Bergman's equation 1.55, both are constant terms that define the equilibrium point of the equation, but they are considered in different ways. In the Bergman model, the basal level is stated explicitly, while in the Panunzi model it is seen as a hepatic balance of glucose. Both approaches are true and useful, and the identification is done in the same way in both examples.

Bergman model exhibits an odd behavior when simulating a diabetic patient to whom basal insulin infusion is removed. The correct physiologic behavior to a empty insulin compartment in a diabetic patient is for blood glucose to increase asymptotically. In Figure 1.11 a postprandial simulation with Bergman's model is shown, including a long pre-prandial stage of 10 hours in which the insulin infusion is stopped; then, at minute

600, insulin infusion is restored to its nominal value and a 75 grams of a mixed meal ingested (simulated by the UVA model) simultaneously to a 7.5 units of insulin bolus injected.

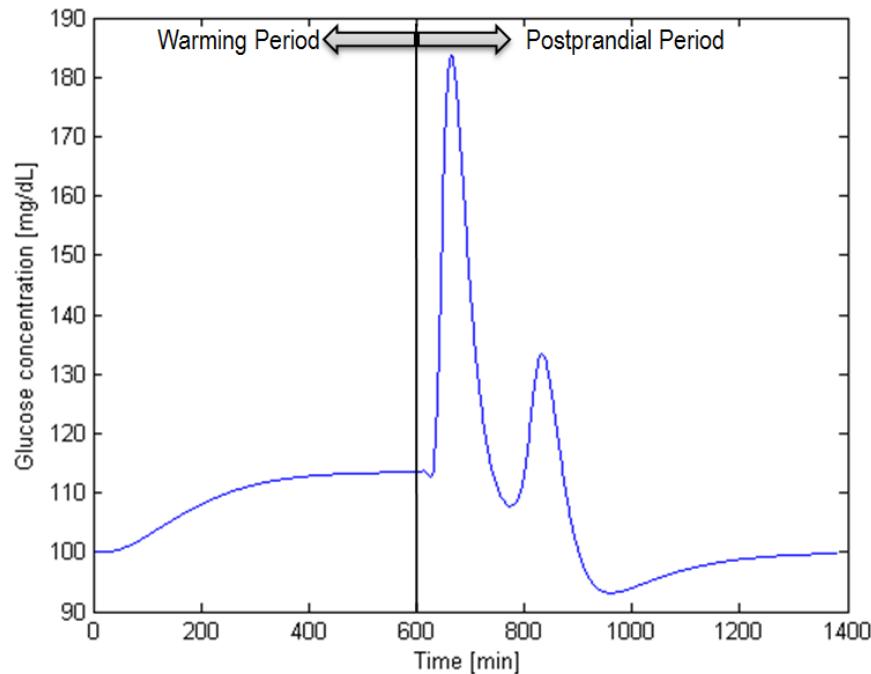


Figure 1.11: Bergman model simulation of a stop in the basal insulin infusion with nominal values of glucose effectiveness. Basal infusion is restored when the postprandial period begins

A small increase in glucose level can be seen as a consequence of the elimination of insulin infusion. Basal insulin removal causes a change in the set-point of the output variable, *i.e.* glucose concentration. If basal insulin infusion stops in a model with smaller glucose effectiveness, the results are shown in Figure 1.12.

The behavior observed in the warming period of Figures 1.11 and 1.12 is unrealistic. A diabetic patient whose insulin infusion is removed should become an unstable process, not just change its equilibrium point. Panunzi's model, in its equation (1.56) becomes a pure integrator if insulin is removed, making the system unstably increasing. Panunzi's model seems more suited to the physiology of glucose in diabetic patients in this case, but it also has its disadvantages. As it was said before, the Panunzi model represents the delay

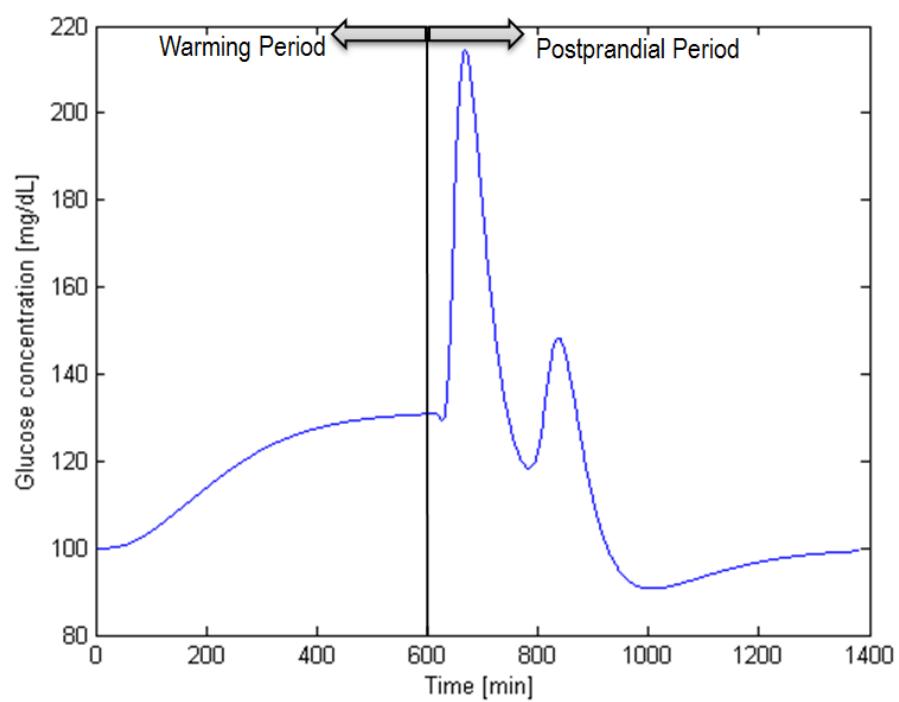


Figure 1.12: Bergman model simulation of a stop in the basal insulin infusion with 50% reduction in glucose effectiveness

of insulin action in the equation of secretion, but that equation (1.32) is eliminated when simulating diabetic patients because of the absolute insulin deficiency characteristic of the diabetic persons, so the delay effect is removed as well. If the delay is not reinstated somehow, the model simulates much faster dynamics when applying the bolus insulin in front of the simultaneous meal ingestion, making blood glucose to drop immediately after the injection, as can be seen in Figure 1.13.

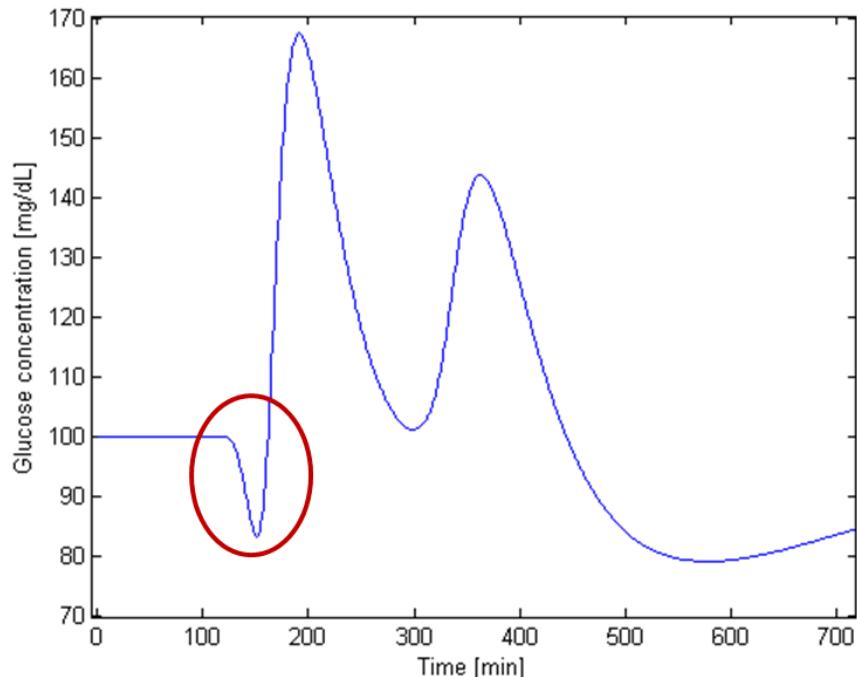


Figure 1.13: Panunzi model simulation of a postprandial period. Meal ingestion occurs at minute 120. Blood glucose drop after insulin bolus administration and before meal glucose appears into bloodstream is circled

This phenomenon is too pronounced and unrealistic. Even though insulin dynamics can, and often are, faster than gastrointestinal absorption, the effect on the displayed simulation is too abrupt and can yield to danger to the simulated patient. However, this effect can be easily avoided by adding a simple dynamic delay equation (1.58) to Panunzi's

model, resulting in the next system of equations:

$$\dot{G}(t) = -K_{xgl}X(t)G(t) + \frac{T_{gh}}{V_g} \quad (1.57)$$

$$\dot{X}(t) = -k_i[X(t) - I(t)] \quad (1.58)$$

With the addition of a new parameter to the model. This new parameter is not tunned and requires a nominal value to be set. This can be done simply by reviewing related literature. In [30] Helms & Kelley quantify the average insulin action delay with a 20 minutes settling time. By looking at the new delay equation as a first order model, the time constant of the system can be calculated by measuring the time at 63% of the settling time, which reads a time constant of 12.6 minutes, which makes the parameter  $k_i$  to be approximately equal to  $0.08 \text{ min}^{-1}$ . The response to a meal and a bolus in that case is shown in Figure 1.14, where the dropping of glucose just after the bolus time is almost non present.

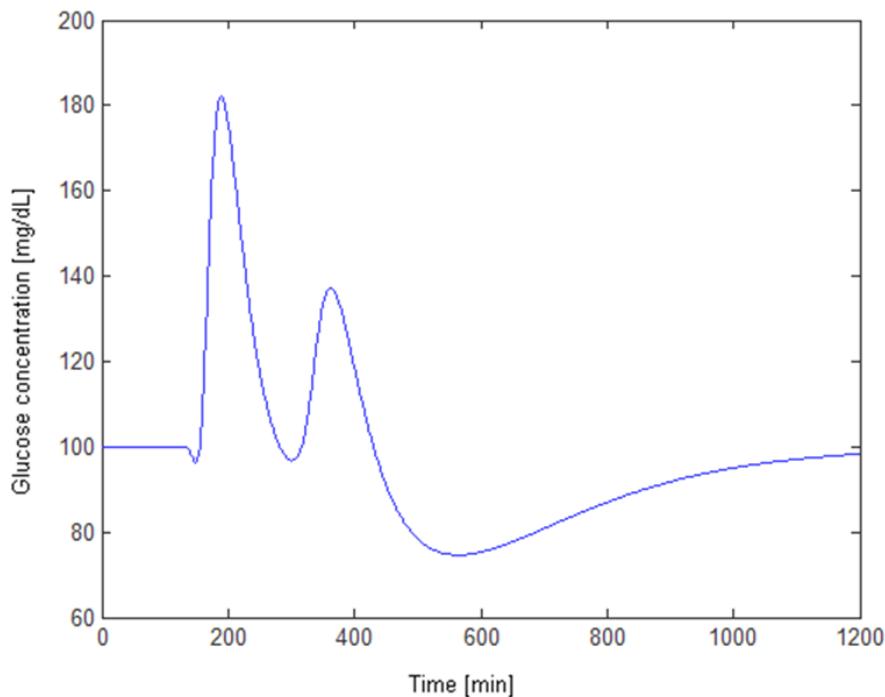


Figure 1.14: Modified Panunzi model simulation of a postprandial period. Meal ingestion occurs at minute 120.

The shown model is completely satisfactory for identification purposes and control.

In order to check the model's reliability we look upon the work of 1994 by Torlone et al. [79], where they showed a series of experiments in which insulin bolus is injected few minutes prior to an intravenous glucose infusion (infused in order to avoid hypoglycemia). This experiment shows clearly the effect and dynamics of insulin, and how it makes glucose to decrease in a similar way to an impulse response. In Figure 1.15 insulin action is applied separately from any other action for almost one hour before the glucose infusion starts.

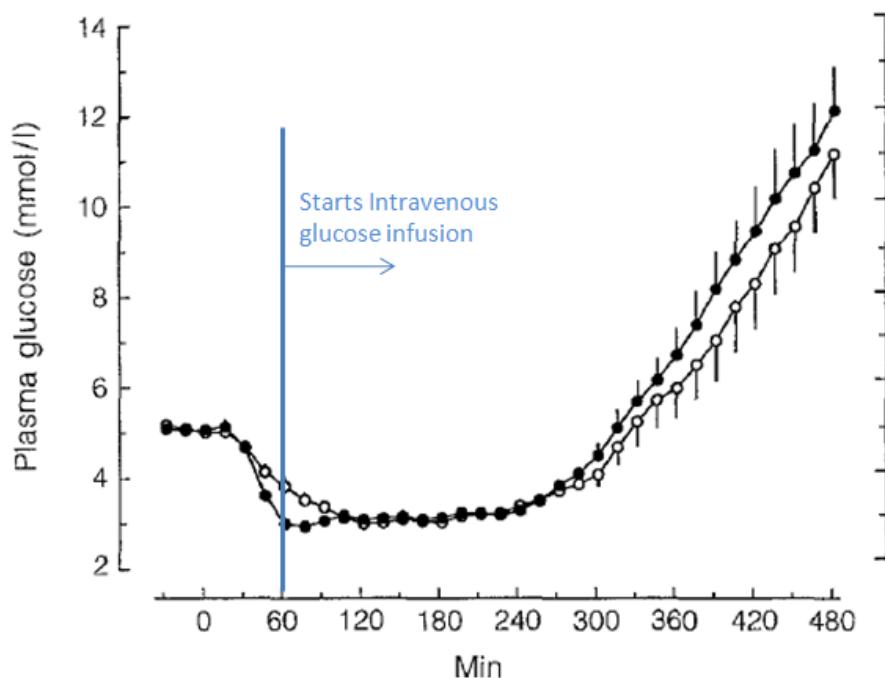


Figure 1.15: Insulin impulse response of glucose after bolus and counteraction with intravenous glucose infusion [79]. Black dots represent the blood glucose measurements for monomeric insulin analog, and the white dots for human insulin

From the work done by Torlone *et al.* it can be extracted that there is no significant change in blood glucose until approximately 15 minutes after bolus injection, and then glucose drops steadily for 30 minutes until approximately  $70\text{ mg/dl}$  (approximately  $3\text{ mmol/l}$ ). A very similar set-up was used for simulating the modification of Panunzis' model including insulin delay; an insulin bolus was applied and the effect on the blood glucose measured, as displayed in Figure 1.16, but no intravenous glucose infusion was simulated.

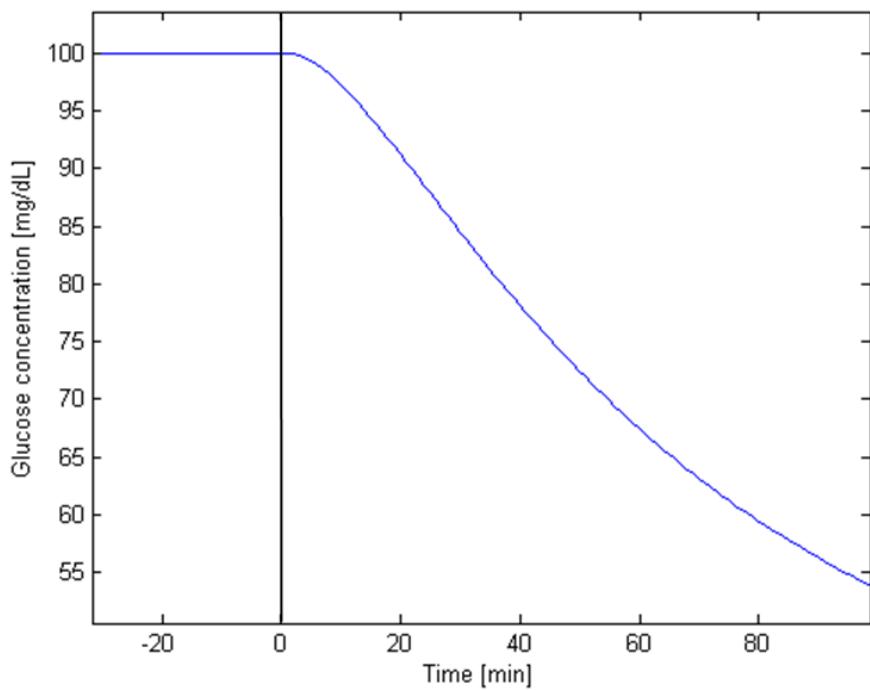


Figure 1.16: Variation of blood glucose ( $mg/dl$ ) in Panunzi's model simulating a drop of glucose after insulin injection. Insulin is given at time 0

In simulation, the response is exactly as expected: no significant change in blood glucose in the first minutes, and in the next half hour, glucose level drops to almost 70 mg/dl. This new dynamics make the new model much more reliable than Bergman or Panunzi's models in the referent to insulin simulation.

Another simple feature has also been included in the modified Panunzi model. Endogenous hepatic glucose production is one of the parameters of the model,  $T_{gh}$ , and it is considered constant. This parameter is actually variable, because endogenous hepatic production is suppressed by insulin. A simple variation of the model has been introduced, considering the  $T_{gh}$  parameter to be reduced when plasma insulin surpassed a defined threshold, and defined as a piecewise function as follows:

$$T_{gh} = \begin{cases} T_{gh0} & \text{if } I(t) < T_{ghthr} \text{ mIU/l} \\ K_{gh} \cdot T_{gh0} & \text{otherwise} \end{cases} \quad (1.59)$$

With the described changes applied to the Panunzi model, we believe a very useful yet simple model was developed and can be used for identification purposes in the following thesis. This resulting model will be mentioned in this thesis as "Modified Panunzi Model".

# Chapter 2

## Identification in Diabetes

Insulin treatment is currently a necessity for T1DM patients. The insulin dose must be adjusted to every patient specifically and for that physicians have traditionally characterized the patients using clinic parameters tuned heuristically. Such parameters (like insulin sensitivity or insulin-to-carbohydrate ratio) measure the influence of insulin on glucose level directly and, in many cases, constantly throughout the patients life.

In the artificial pancreas context the characterization of a patient is required to be much more accurate. Model individualization is required for every patient in order to get accurate enough predictions of the blood glucose levels. Using mathematical models for glucose prediction has been present in diabetes literature since the first models where published, but few satisfying results have been achieved for long term predictions.

Patient identification deals with inter-patient (between patients) variability. From this point onward when referring to variability only intra-patient (within the patient) variability will be accounted.

In this chapter I will review the most important and latest findings published in literature related to model individualization for diabetes.

## 2.1 Patient Identification

Glucose prediction from identification of models using experimental glucose data can be classified according to the nature of the model used. As described in the models section 1 models for diabetes can be data-driven or physiology-driven.

From the engineering point of view, Ståhl and Johansson published a pilot identification study [77] using modified models from literature and data-driven models. This paper describes the experience of the first author, whom was prescribed diabetes shortly before the study, analyzing and modeling his own glucose profile over a period of several weeks. They analyzed in detail the characteristics of the data-set, including auto-correlation function, probabilistic distribution of the samples, sampling times, and statistical properties. They used models in literature for the input models, using in both cases (insulin and meal) two different pathways for fast and slow inputs. As for the endogenous models, they tested a battery of grey-box, data-driven models, including ARX, ARMAX, General Transfer Function Models, and state space models of increasing complexity. They used the models for validation of the identified patient under different prediction horizons. As seen in Figure 2.1, the study concluded that GTFM were well suited for prediction up to 2 hours ahead (8 simulation steps), but further horizons of prediction are still not possible using any of the models used.

The study performed by Ståhl and Johansson was very useful in the diabetes identification scene as a pilot study, but it was nonetheless limited in many aspects. The data-set used was comprised only of data from one patient, thus neglecting completely the population variation in the diabetes metabolism. The study also only used self-measured capillary samples with irregular sampling times. Capillary blood glucose estimation can be subject to many errors and disturbances, and it is also too sparse to be able to capture glucose dynamics. More frequent sampling and more reliable measurements can improve the outcome of the study.

Johansson and his colleagues in Lund continued the investigation further on, publishing another study [11] where the effect of subspaced based models was considered. In this case, the data was taken from a patient that was wearing a CGM, monitored during three days under CSII therapy, which provided with much more frequent data than in the previous study. Again, the models used for the inputs of the system were rather complex physiology-based models, like the UVA group meal model and the Cambridge insulin model described in chapters 1.2.1 and 1.1.2 respectively. The rational behind this decision lies in the fact that the perturbations of the system (i.e. meals and insulin treatment)

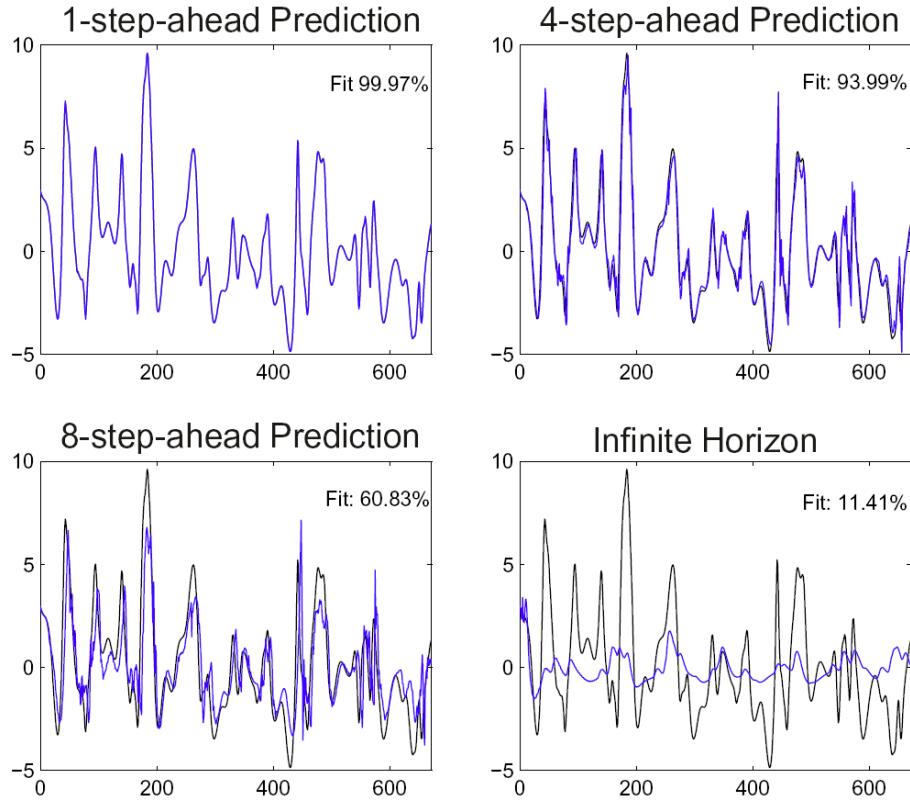


Figure 2.1: Illustration of the degradation of predictive ability with the prediction horizon: upper right  $\tau = 1$  (15 minutes); upper left  $\tau = 4$  (1 hour); lower right  $\tau = 8$  (2 hours); lower left  $\tau = \infty$

have an enormous impact on the diabetic patient glucose, in many cases overcoming the influence of the actual glucose homeostasis, thus needing of more detailed description in the models used. This work showed that predictions with horizons larger than 30 minutes were very imprecise for CGM measurements. CGM involves more frequent sampling, but it introduces large errors in the measured blood glucose concentration, leading to great difficulties in the predictions. Nevertheless, this study introduces separation between meals and its related insulin dose and viceversa for easier identification, which is a concept that will be investigated later in this thesis.

Rollins *et al.* published another self study, where a type 2 diabetic patient (Rollins himself) was monitored for several weeks [73] without insulin treatment. This study focused on predicting glucose by using many different measurements of the patients

metabolism. The patient followed a normal life routine, wearing sensors of temperature, body activity, heat flux and CGM, as well as a diary with all relevant information about meals. They used data-driven models with multiple inputs in order to fit the patient's data. Model validation was performed against CGM estimations, as seen in figure 2.2.

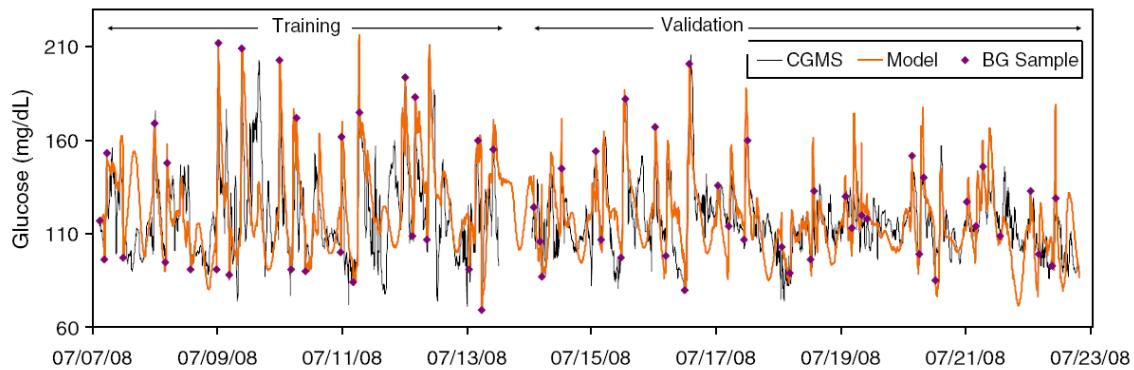


Figure 2.2: Model data fitting and validation of one of the models tested. Seven days were used for data fitting and 9 days for validation

Rollins showed that very good predictions ( $r_{fit}$  up to 0.70) can be achieved if more data from the patient's metabolism and daily rhythm was provided. Unfortunately, these results are not extendable to type 1 diabetic patients since insulin treatments induce enormous perturbations to the system which were not modeled in this paper. Also, only 1 patient was monitored for this study, and diabetic patients may be reluctant to wear all the devices needed for full metabolic monitoring required for this study.

Georgia *et al.* studied the possibility of using a combination of physiological models as inputs to the system, and vector machine for regression as predictor of blood glucose [27]. Their study comprised seven patients with 10 days average monitoring period. They incorporated the possibility of using an exercise model as an input to the glucose predictor. They showed results similar to those of [73], but slightly less accurate predictions to those in[77], but comparisons between these studies are loosely sustained since the two previous studies only simulate one patient. No significant improvement was observed when including the exercise model into the prediction method.

In the work of Cameron *et al.* on prediction of blood glucose for MPC controllers, a comparison of multiple models is performed [10]. Four variations of a very simple data-driven model are tested for their predictive capabilities. The variations affect the meal prediction of the model in different ways, with increasing complexity:

- No meal detection. Meals occur without announcement or reaction of the predictive model.
- Meal detection. Probabilistic detection of meals is implemented in the prediction model.
- Meal detection and anticipation. Meals are predicted and action is anticipated to the actual meal effect on glucose.
- Meal announcement. Meal time is assumed to be input by the user.

Prediction horizons of up to 5 hours are tested for all the variations of the model, both on simulated meal data and 19 days of experimental clinical data. Prediction capabilities are computed with mean error and root mean squared error (RMSE). Predictions are evaluated in Figure 2.3 against prediction horizon on experimental data, and against time for a 2-hour prediction horizon in simulated meal data. This study presents large prediction errors (100 mg/dL) even for relatively short prediction horizon, but prediction performance increases as more information is estimated or provided to the model, being the meal announcement variation the most successful in terms of prediction capabilities.

Palerm *et al.* studied the possibility of using physiology models for identification of diabetic patients [62]. They used a modification of the Cambridge model in data from five diabetic patients in a mixed meal response. The measurements were taken using a YSI 2300 STAT Plus™ (YSI Inc., Yellow Springs, Ohio) very frequently (5 minutes period). This study incorporates an identifiability study of the Cambridge model's parameters, leading to a 4 parameter estimation for a well posed identification problem, at least locally. Given the non-linearities of the Cambridge model, global solvers ought to be used for the data fitting. In this case, Palerm *et al.* used a hybrid method combining both global and local solvers to enhance identification speed [72]. As opposed to the study performed by Johansson and his colleagues, in this case frequent reliable data was fitted, and several patients were individually identified and compared. Yet, validation results were not satisfactory. In figure 2.4 they showed a patient as an example of the performance of the methods used. In the left panel, the patient's data was fitted successfully for all the postprandial period. However, when using the identified model for validation, the prediction of the model did not closely resemble the data gathered.

The work performed by Palerm *et al.* supposed a solid first approach to the identification problem, but still, unsuccessful. The predictions were not able to mimic the response of the patients in different days, which can be caused either by model mismatch

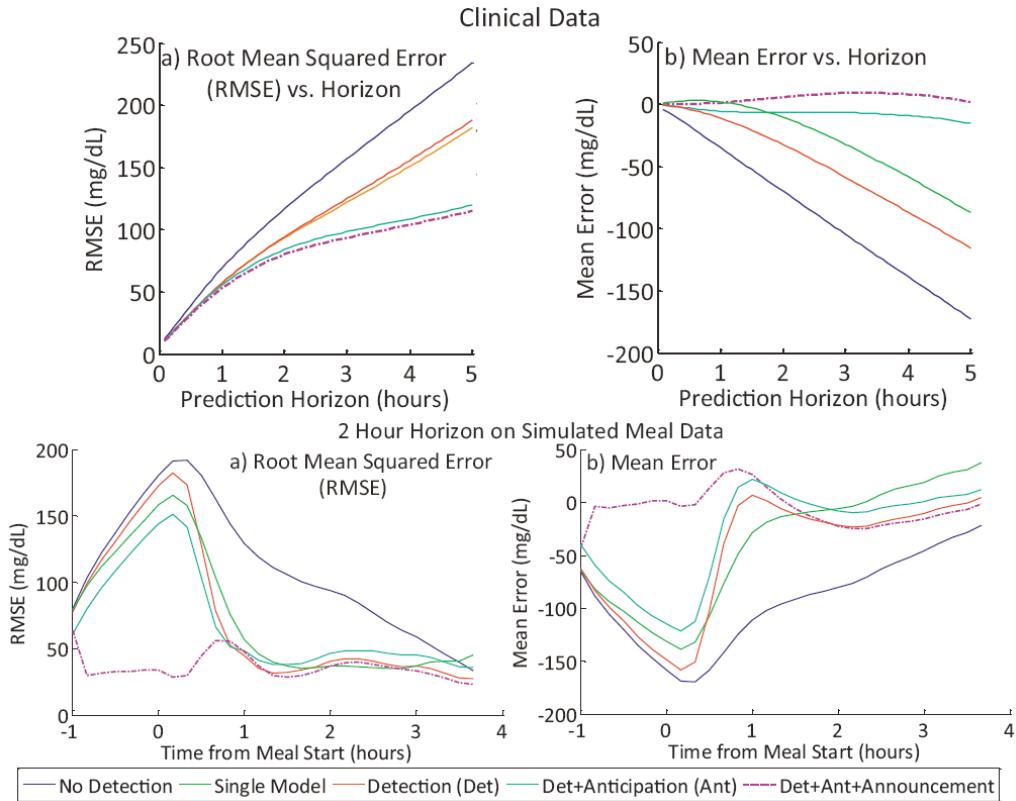


Figure 2.3: Prediction results for all the variations of the model tested by Cameron *et al.* Top pannel presents the prediction error around mealtimes on simulated data. Bottom pannel shows prediction error vs. prediction horizon across all 19 days of clinical data.

to the actual physiology or variations in the parameters within the same patient, or even maybe the incorrect use of data for identification.

Finan *et al.* proposed a comparison of models identification for diabetic patients based on ARX models [25]. They used data from 9 type 1 diabetic patients for identification comparing every prediction to that of a Zero Order Holder (ZOH). The ZOH predictions consisted simply in maintaining the glucose level for as long as the prediction horizon is. The predictions of the models were tested for prediction horizons of 30, 60 and 90 minutes. The study concluded that non significant improvement on the prediction capabilities (Relative Mean Square Error of the validation days) was observed from using ARX models instead of a ZOH. This conclusion is of great importance for identification in type 1 diabetic patients, since it discourages the use of pure data-based models

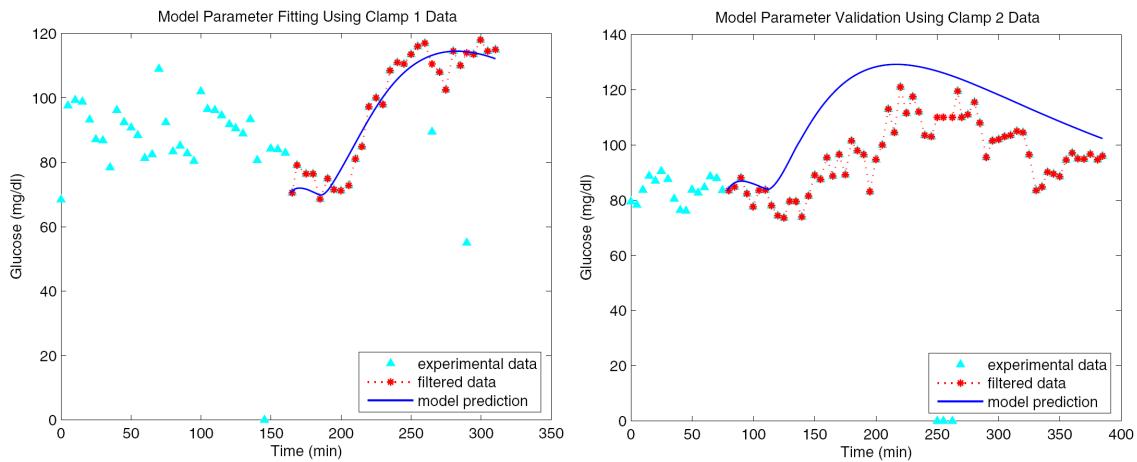


Figure 2.4: Model prediction versus data for one of the patients. Left graph shows the fitting of the data. Right graph shows the validation data on the same patient

in the prediction of glucose concentration. The authors conclude that non-modeled disturbances hinder greatly the performance of the selected models, and that getting reliable quantitative measures of exercise and stress levels would result in better quality predictions.

Concerning the lack of identifiability showed so far by any of the available models in literature, Simone del Favero *et al.* approached the identification problem from a completely new angle [20]. They proposed a method for identification using clinical index instead of purely mathematical quadratic indexes for data fitting. The clinical index consisted in a weighted index based on the glucose prediction danger to the patient. The reasoning behind this kind of fitting is that predictions for clinical patients are not worse as they differ from actual data, but as they unable to detect situations in which the patient may be endangered. For example in the case were a patient is heading towards hypoglycemia and the predictions are those of high glycaemic levels.

## 2.2 Experiment Design for Artificial Pancreas

Model identification is limited by both the model complexity and the quantity of information present of the data used. The current state of models for identification has already been stated. In this chapter we will focus on the improvement of data acquisition

techniques in the diabetes environment.

Predictions of the identified models are highly dependent on the identifiability of the model itself. Identifiability of a model is the facility to repeat an identification of its parameters under similar circumstances. Identifiability of a model is mainly dependent on the models structure, but it can also be improved by performing experiments that ease the process of the identification. The process of preparing data gathering experiments in order to enhance the identifiability of the model is called Optimal Experiment Design. Optimal Experiment Design is rare on the diabetes context due to the fact that data acquisition usually occurs in a clinical environment, where there are physical and ethical limitations to the experimental conditions.

There is a set of standard clinical tests that help clinicians to either diagnose diabetes or to identify clinical parameters of glucose homeostasis:

- *Oral Glucose Tolerance Test (OGTT)*. This test is used for diagnostics of diabetes. The test starts in the morning with the patient following the usual physical activity as in a normal day. The previous evening a meal of 30-50 g of carbohydrates is consumed. In the morning, a fasting blood sample is collected and then a solution of 75 g of glucose in water is drunk over a period of 5 minutes. Blood is then monitored afterwards at intervals. The most common interval though is a single sample (plus the fasting sample) 2 hours after the glucose solution ingest.
- *Intravenous Glucose Tolerance Test (IVGTT)*. This test is used for measuring the pancreas activity either in a healthy person or a diabetic patient. The glucose is administered in a antecubital vein over 60 seconds using a dose of 300 mg/kg. Glucose (and potentially insulin) is monitored starting at the moment of the infusion, with a fast sampling period (5 minutes or less) in the first period of the experiment, and more distanced measured at the end, lasting the whole of the experiment at least 3 hours. This experiment can be modified in order to measure the insulin diffusion and absorption into plasma as well. In that case, an insulin intravenous is also delivered after the meal.
- *Postprandial Glucose Test*. This test is used to screen for diabetes and to evaluate treatments of therapies in diabetic subjects. The patient eats a balanced meal of measured components with a carbohydrate content of 100 g or more. The patient is then monitored for a minimum of two hours, and in case of necessity of testing a therapy, it is applied normally to the patient. This test is the easiest to perform

since it follows the daily routine of the patients with a tight monitoring of blood glucose.

- *Clamp Glucose Test.* Clamp tests are a battery of experiments that consist in artificially keeping blood glucose levels steady by intravenously infusing either glucose solution or insulin. They are difficult to perform as they require of an skilled doctor adjusting the glucose infusion levels to the appropriate concentration in order to lead blood glucose to the desired range. An example of clamp test is the euglycaemic hyperinsulinemic clamp, which consists on administering a large insulin infusion while maintaining glycemia in a controlled range of 90-100 mg/dL. This test permits the quantification of the insulin resistance of the patient without risking the patient to go to hypoglycemia.

Much more complex experiments can be performed for measuring metabolic fluxes in diabetic patients, as for example methods with radioactive tracers can help in the characterization of emptying rates in the digestive tract. In 2003 Basu *et al.* described a triple-tracer method for assessing the postprandial glucose metabolism [2], which was later used by Dalla Man *et al.* for developing a model of the gastrointestinal absorption of glucose [18]. This type of tests though are extremely difficult to perform, even with single tracers; they require specialized instruments for the detection of the isotopes, and they are also much more intrusive to the patients involved

The field of optimal experiment design for diabetes has been developed greatly by Federico Galvanin and his team. They published in 2009 a design of a postprandial monitoring experiment with a variating insulin infusion throughout the post-prandial period. The approach they used leads to a results dependent on the model used, and they based it in the Cambridge model described before. They based their design on the optimality of 5 parameters of the endogenous glucose-insulin model, attempting to minimize the uncertainty in the identification of those parameters. After *a priori* sensitivity analysis, they realized that two of the 5 parameters are highly correlated in the postprandial period and cannot be identified together so they reformulated the problem as the optimization of identifiability of 4 parameters, one of which was the ratio between the unidentifiable parameters. The outcome of the experiment design was a 2 meal monitoring where they tunned the CHO content of the meals and the insulin infusion from an insulin pump. The insulin infusion was defined as a piecewise constant function with 12 levels and 11 level switching times. The outcome of the experiment was relatively complex, as shown in Figure 2.5.

The qualitative result shown in the experiment design proposed by Galvanin *et al.*

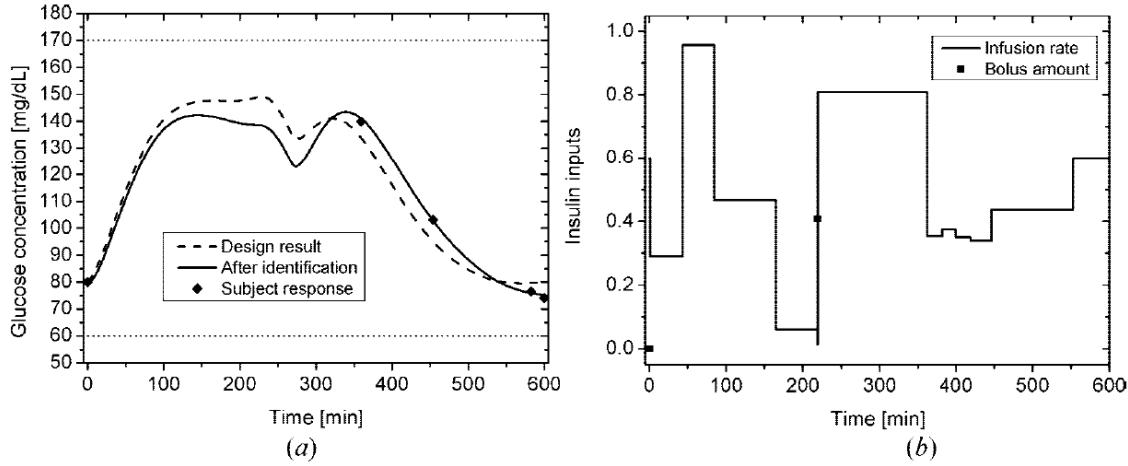


Figure 2.5: In the left panel the result of a simulation of the virtual patient is shown for the two meals period. In the right panel can be seen the piece-wise function result of the experiment design for the insulin infusion

is actually difficult to implement in a real experiment. Given that the models used are inaccurate, it is not advised to take the results of the experiment design accurately. Nevertheless, the results of this paper are very useful: it demonstrates that identifiability in diabetes is increased for longer experiments, and it shows that exciting the insulin input, statistically satisfactory parameter estimation can be achieved.

Later on, Galvanin *et al.* published another work on optimal experiment design using on-line updates to the identifiability of the parameters. They realized from previous experimentation that model-based experiment design was too dependent on the model used. They decided to approach the issue in this paper introducing a model mismatch, based on the simple idea of simulating the data with one literature physiological model (in this case de UVA model) and identifying with another (the Cambridge model). The experiment design outcome is very similar to that of the previous work, with step shifts on the insulin infusion to the patient. The new design is depicted in Figure 2.6.

Galvanin *et al.* concluded the study with a feasibility analysis of the experiments designed. When introducing the model mismatch, the classic model based experiment design was unable to provide feasible experiments. When using the on-line experiment redesign and identification, feasibility was achieved, and identifiability of the parameters was greatly improved. The amount of glucose and insulin delivered with the new method was much more aggressive. Delivering aggressive treatments and overextending on the

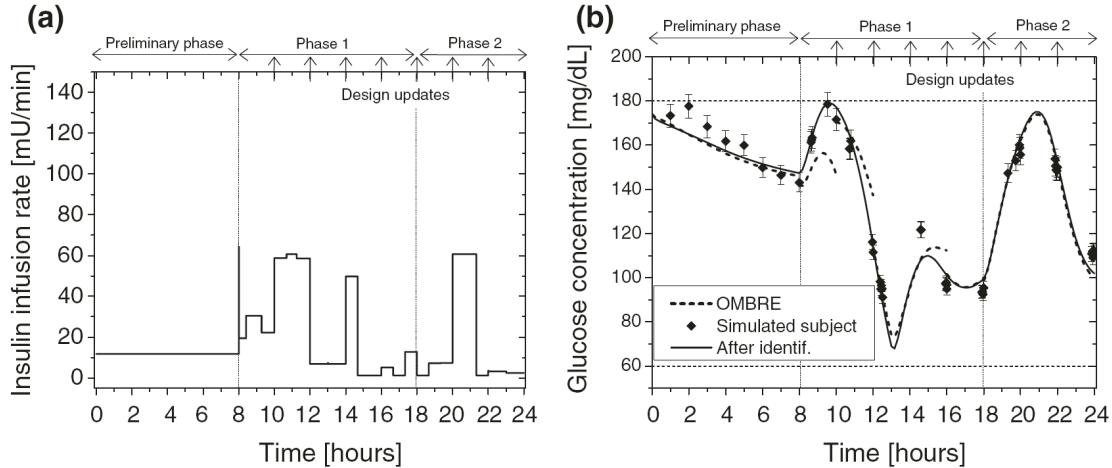


Figure 2.6: In the (a) panel the insulin infusion is shown for the whole duration of the experiment. In the right panel the blood glucose concentration of the patient is displayed, including the identified model's predictions and simulations.

metabolic excitation on a diabetic patient can be very dangerous when implementing the experiments, and results as such have to be double-checked to avoid danger to the patients.

Cobelli and Thomaseth published a study of optimal inputs for the Bergman minimal model using also optimality of identification of the parameters of the model [14]. They focus on the design on finding optimal continuous functions of excitement to the endogenous model. They concluded that the outcomes of the experiment are very interesting from the modeling point of view, but of limited relevance in the diabetes context. The optimal inputs obtained are too difficult to realize in clinical practice. They also imply that delayed insulin inputs can be useful for identification of certain parameters of the model.

Few other relevant works on the field of experiment design have been published. However, modifications on the traditional inputs with the aim of improving data quality is often seen in research of the artificial pancreas. Percival *et al.* for example used delayed insulin bolus (2 hours after the meal) for the data acquisition of glucose profiles for a control algorithm [65]. The scheme for the experiment monitoring and inputs can be seen in Figure 2.7

Even though the clear objective of Percival *et al.* when shifting the bolus time was to

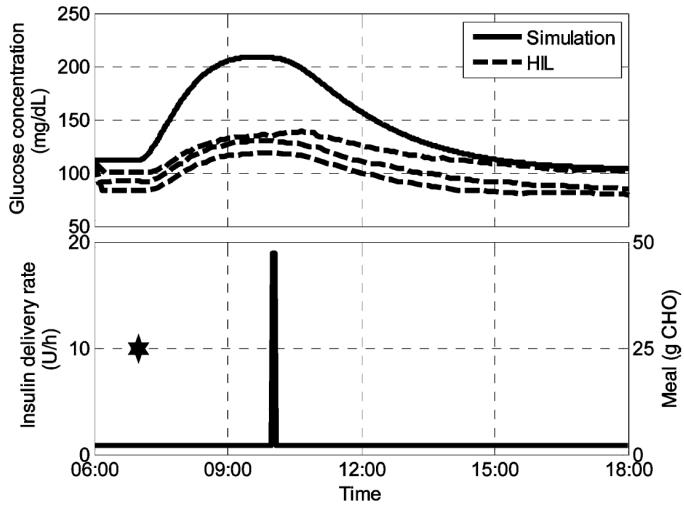


Figure 2.7: Top panel shows the glucose simulation and CGM estimations for several calibration methods. Bottom panel corresponds to the inputs of the system for the same period. The star points at the meal ingestion time, and the impulse is the insulin bolus time.

enhance the identifiability of the model, there is no further justification to it. In general, diabetes research is lacking on simple experiments designed for the day-by-day routine of the diabetic patient; simple enough experiments that will imply the patient participation and understanding of the disease he is living with, and the research related to it.

## 2.3 Uncertainty and Interval Identification in Diabetes

Individualization of models for diabetes is greatly hindered by the fact that few models in literature consider intra-patient variability. In the models described in Section 1 every model parameters stands for a physiological constant or function, and it is assumed to be characteristic to each patient. Most of the model proposed are published with some distribution in their parameters that resemble population values. These population values are an accepted quantification of the inter-patient variability for the parameter evaluated. However, characterization of a patient is dependent on the current metabolism of the patient, which is highly variable. This intra-patient variability has been regarded in literature with little detail.

One of the main sources of metabolic variability is the circadian rhythm of the patient. It is well established that the circadian variations of diabetic patients cause above average glucose levels in the early morning (event known as “Dawn Phenomenon”), among other effects. The circadian effect has been quantified in the insulin dependance by Scheiner and Boyer [76], showing a higher insulin relative dosage needed in the morning in type 1 diabetic patients. Higher insulin demand combined with constant basal infusion rates directly results in high morning glycemia, such as the reported in the Dawn Phenomenon.

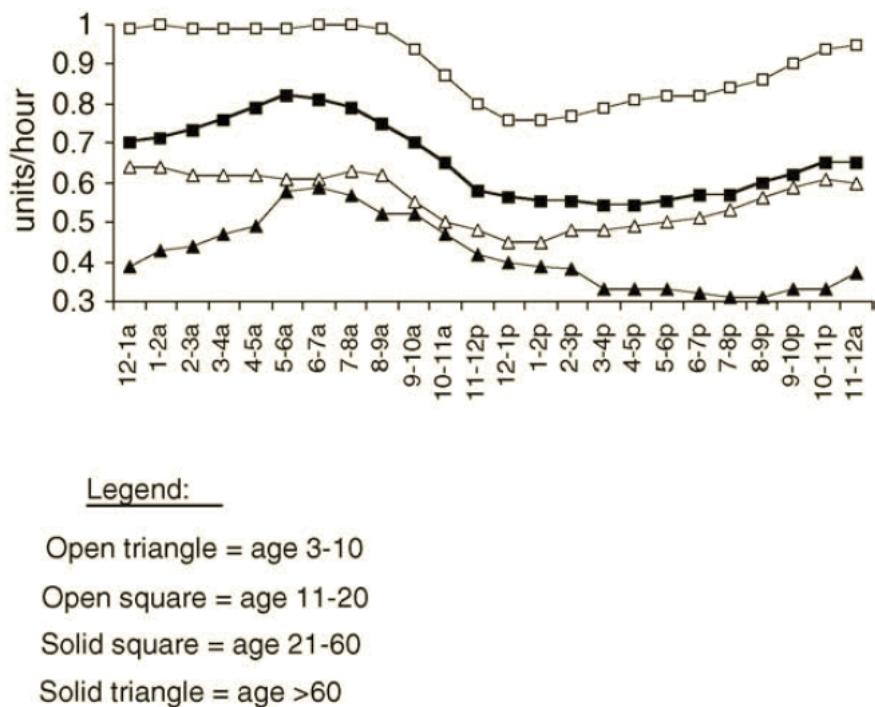


Figure 2.8: Average hourly basal rate values by age group

Figure 2.8 shows how insulin basal rate decreases for every age group after breakfast time, showing a periodic behavior of one of the better known physiologic parameters in diabetes: the insulin sensitivity. Furthermore, this work emphasizes the great difference in the population values when distributed by age groups, and suggests that a single probability distribution is not sufficient to characterize one parameter for the type 1 diabetic population.

Understanding that physiologic parameters are not correctly characterized by only a number, several authors in recent years have been adapting physiological models in

order to simulate patients with variability inherently considered. Using interval analysis methods and models is a new way of considering those parameters. Interval models consider internal parameters as “ranges” instead of punctual numbers. The parameter at a given time is actually unknown, but it is known to be comprised within the boundaries of the interval that defines it. The outputs of interval models are also intervals, and in the case of diabetes, the main output being glucose, means that we have to work with uncertain ranges of glucose.

Calm *et al.* published work based on the Cambridge where modal interval analysis is applied for the plasma glucose prediction to face uncertain food intake and patient parameters such as insulin hepatic and peripheral sensitivity [9]. The model was later validated against Monte Carlo simulations proving it to be a much more efficient in computation [8]. The interval model developed was tested against several scenarios where parameters in the insulin absorption and meal absorption models were considered to have between a 10% and 20% variation.

In Figure 2.9 a set of simulations with the proposed model is shown for a three meal experiment and overnight period. In general, wider variability considered in the parameters yield wider glucose outcomes. It can be observed that even relatively small variations considered in the parameters of the model (10%) yield wide glucose prediction bands (approximately 100 mg/dL). Considering that sensitivity parameters of a real diabetic patient can be much more variable than the 10% considered by this study, even larger glucose prediction bands are expected when characterizing diabetic patients in an experimental setting.

Another approach to interval models simulation was undertaken by de Pereda *et al.* on the Cambridge model [19]. They tackled the uncertainty analysis by using monotonicity analysis of the states of the model and its critical points. The bounds provided by this type of analysis are very similar to those of the proposed by Calm *et al.*, but with the advantage of considering more parameters subject to uncertainty, as for example the insulin absorption lag time  $t_{maxI}$ . They also proposed the use of the uncertainty model as a short-term prediction tool, showing that for smaller prediction horizon scenarios, band widths do not grow excessively, and are suitable to use in regulation algorithms such as Model Predictive Control. The work done by de Pereda *et al.* is very recent (2012), and it still has to be tested as a suitable method for identification and control.

To our knowledge, Kirchsteiger *et al.* published the only identification procedure with consideration of uncertainty so far [44]. This work was initially based on a simple

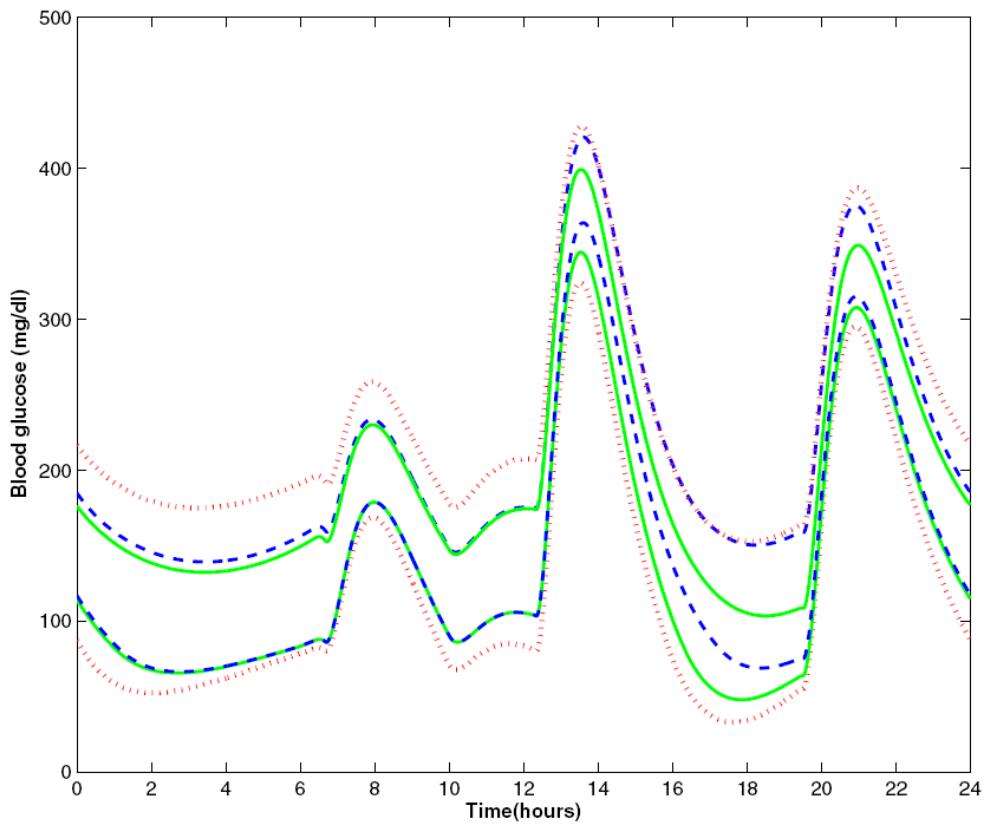


Figure 2.9: Envelopes obtained with the interval model for three different scenarios: 10% variation in insulin sensitivity parameters (solid line), 20% variation in insulin sensitivity parameters (dotted line) and 20% variation with modified insulin bolus (dashed line)

data-based transfer function model with 6 interval parameters. The identification was performed on data from 9 patients and three postprandial periods for each patient. After some trials, a simplified model structure was proposed for identification, due to identifiability issues, reducing it to only 4 interval parameters. The identification was posed as an optimization problem, where not only the fitting error was identified for 3 days, but also the standard deviation of each parameter considering the three postprandial periods. This procedure kept the parameters from resulting in too wide intervals, while still fitting the data of the patient.

Figure 2.10 shows one case of the identification performed by Kirchsteiger *et al.*. The model bands encompass all the data of the patient. This is not the case for all the patients,

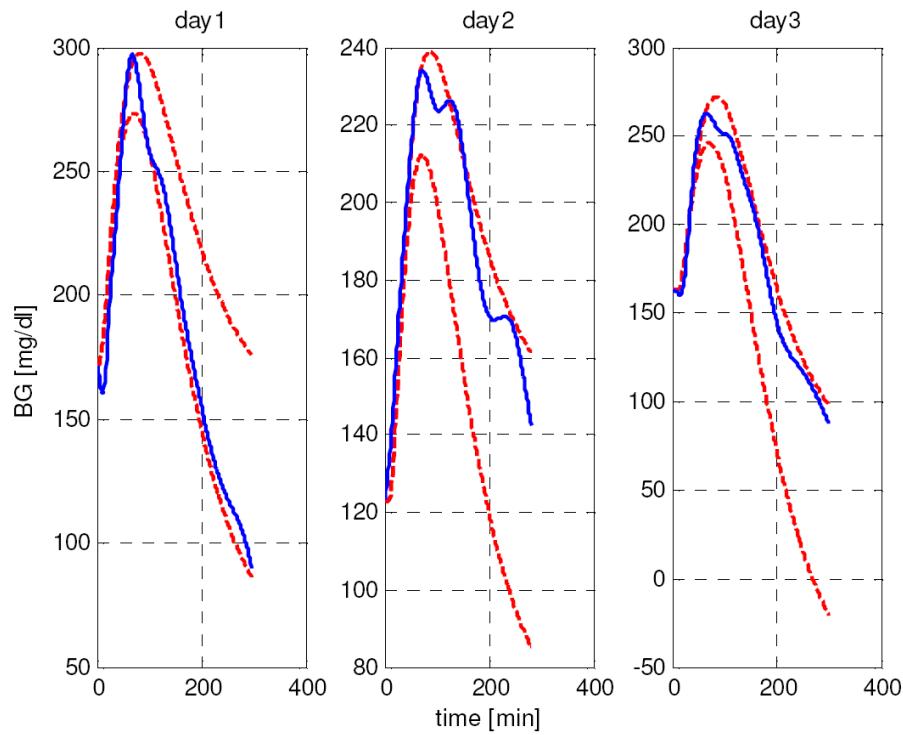


Figure 2.10: Three-day identification of a diabetic patient’s breakfasts. Blue solid line represents the spline-interpolated experimental glucose measurements. Red dashed lines are the upper and lower bounds given by the model

but it is a good example of how a satisfactory identification looks. Even though the idea of identification of 3 separate days with common parameters is appealing, the study lacked of the proof of cross-validation on the data. Validation on single meals was performed, but not using the interval parameters, neither selecting a day *a priori* as the validation day. It is unclear whether the day used for validation was just the 3rd day chronologically or the best case validation. Also, the data source on this study was treated without consideration of the possible errors introduced, which is a very important issue if not using glucose reference data.

# Chapter 3

## Parameter Estimation

Identification of a model's parameters is an inverse problem. As such, identification consists in the characterization of the inputs and parameters of a system given the outputs. Frequently, *a priori* information on the system can be used to narrow the search path of the problem, but this type of information is not enough to allow for a complete derivation of the models. In the glucose-insulin model identification paradigm, the inverse problem is usually reduced to finding the parameters of a model which characterize an individual given the glucose profile of that patient, and the patients conditions such as meal characteristics and insulin treatments.

Parameter estimation is usually posed as an optimization problem. Its solution relies on optimization algorithms attempting to optimize an index of the fit of the model's output to the available data. Usually scalar indexes are considered for characterizing the deviation of the model from the data. Such index can be formulated, for example, as a quadratic error depending on a set of parameters  $p$  that has to be minimized:

$$J(p) = \sum_{i=1}^N (y_i(p) - \tilde{y}_i)^T Q_i (y_i(p) - \tilde{y}_i) \quad (3.1)$$

where  $y_i(p)$  are the model predictions and  $\tilde{y}_i$  are the experimental measurements, therefore, the data to be fitted.  $Q_i$  is the data weighting matrix, which permits to fit more accurately some data in detriment to other data samples.  $Q_i$  is usually chosen as a diagonal matrix, due to the assumption that data samples are independently correlated

to each other. The choice of the diagonal values in the matrix are usually referred to as “weights” of the data samples. This quadratic error index reflects the most classical approach to parameter estimation.

A usual compromise choice for the ponderation matrix is to make the relative errors of each data sample weight the same in the cost index in order to fit all data points equally in the time axis. This is achieved by forcing the diagonal of  $Q_i$  to be equal to  $1/\tilde{y}_i^2$ . This type of ponderation matrix has the disadvantage of underestimating large errors in big values (hyperglycemic region) of the output of the model than in small values (hypoglycemic case). Usually it preferred that all errors are normalized in the optimization index, and this is achieved with a different ponderation matrix, like for example the case where all the elements of the diagonal are equal to  $1/\max(\tilde{y}_i)^2$ . This way the contribution of every data sample is standardized in the glucose concentration axis.

The most common choice of optimization methods for parameter estimation are simple local optimization algorithms, eventually iteratively for easier convergence. There are many drawback to the use of local optimization though:

1. The initial value of the parameters to be chosen for the identification heavily influences the final outcome of the optimization. The choice of these values relies strongly in guesswork.
2. Convergence to the global optimum of the problem is achievable, but not guaranteed.
3. There may be several sets of parameters that yield similar optimization index values. This problem may happen with a very flat solution region but it can also happen for different local “valleys” of the index function.

These problems are very much present for every model shown in chapter 1, and every diabetes model in literature whatsoever. The solution of the identification problem in diabetes has to be focused in the practical aspect of the optimization. It is established in the scientific community that modeling of the glucose system, even though are often based on physiological aspects of the glucose metabolism, are far away from being accurate. The approach to patient identification is much more focused in the efficacy of the predictions of the model rather than interpretation of the identified parameters. It is then fine for an identification in the diabetes paradigm to achieve a flat solution in the parameter space, if that solution is a global optimum. It is not acceptable though falling

on local optimum values of the index, which are likely to provide worst predictions than a global solution. Also, searching for a single optimal value of the parameters set is often not the best option, since there may be a full set of acceptable solutions that fall under feasibility conditions. Use of global search optimization methods are encouraged for this kind of problems.

Identification feasibility or identifiability of a model is the property that determines whether the model is suitable to characterize a set of data and in which terms. Identifiability analysis must be performed on an identification study to discard possible incongruences of the model's structure or incompatibilities with the data. Improving identifiability of the models proposed is one of the goals of this thesis, and optimal experiment design will be introduced as one of the main aid tools for diabetic patient identification.

Given that the endogenous glucose system is subject to great parametric variations, uncertainty on the parameter estimation has to be assessed. In local classical methods, uncertainty on the estimate is evaluated (if at all) by using asymptotic properties of the estimator that rely on a number of doubtful assumptions, like assumption of linearity of the parameter's space. This type of uncertainty measurements are not reliable for the diabetes environment, and the use of new estimation methods that naturally take uncertainty into consideration is at its utmost importance.

A possibility for solving the problem of uncertainty in the parametric space is interval analysis. Interval arithmetic and calculus has been a very active field in scientific computation for the last 40 years (see for example [58]). Intervals are defined as a new kind of numbers, on which all classic arithmetic operations can be performed. An interval  $[x]$  is defined as a closed, bounded and connected set of numbers:

$$[x] = [x^-, x^+] = \{x \mid x^- \leq x \leq x^+\} \quad (3.2)$$

$x^-$  (lower) and  $x^+$  (upper) are the limits or boundaries of the interval, and are the scalar magnitudes that define it. Another scalar characteristic of the interval that is derived for the boundaries and will be of special interest is its width  $w([x])$ , and it is defined as follows:

$$w([x]) = x^+ - x^- \quad (3.3)$$

Interval analysis do not impose any restriction on the intern possible values of the interval further than continuity. This feature makes interval computation very suited to work with uncertain magnitudes naturally by identifying not only scalar values of the model's parameters, but interval counterparts for every parameter in the models presented in previous chapters. Identifying interval parameters allows for the characterization of many sources of uncertainty:

- *Parametric uncertainty.* Parameter feasibility vicinity is directly identified by using intervals and no further assumptions are needed.
- *Model mismatches.* Unmodelled dynamics can be (artificially) included as larger intervals in the interval identification.
- *Measurement errors.* Sensor-induced uncertainty is common in the diabetes paradigm, especially in continuous monitoring systems. Interval identification can be used either for consideration of this error (assuming the error distribution is known) in the identification algorithm, or by aggregating the uncertainty in an identified parameter.
- *Parameter variability.* Parameters that are considered in model as scalar values can in fact be variant in time or in the states space, and as such identified with uncertainty in interval quantities.

Interval parameters in glucose metabolism models produce interval glucose outputs. The power of using interval analysis for simulation relies in the fact that glucose envelopes of an interval model will bound all possible *scalar* simulations of the (scalar) parameters within the bounds of the interval parameters used. This guaranteed outcome of an interval model is true not only for simulations of single parameters, but also for any combination of parameters included in the intervalization of the model. In Figure 3.1, a very simple simulation of the Cambridge model is presented, picturing the envelopes in blue. In green a thousand simulations of the scalar model are plotted where the uncertain parameter is randomly generated within the boundaries of the interval parameter previously simulated in a Montecarlo simulation. As can be clearly observed, every possible solution of the original model is contained within the envelopes of the interval model, as expected.

Intervalization of a given scalar-based model is not straightforward. The computation of both envelopes has to be taken in consideration and depending on the nature of the

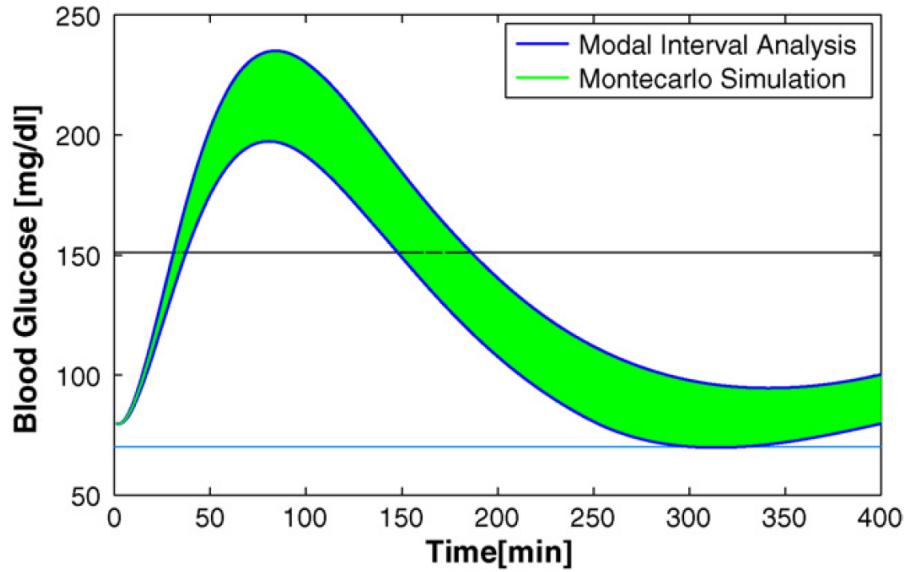


Figure 3.1: Envelopes obtained for the interval simulation presented by Calm *et al.* [8] of the Cambridge model and considering 10% uncertainty in the meal size

parameters considered uncertain, the calculations can be very complicated. Fortunately, interval counterparts of every model proposed in chapter 1 can be found in literature. In this thesis the intervalization of glucose models will not be reviewed, and instead, it will be taken as given. An extensive review of the models proposed in here was performed by Maria Garcia-Jaramillo in her PhD thesis [40], including the intervalization process of every equation and the final interval system of equations of every model. More recently, the Cambridge model was further improved in its interval form by de Pereda *et al.* [19], including the possibility of simulating as intervals more parameters than the interval version presented by Garcia-Jaramillo. Both versions of the model will be mentioned in the following thesis.

In the rest of this chapter I will review the optimal experiment design and identifiability improvement for the models described in the previous chapters. Then I will look on the interval identification process, and the most used identification algorithms are used in literature. Finally I will quickly enumerate the optimization tools and software used for the development of the work that produced this thesis.

### 3.1 Optimization of Identifiability

Identifiability analysis can be divided into two stages [81]: identifiability *a priori* and identifiability *a posteriori*. Identifiability *a priori* is the feasibility of identification of the model's parameters considering infinite noise-free data is available. It is a structural property of the model and the lack of it represents a model that can produce the same output with several sets of parameters given a determined input. Identifiability *a posteriori* incorporates the data into the identifiability analysis. It analyzes the influence of data noise and uncertainty in the estimation of parameters, and its main outcome is the reliability of an identification.

*A posteriori* identification methodology often relies in the analysis of the model's parameters sensitivities to the output variations. In the following, the classic method based on the Fisher Information Matrix will be reviewed. Analyzing with detail the Fisher Information Matrix (FIM), all information about local identifiability of the model can be extracted, due to the fact that this method's application implies linearization of the model.

To understand the method, let's take the index defined in equation 3.1. The statistical estimation of the index for a set of parameters slightly different of the optimal is given by:

$$E[J(p + \delta p)] \cong \delta p^T \left[ \sum_{i=1}^N \left( \frac{\partial z}{\partial p}(t_i) \right)^T Q_i \left( \frac{\partial z}{\partial p}(t_i) \right) \right] \delta p + \sum_{i=1}^N \text{tr}(V_i Q_i) \quad (3.4)$$

where  $V_i$  represents the covariance matrix of the measurement errors ( $Q_i$  is typically chosen as  $V_i^{-1}$ ),  $z$  is the model output and  $N$  is the number of data samples. The term between brackets is the Fisher Information Matrix and it expresses the quantity of information contained in the experimental data, as explained in detail by Ljung [53]:

$$\text{FIM} = \sum_{i=1}^N \left( \frac{\partial z}{\partial p}(t_i) \right)^T Q_i \left( \frac{\partial z}{\partial p}(t_i) \right). \quad (3.5)$$

The terms  $\partial z / \partial p$  are the sensitivity functions of the each parameter  $p$  and they are of great importance for the evaluation of the practical identifiability since they gather the parameter influence on the output of the system. The FIM is a square matrix with

dimension equal to the number of parameters that are to be identified. The inverse of the FIM is an approximation of the covariance matrix of the estimation error of the model parameters:

$$C = FIM^{-1} = \left[ \sum_{i=1}^N \left( \frac{\partial z}{\partial p}(t_i) \right)^T Q_i \left( \frac{\partial z}{\partial p}(t_i) \right) \right]^{-1} \quad (3.6)$$

As The diagonal of  $C$  contains the information of the confidence interval in the estimation of every parameter. The statistic expected value of the error of an estimation (which is a measure of the confidence interval of the model's parameters) is actually bounded by the matrix  $C$  following the Cramer-Rao inequality, as introduced originally in the classic references [16] and [69].

$$E [\hat{p} - p] [\hat{p} - p]^T \geq C \quad (3.7)$$

where  $E$  stands for the statistic expected value and  $\hat{p}$  is a estimation of the parameters  $p$ . For further explanations and proofs of the Cramer-Rao inequality, the reader is referred to Ljung's work [53]. The extreme value of the inequality, where the confidence intervals match the inverse of the FIM, is the Cramer-Rao Bound. Given that the FIM is know and that it is invertible, the coefficient of variation for the parameter  $p_i$  being identified is calculated as:

$$CV_i = \frac{\sqrt{C_{ii}}}{p_i} \quad (3.8)$$

The interpretation of this Cramer-Rao limit is simple: if, for example, a parameter  $p_i$  has a CV of 0.4 it means that, for the measurement error which variance was considered in the calculation of  $Q_i$ , successive parameter estimations will have a deviance of 40% of the true parameter value. There is more useful information to be drawn off the FIM, like the correlation matrix. This matrix, the elements of which are the approximated coefficients of correlation between the  $i^{th}$  and the  $j^{th}$  parameters, is defined as:

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}} \quad (3.9)$$

Analyzing the correlation matrix gives and idea of the compensation effect of the changes in the values of the parameters over the model output. If two parameters,  $p_i$  and  $p_j$ , are highly correlated, a change in the output due to a change in parameter  $p_i$  can be hidden by the appropriate change in  $p_j$ . Very strong correlations between different parameters are a display of poor identifiability of a model. A correlation of value 1 between parameters is a sign of a structural problem of identifiability of the model. Thus, FIM analysis can also be used in the process of *a priori* identifiability analysis.

In practical terms, the identifiability analysis must be done applying some simplifications. The sensitivities of the parameters have to be calculated by approximating the derivatives of the output with first order approximations. In general, application of the analytic expression of the derivative function is the correct way of performing the FIM calculation. With the models exposed before though, the analytical expression of the derivative with respect to the parameters of the model's output is not feasible.

Usually, the sensitivity function is calculated by linearizing the model around the nominal parameter, and obtaining the symmetric first order difference. In practice, two simulations of the model are calculated, one with a positive variation of the selected parameter, and the other with a negative variation and then the sensitivity function is calculated with equation (3.10).

$$S_p = \frac{z(p + \Delta p) - z(p - \Delta p)}{2 \cdot \Delta p} \quad (3.10)$$

Also, if the FIM results to be singular it can not be inverted, and that is a sign of non-identifiability. In fact, it is a sign of *a priori* non-identifiability, so that analysis is being carried out with this method as well. Usually, when working with noisy measurements, it is really difficult to get a singular FIM, and yet it will be difficult to identify any parameter because the FIM is bad-conditioned. The condition number of the FIM has to be analyzed to overcome this computation problem, and its magnitude checked to analyze if it permits inversion of the matrix and therefore identifiability of the model.

Identifiability of a model depends on model structure, and the data used for identification. Identifiability of a model can be improved by conditioning the data used for identification, and designing optimal experiments for the model used. So far, identifiability has been analyzed as a property of a parameter, quantified as the confidence interval in the estimation of each parameter. That information is obtained from the Fisher

Information Matrix (or better, from its inverse), which summarizes the information of all the parameters of the model for a given experiment. The problem of optimal experiment design can be expressed then as an optimization problem of finding the minimum value of a certain scalar function of the FIM that optimizes the identifiability of the model. That scalar function is called optimality criteria, and its general expression is:

$$j(\Xi) = \phi[F(p, \Xi)] \quad (3.11)$$

where  $\phi$  is a scalar function. The evaluation of the Fisher Information Matrix is a function  $F$  of  $p$ , the parameters vector, and  $\Xi$ , the experiment conditions to be optimized.

There are several criteria that can be used in this case, as seen in [26]:

- D-optimality, in which the scalar function chosen is the determinant of the FIM. The three following equations are equivalent, and all of them define the D-optimality criterion and whether it has to be maximized or minimized in order to improve identifiability:

$$\Xi = \text{argmin}_{\Xi} j_D(\Xi) = \text{argmin}_{\Xi} \det(F^{-1}(p, \Xi)) \quad (3.12)$$

$$\Xi = \text{argmax}_{\Xi} j_D(\Xi) = \text{argmax}_{\Xi} \det(F(p, \Xi)) \quad (3.13)$$

$$\Xi = \text{argmax}_{\Xi} j_D(\Xi) = \text{argmax}_{\Xi} \ln(\det(F(p, \Xi))) \quad (3.14)$$

- E-optimality, in which the function is the smallest eigenvalue of the information matrix, and it has to be maximized.
- A-optimality, in which the problem is solved by maximizing the trace of the information matrix.

In order to better understand the meaning these criteria an geometrical analogy is needed. If every parameter is placed in an axis of the geometrical space, then the region defined by the confidence intervals of each one of the parameters defines an ellipsoid the axis of which are given by the eigenvalues of the inverse of the information matrix. Given that the objective of the optimization is to minimize all the regions of confidence, every axis of the ellipsoid have to be minimized, or equivalently, the volume of the ellipsoid has to be minimized, and that volume is exactly the determinant of the inverse of the FIM. The rest of the criteria have similar meanings that are summarized in Figure 3.2.

The question of which criteria to be used arises now. The D-optimal is the most

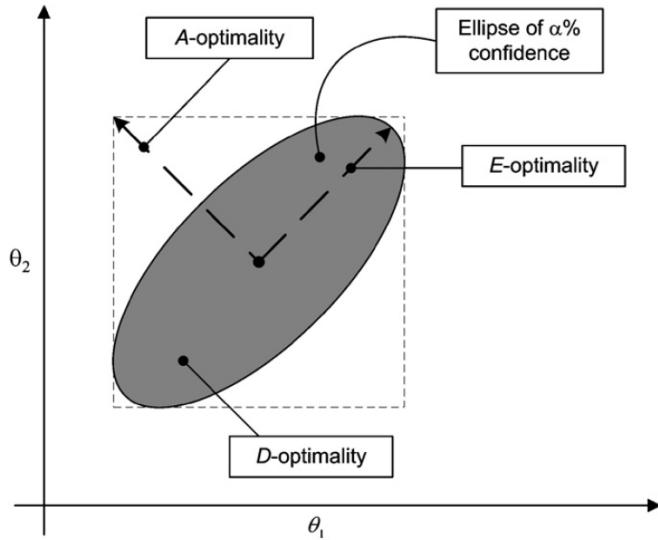


Figure 3.2: Each criterion reduces the confidences intervals attempting to minimize one singular scalar value [26]

used of the three standard criteria cited above. This is due to some exclusive appealing properties of the criterion [26]:

- Easy geometrical interpretation, as seen in Figure 3.2.
- Invariance with respect to non-degenerated transformation applied to the model parameters, such as rescaling. This property is applied in equation (3.14) in order to work, during the optimization process, with smaller quantities.
- Yielding to optimal experiments which correspond to replications of a small number of different experimental conditions.
- Optimal experiment design yields to non-singular FIM.

The main drawback this criterion has is that it gives too much importance to the parameter to which the model is most sensitive. Geometrically, this problem is equivalent to the idea of trying to minimize the volume of the ellipsoid by reducing mostly its bigger axis.

There are other criteria available for different purposes, like the modified E-optimality, that tries to maximize the FIM condition number, aiming to make the confidence ellipsoid as spherical as possible [80]. The correct application of this criterion would be in a case of strong parameter correlations, and the design will yield to a decoupled identification in parameters. That is not the case in this thesis though. D-optimality will be the one used for the experiments designed for diabetic patients, and for every model tested.

The optimization problem associated is non-linear, and global solvers are suggested for its solving. The choice of experimental parameters for the optimization is as wide as experiments exist. One of the most common choices is to optimize the sampling times of a given experiment. The choice of these parameters usually relies on the person performing of the experiment design, but it conditions the characteristics of the optimization. Too many parameters may lead to long, unrealistic convergence times on the optimization algorithm, or even non convergence at all. Also, if many experimental parameters are introduced experimental setup may be too complicated to perform. Given this arguments, and especially in the diabetes context, the number of parameters to be optimized should remain as low as possible for achieving the minimum identifiability needed.

## 3.2 Interval Identification

Interval use for identification has been developed since the 1990s due to its ability to cope with uncertainty either in the structure of the model and also in the measurements and parameters. Focusing on the measurement error, a parameter estimation methodology outstands over all others: the error-bounded estimation.

### 3.2.1 Error-bounded Estimation

Error bounded estimation assumes that an error exists in the measured variable, and that there exist a trusted estimation of this error. The estimated error must be within bounds that are acceptable for the good working of the system. One may define the error as:

$$e(p) = \tilde{y} - y(p) \tag{3.15}$$

where  $\tilde{y}$  is the vector of experimental measurements, and  $y(p)$  is the model output, which is dependent on the parameter vector  $p$ . The problem of bounded error estimation considers that the output errors lie within acceptable bounds:

$$e^- \leq e(p) \leq e^+ \quad (3.16)$$

where  $e^-$  and  $e^+$  are the lower and upper acceptable bounds of the error. The identification problem relies on finding all possible values of the parameter vector  $p$  that produce outputs that fall within the acceptable bounds. The subsequent problem is a set inversion problem, and it can be solved by set inversion algorithms like the SIVIA (Set Inversion Via Interval Analysis) algorithm presented in Jaulin *et al.* [41].

The SIVIA algorithm divides the parameter space into interval “boxes”, i.e multidimensional intervals, and evaluates them in the output space for compliance with the desired characteristics. Given that the interval analysis produces guaranteed solutions of the output of the model to all the values of the parameter space inside the evaluated box, classification of the boxes evaluated can be divided into three categories:

- Guaranteed solutions. Boxes that fall within the boundaries of acceptable constraints.
- Guaranteed non-solutions. Boxes that results unacceptable for the constraints given.
- Indeterminate solutions.

The algorithm works iteratively classifying the boxes into these three categories and dividing indeterminate boxes into smaller boxes for further classification. The algorithm searches all the parameter space. The resolution of the algorithm is given by the acceptable size of the indeterminate boxes in the parameters space. These boxes compose the frontier of the parameter space that produces outcomes of the model in the acceptability constraints.

The SIVIA methodology is summarized in Figure 3.3. The images in the output space are related to boxes in the input (or parameter) space. The boxes in the input space are divided each time they are classified as undetermined. The first evaluation to be performed correspond to the whole parameter space, which will be evaluated as

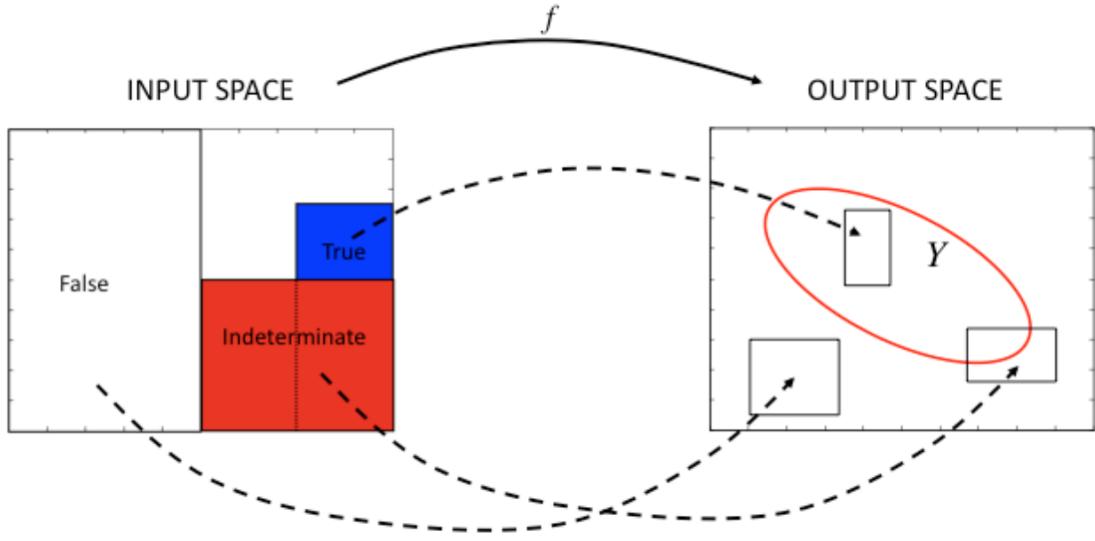


Figure 3.3: Basics of the SIVIA algorithm. The boxes that produce images in the output space that fall within the acceptability range (red ellipse) are classified as true solutions, in blue. Boxes out of the acceptability zone are classified as false solutions and marked in white. Red boxes are undetermined.

undetermined if the acceptability criterion embraces only some part of the output space. the parameter space is then divided using a predefined criterion to the choice of the user into smaller boxes, and reevaluated. The algorithm then works as a tree-search algorithm, finishing a branch when it is classified into true or false solutions, or if the resolution threshold for the search boxes is met.

The method is clearly understood when looking at an example. A classic example of the performance of error-bounded estimation will be reproduced next in order to better display this set inversion problem. A simple two parameter model is to be fit to data with consideration of uncertainty in it. The data for the example was first found in [56]. The model to be fit responds to the equation:

$$y(p_1, p_2, t_i) = 20 \exp(-p_1 t_i) - 8 \exp(-p_2 t_i) \quad (3.17)$$

where  $p_1$  and  $p_2$  are the parameters to be identified with uncertainty. The data for the identification was sampled at different times, and it presented different relative error known *a priori* following  $[e] = [-e_{max}, e_{max}]$  with  $e_{max} = 0.5 |\tilde{y}| + 1$ , where  $[e]$  is the

vector of acceptable intervals for the error and  $\tilde{y}$  are the experimental measurements. Notice that the error estimation is considered as given for the problem. The vector of sampling times and related experimental measurements with the error interval associated to each sample are displayed in Table 3.1. The same samples are displayed in Figure 3.4 in a Cartesian plot for easier visualization of the problem.

$i$	$t_i$	$\tilde{y}$
1	0.7	[2.7, 12.1]
2	1.5	[1.04, 7.14]
3	2.25	[-0.13, 3.61]
4	3	[-0.95, 1.15]
5	6	[-4.85, -0.29]
6	9	[-5.06, -0.36]
7	13	[-4.1, -0.04]
8	17	[-3.16, 0.3]
9	21	[-2.5, 0.51]
10	25	[-2, 0.67]

Table 3.1: Measurement times and corresponding interval data

The error bounded identification for this data-set is defined as the set of parameters  $p_1$  and  $p_2$  that create curves from equation 3.17 that cross all the vertical bars in Figure 3.4, which represent the experimental values. The set of solutions from the error-bounded identification then represents all possible model outcomes that are compliant with the error estimation considered. The solution set in the parameter space is shown in Figure 3.5. The thinner boxes represent the boundary of the optimization algorithm where the resolution threshold was met. All the boxes within the boundary are guaranteed solutions to the problem, and the boxes out of the boundary are guaranteed not to be solutions.

Even though error-bounded estimation is very extended within the scientific community, it depends heavily on the estimation of the error assumed in the measurements. In the context of diabetes monitoring, errors associated to continuous glucose monitoring are often too large to be useful [55], given the complexity of the models associated. In the hospital environment is possible to obtain much more reliable measurements in a continuous monitoring, with measurement errors smaller than the 2% [60], which can always be neglected for identification.

Error-bounded estimation is very much focused in the search of feasibility of a group of parameters with the uncertainty of the measure, although it can also consider

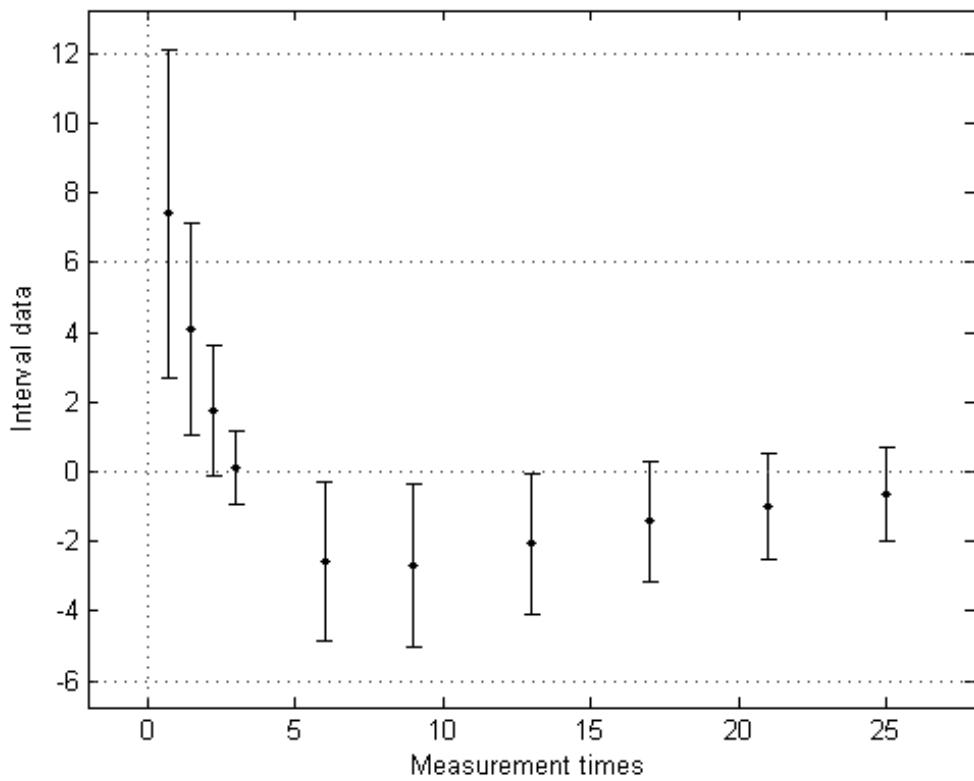


Figure 3.4: Acceptable data measurements including data uncertainty displayed as vertical bars.

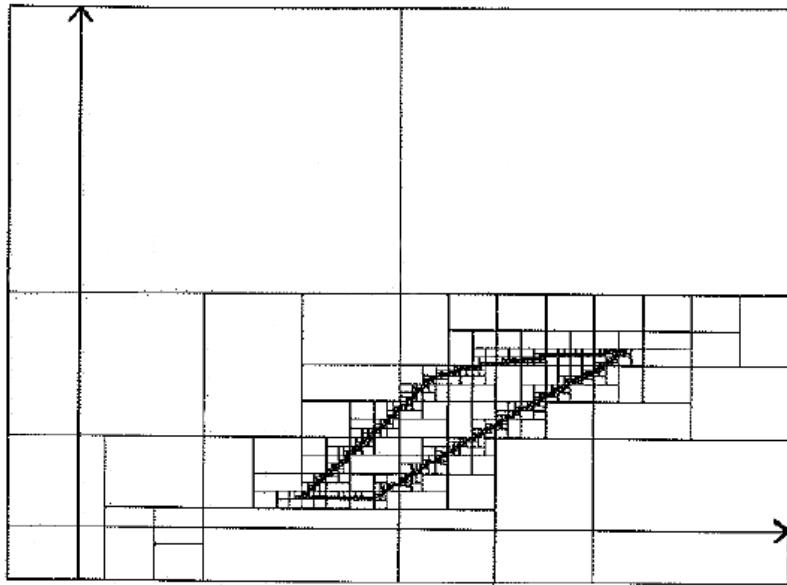


Figure 3.5: Parameter space for the parameters  $p_1$  and  $p_2$ . The frame corresponds to the search domain  $[-0.1, 1]^2$

parametric variability. From control-oriented identification, finding sets of parameters that match the estimated uncertainty in the output is not as important as finding all possible values of the parametric space the model varies within in order to tune a matching controller.

### 3.3 Optimization and software

Identification and optimization is usually a computation intensive task. Global non-convex optimization is especially demanding in computation requirements, and it has been established that identification on diabetic patients and models relies on global optimum solutions. There is a wide variety of optimization methods available in literature, but in the following lines we will quickly review the algorithms to be used in this thesis.

All data analysis and computations were done in the Matlab environment, release 2012a (Mathworks, Natick, MA).

### **3.3.1 Scatter search for Matlab**

Scatter search for Matlab (SSM from now on) is a global optimizer based on statistical principles and geometrical analysis of the parameter space. This search algorithm has already been used in the artificial pancreas environment with the objective of patient identification, by Cesar Palerm in Santa Barbara [62].

SSM optimizer is a project of the CSIC (Centro Superior de Investigaciones Cientificas) and the university of Vigo. It is a global optimizer, and as such it is easily comparable with genetic algorithms. However, SSM does not use codification of the population as genetic algorithms do, although it does work by spawning new generations (offspring) of the function optimum by combining the properties of the previous (parents) population. SSM does not generate random “mutations” on the population, as opposed to genetic algorithms, although it renews the existing individuals by adding new random samples to the new generations of optimal solutions in each algorithm iteration. For the details of the inner working of the algorithm and better understanding of the searcher check out Julio Banga’s group papers in 2006 [71] and 2007 [21].

This optimizer has the advantage of using local solvers to refine the search when it seems to have found some optimum solution. The local solver to be used can be chosen from a list available in the SSM’s documentation. The local solver chosen was the *fmincon* solver, which is implemented in Matlab’s Optimization Toolbox, along with many other optimizers and aid tools for solving any optimization problem in the Matlab environment. This local solver fits into the group of solvers denominated as “Nonlinear programming solvers”, or “Constrained nonlinear optimizers”. This sort of algorithms are deterministic solvers that use first and second derivatives of the objective function, along with some heuristics to cope with the various problems that deterministic local searchers have. Abundant information about this kind of algorithms can be found in Coleman and Li paper of 1996 [15], Powell’s conference in 1978 [66], or for a more general reference see Bazaraa’s book [3].

### **3.3.2 Covariance Matrix Adaptive Estimation**

A very well known global optimization method within the scientific community is the algorithm of estimation based on adaptations of the matrix of covariance of the sample population (CMAES). CMAES performs very fast optimizations on a single objective even for large parameters spaces. The optimization performed by CMAES is based on

the update of generations of sampling individuals, in a similar way as an evolutionary algorithm, although the update process and the randomly generated individuals are handled differently.

CMAES uses statistical properties of the populations and updates them in each iteration based on the characteristics of the population and the explored parameter space. The adaptation objective in on the covariance matrix adaptation moves the sampling population in order to better fit it to the contour lines of the objective function being minimized. It is assumed that the optimal covariance matrix equals the inverse of the Hessian matrix, although this is only strictly true in convex-quadratic functions. For general optimization functions, the CMAES methodology aims to approximate the inverse of the equivalent of the Hessian matrix for the objective function properties. For further explanation on the CMAES rational the reader is referred to the works of Hansen in [28] and [29].

Optimizations using interval parameters may consist of two independent parameter sets, upper and lower bounds, where obviously all the lower bounds must be smaller than their respective upper bounds. These greater-or-equal restrictions introduce further constraints on the optimization index. Unfortunately, the build released by Hansen *et al.* [29] does not implicitly consider restrictions neither in the outputs nor in the inputs, so they have to be integrated in the cost index with a penalty method.

### 3.3.3 $\epsilon$ - MOGA Evolutionary Algorithm

Classic parametric identification results in a single “optimal” point in the parameter space, being insufficient for the case of a time-varying model based on poor prediction capabilities. Many problems in engineering can be translated to multiobjective optimization problem, including the diabetic patient identification problem. In the case of identification in presence of uncertainty using interval models, a problem of optimization of two objectives can be suggested: minimization of glucose interval width and minimization of the fitting error. This problem will be explained in detail in the following thesis, and multiobjective optimization algorithms are used to find it’s solution. A brief explanation of these kind of algorithms follows, along with the details of one particular methodology used in this work.

Multiobjective identification is the process of simultaneously optimizing two or more conflicting objectives subject to certain constraints. The solution of this kind of problems

is not a single optimal point in the objectives search space, but a family of solutions called a Pareto Front (PF). PF present advantages over single objective optimization techniques:

- Provide the designer with the possibility of a better selection of the final solution, presenting a wide variety of possibilities that in many cases include the single objective solution of an analog problem.
- Are representative of the whole space of design variables, all of them optimal.
- Require of no *tradeoff* parameters to be tuned in the algorithm.

PF are created on the basis that each individual in the family is non-dominated by the other individuals. An individual is dominated by other in the population if its evaluation in *both* of the optimization objectives is worse. Therefore, an individual is considered non-dominated and included in the PF by the algorithm if it is best solutions for both objectives in a particular region of the objective space i.e. it cannot be replaced by other point in the objective space for improving an objective, without worsening another one.

Evolutionary algorithms are popular solvers for multiobjective problems because of Pareto front groups being susceptible to evolution rules, and both the solver and the problem being population based methods. The  $\epsilon$ -MOGA evolutionary algorithm developed by [31] takes the concept of dominance an step further with the notion of  $\epsilon$ -dominance.  $\epsilon$ -dominance is based on the domination concept explained before, but it adds a new constraint to the PF individuals: each individual must be isolated within a box of previously defined magnitude.

The concept of  $\epsilon$ -dominance is clearly displayed in Figure 3.6, where the objective space is overlapping a grid of equally dimensioned boxes. Dominance is then applied to the box in which the Pareto individual is placed, and not just to the point in the objective space. This box-dominance is the  $\epsilon$ -dominance of the individual of the PF in that box.

The  $\epsilon$ -MOGA algorithm is able to characterize all kind of PF, including non-convex, non-linear discontinuous search spaces, which is the case for diabetic patient models optimization. The algorithm converges faster than similar evolutionary algorithms, and supposes a smaller computational burden. The main achievement of this algorithm is the fact that the PF obtained from  $\epsilon$ -dominance are well-distributed in the objective space independently of the problem, which is not the case for any other multiobjective optimization algorithm. A well distributed PF is of great help in the interpretation of the

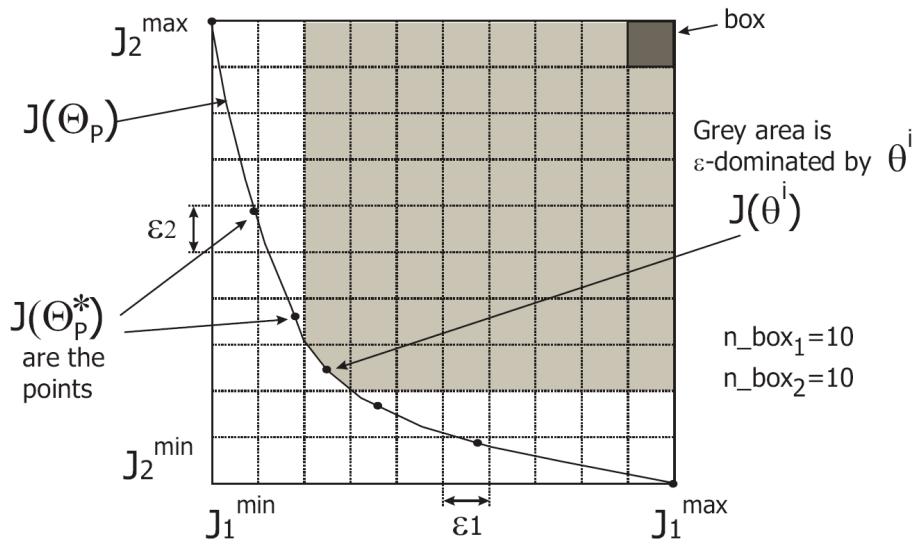


Figure 3.6: Display of the concept of  $\epsilon$ -dominance.  $J(\Theta_p)$  represents the PF interpolation and  $J(\Theta_p^*)$  the actual PF individuals.  $J_1$  and  $J_2$  are the objective variables, and  $\epsilon_1$  and  $\epsilon_2$  are the box widths.

results, and can be of utility when choosing an individual out of the PF with a decision making methodology.

# Chapter 4

## Conclusions

Identification studies for type 1 diabetic patients are rarely found in literature. In this part of the thesis I reviewed the most significative cases of patient individualization, but unfortunately not one of them displays satisfactory prediction capabilities. In summary, individualization of type 1 diabetic models is still work in progress and special effort must be done in order to develop new strategies for the achievement of a prediction method of the glucose of a diabetic patient. A summary of the reviewed studies can be found in table 4.1.

Reference	Model	Achievements	Drawbacks
[77]	Data based	Real life measurements Up to 2 hour prediction	Only 1 patient
[11]	Data based	Use of CGM	Only 1 patient 30 minutes prediction
[73]	Data based	Use of multiple sensors Fit $R = 0.7$	Only 1 patient No insulin disturbance
[27]	Data based	7 patients	Poor predictions
[10]	Data based	Experimental validation	Poor predictions
[62]	First principles	Good data fitting Physiology based models	Poor validation
[25]	Data based	Comparison against ZOH	No improvement over ZOH
[44]	Data based	Interval usage	No validation presented

Table 4.1: Summary of identification studies

Overlooking the studies found in literature it is immediately clear that no studies with many patients have been conducted, and multiple patient studies present either no validation of the identification or a poor validation result. Only one study is shown using physiology based models, and even though validation data is presented, the results are not satisfactory. In that study, no consideration of uncertainty was performed. In this thesis a full cross-validation study on multiple patients is presented using first principles models. By adding uncertainty on the identification process I expect validation results to be improved in the whole population of patients examined.

Lack of successful prediction and identification studies published in literature suggest that all of the pilot individualization studies performed by the scientific community have been unsuccessful. This thought hints that individualization studies may be much more challenging for the diabetes paradigm than other classic engineering problems. These facts did not dissuade the author to advance as much as possible towards the solution of the identification problem, and in fact, new identification strategies are explored in this thesis that we believe strongly encourage the scientific community to push the limits of patient individualization forward.

The work developed in this thesis is based upon the pilot identification study presented in the master's thesis by the same author [51]. In the master's thesis, an extensive review of models for diabetic patients was done, displaying all equations and characteristics of almost all the models found in literature. A brief introduction was also performed on practical identifiability of some of the models, as well as some results on identification on experimental patient's data. No uncertainty was considered in the identification in that case, resulting many times in poor prediction capabilities of the models individualized.

The results of the master's thesis led to negative conclusions, so no publications were associated. Following this reasoning, and given that few identification studies in literature show validation data, and being patient individualization such an important matter, we can only assume that many groups have already tried to perform pilot identification studies with few success. We would like to encourage readers to scientifically acknowledge that non-successful studies, should be supported by a solid theoretical methodology, are as important for publication as successful experiments. I believe enormous efforts and resources are being spent in repeating experiments with negative results, not only in the diabetes context, and especially within the smaller research groups, because of negative results publication avoidance.

Model accuracy is an open issue when dealing with glucose predictions. No proposed

model of those reviewed in this thesis has shown exceeding performance over the others, and each one of them has been designed following a different set of directions. Models based on physiology are the most extended models in control and identification in diabetes, but they all lack on validation results. Experiment desing based on models can be used for a better use of the models proposed.

Almost no focus has been given to uncertainty consideration in the identification context in literature, despite the public knowledge of enormous variations in some of the physiology parameters, and large errors registered in continuous monitoring studies. I believe uncertainty treatment is key for better prediction of glucose concentration of diabetic patients, even using current models based in phsyiology. Interval models appear as the perfect choice for these type of problem, and will be used in this thesis to develope reliable identification methodologies. Of course, as modeling in diabetes advances, predictions can be more and more accurate, with less uncertainty from unmodeled dynamics to be considered by the intervals, but we consider this to be work to be done in parallel to the identification methodologies, which are the focus of this thesis.

Interval analysis and error-bounded estimation is a well stablished methodology, but have never been used in the diabetes context. This is most likely a consequence of the fact that estimating error incidence in glocuse monitoring is usually very difficult and resulting errors are too large to be useful. Sets of parameters that result from bounding all the possible errors in the glucose concentration space are surely too large to be used in prediction studies and controller design. An interval approach more focused on the modeling of variability is presented in this thesis, where the focus of the interval model is not to simulate all feasible responses to a single simulation study but to bound several similar experiments on the same patient. With this repetition of days on the experiment, patient variability is expected to be present and interval models are expected to capture it. Coping with error in the measurement is handled by acknowledging a compromise between data fitting and measurement error. This type of interval bounding also focus on the development of robust controllers because it provides not only patient characterization but also relative uncertainty measures for every different patient, which is closely related to robust controller design parameters.

A quick visual outline of this thesis contribution can be seen in Figure 4.1. Virtual patient identification is needed for proving the benefits of the proposed identification methodology. The first contributions are comprised in the second part of this thesis, and it focuses on the improvement of the conditions of identification for both the identification of virtual patients and in the experimental condition. Experimental design is carried out using multiple models in order to obtain rich quality data from the

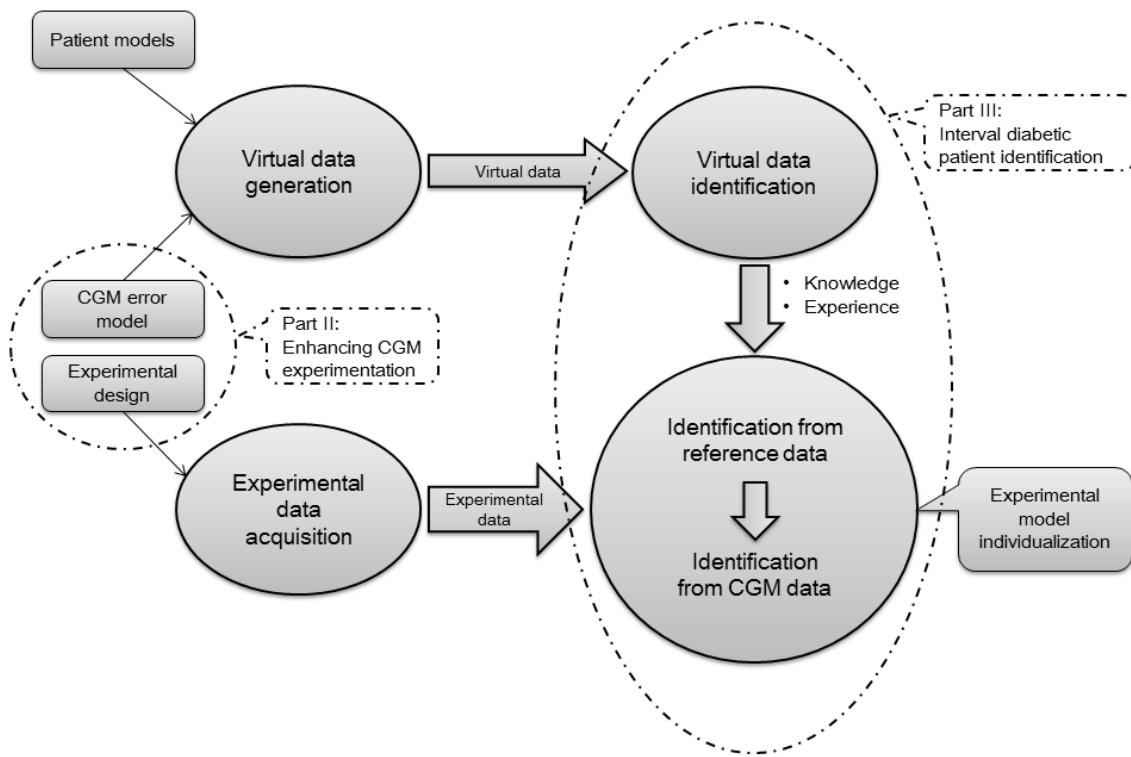


Figure 4.1: Workflow schematics for this thesis. The conceptual blocks that correspond to the subsequent parts of this thesis are highlighted in dashed circles.

diabetic patients. Diabetic patients hospital data is then gathered and analyzed for CGM error modeling, in order to obtain a much more reliable virtual patient simulation, including daily monitoring using a simulated CGM. The last part of this thesis is the core identification study, on both the virtual and experimental patient's cohorts, using the proposed methodology adequately in each case.

## **Part II**

# **Enhancing CGM Experiments**

# Chapter 5

## Optimal Experiment Design

Optimal experiment design requires of the designer to determine the experimental parameters to be used for optimization of identifiability. The diabetic patient model's inputs are the variables that can be adjusted for the adequation of the experiment. For obtaining an optimal experiment, the best combination of the inputs regarding at the identifiability of the model is to be found.

Another variable usually considered in the optimization of the inputs is the sampling period of the measurements. In the experiments to be performed in this thesis, the measurements are considered to be sampled by a CGM, so the measurements are taken with the sampling period of the monitor, and the number of samples linearly increases with the experiment's length. The design problem is then transformed into how long the experiment should last. Usually, identifiability of the model increases with the number of samples because the experiment's data gets richer, so the boundary of the experiment time is not of identifiability issues but of the patient's comfort. After medical consultation and examination of the resources for the experiments, it was decided that the length of the experiment was to be fixed at five hours after a meal.

The most common inputs of a complete glucoregulatory model are two, the meals and the insulin infusion (supposing insulin pump treatment). The complexity of the meal absorption and the absence of satisfactory models has historically made meals to be considered as a disturbance, but in this work it is treated as an input to be designed. Insulin postprandial infusion is considered as an input as well. Both inputs can be parameterized as:

- Parameterization of the meal system input:
  1. Meal size
  2. Meal composition
- Parameterization of the insulin system input:
  1. Insulin bolus size
  2. Insulin bolus time (relative to meal time)
  3. Basal infusion rate

Regarding meal composition, the published parameters in literature for the digestive model analyzed in here, the UVA gastrointestinal model, were identified under the specific conditions listed below:

- Caloric input: 10 kcal/kg
- Carbohydrate: 45%
- Protein: 15%
- Fat: 40%
- Glucose estimation:  $90 \pm 5$  g

Glucose estimation represents the size of the meal. The rest of the parameters determine the gastric emptying profile, and the identification performed in [17] produced the parameter values shown in Table 1.5 based on the specified meal composition. Unfortunately, no identifications on different meal compositions have been performed on this model. Therefore, if optimum performance of the model is desired, experiments with this model are to be performed with similar composition to that shown in [17]. The size of the meal is one of the parameters for the experiment design.

Regarding at the insulin input, the parameters chosen for optimization of the experiment design are much more complex than the those of the meal ingestion, which is only considered as an impulse. The insulin pump treatment can have any shape imaginable respecting the pump limitations. Designing an absolutely free insulin profile, i.e. a non-parametric input, would be too expensive in computational terms (for more

information see [81]), so a parametric input is chosen. The classic profile of basal-bolus infusion is preserved in this experiment design for simplification purposes and to increase the patient's compliance with the experiment.

Once decided that a traditional treatment is going to be followed, the only parameters to be chosen for the experiment design are the amount of insulin to be given as a bolus, the basal insulin level, and the instant the bolus is administered. In summary, the experiment conditions to be tuned in the optimizaion are:

- $\tau$  - Delay of the insulin bolus with respect the meal time, in minutes. A positive delay means that the insulin dose is given after the meal and vice versa. The boundaries for the optimization problem are set to -60 and 300 minutes. Maximum advancement of the insulin bolus is defined so as to prevent hypoglycemic events.
- Meal - The size of the meal in grams of carbohydrates. The boundaries for the optimization problem are 0 and 100 grams of carbohydrates.
- Bolus - The amount of bolus insulin units to be given. The boundaries for the optimization problem are 0 and 10 insulin units. These boundaries were chosen using an insulin-to-carbohydrate ratio of 1:10 as population mean.

Only one meal per day is considered in order to avoid influences from the circadian variation of insulin sensitivity. The vector  $\Xi$  (as seen in equation 3.11) is a vector of size  $3 \cdot n_d$ , where  $n_d$  is the number of daysf monitoring. This vector is the output of the experiment design.

Given that the problem is highly non-linear (due to non-linearities of the index and the model) SSM global optimizer was used to obtain the solutions, as introduced in the section 3.3.1. The constraints for the optimization problem are the restrictions of the blood glucose in order to keep the patient under safe health conditions. The simulation was forced to start from equilibrium in a blood glucose level of 100 mg/dl. The maximum level of glucose (hyperglycemic bound) was set to 250 mg/dl and the minimum level (hypoglycemic bound) was set to 70 mg/dl. These boundaries are respected at all times during the optimization of the experiment design.

Before the actual experiment desgin, a selection of the most relevant parameters is performed. A very simple analysis of identifiability is carried out on each of the models tested in order to discard non-relevant parameters that may induce optimality problems in the minimization of the experiment design. Repeated iterations of sensitivity analysis

are done over the full model parameter set, iteratively removing any parameter that produces either singularity of the FIM or that presents a coefficient of variation larger than 30%.

In the following lines results for experiment design in diabetes for two different models are shown. Despite the fact that Bergman's model has been proven unreliable for simulating some glucose dynamics, it is desirable to execute the experiment design methodology varying the models used in order to make the experiments as model-independent as possible. Therefore the first experiment design presented is calculated for this model. After that, experiments designed for the modified Panunzi model are displayed. Qualitative information can be extracted from every model's designed experiment. Common conclusions will be synthesized after the experiment design with both models considered and a feasible, patient-friendly protocol will be developed

## 5.1 Experiments designed with Bergman's model

Initially, Bergman's endogenous model requires for the identification of three parameters, but two more are added to this study: distribution volume and basal glucose. Glucose distribution volume affects the glucose concentration with respect to the glucose inputs of the system throughout the postprandial period, and basal glucose states for the initial state of the glucose simulation, which for Bergman's model (equation 1.30) also affects the dynamics of the system. To complete the simulation of a diabetic patient, UVA's gastrointestinal model is implemented along with Cambridge's insulin absorption model. UVA's gastrointestinal model is one of the few models presented in literature with identification on mixed meals instead of glucose solution (even though the mixed meal used is far from being a classic meal), and that is the reason to choose it for the virtual patient's simulation. The Cambridge insulin absorption model was chosen based on the analysis of Willinska *et al.* where it was compared against 11 literature models [83] and showed excellent prediction capabilities. Bergman's endogenous model combined with UVA's glucose absorption model and Cambridge's insulin subcutaneous model sums up a total of 17 parameters, described in Table 5.1.

Considering only blood glucose as the measurable output of the system, identifiability of all parameters is not expected. A three days experiment as the identification period was considered. A sensitivity analysis previous to the experiment design was performed, and the results are shown in Table 5.2. Only three iterations were needed to obtain a feasible

Model	Number	Parameter	Meaning
Endogenous	1	$p_1$	Glucose efficiency
	2	$p_2$	Insulin action rate of disappearance
	3	$p_3$	Insulin action rate of appearance
	4	$Vol_g$	Distribution volume of glucose
	5	$G_b$	Basal glucose
Gastric	6	$K_{abs}$	Glucose absorption rate in the gut
	7	$K_{max}$	Maximum gastric emptying
	8	$K_{min}$	Minimum gastric emptying
	9	$b$	Involved in the gastric emptying
	10	$d$	Involved in the gastric emptying
Insulin	11	$V_i$	Distribution volume of insulin
	12	$k$	Proportion of insulin in the slow channel
	13	$k_{a1}$	Transfer rate in the slow channel
	14	$k_{a2}$	Transfer rate in the fast channel
	15	$k_e$	Insulin elimination
	16	$V_{MAX,LD}$	Michaelis-Menten parameter
	17	$k_{M,LD}$	Michaelis-Menten parameter

Table 5.1: Parameters intergrated in Bergman's model

set of parameters for identification on the Bergman model, but the algorithm initially discarded 8 parameters due to non-intertability of the FIM, which is closely related to structural non-identifiability. Indeed, the structural combination of Bergman model with the selected input models rendered many parameters difficult to identify just from glucose concentration data. Should richer data be available, such as plasma insulin data or tracer data of glucose flux from the gut, the sensitivity analysis would change.

Number	Parameter	Iteration		
		1	2	3
1	$p_1$	-	-	-
2	$p_2$	-	-	-
3	$p_3$	-	-	-
4	$Vol_g$	0,25	0,14	0,13
5	$G_b$	0,12	0,08	0,07
6	$K_{abs}$	-	-	-
7	$K_{max}$	0,25	0,24	0,23
8	$K_{min}$	0,42	-	-
9	$b$	0,16	0,12	0,04
10	$d$	0,20	0,18	0,09
11	$V_i$	-	-	-
12	$k$	0,34	0,33	-
13	$k_{a1}$	0,31	0,22	0,09
14	$k_{a2}$	-	-	-
15	$k_e$	0,22	0,14	0,12
16	$V_{MAX,LD}$	-	-	-
17	$k_{M,LD}$	-	-	-

Table 5.2: Bergman's model parameter coefficient of variation (CV) for a three meals simulation. Each column represents an iteration in the sensitivity analysis. Parameters with null CV have been fixed to nominal values due to non-identifiability. Parameters fixed at iteration 1 are cause of structural non-identifiability

The final number of parameters identifiable, out of the initial 17, is 7. Only two of those parameters are part of the endogenous model, while three characterize the gastrointestinal model, and the other two are part of the insulin model, as can be seen in Figure 5.1.

Those 7 parameters are used to optimally express the identifiability of the model. The experiments design next based on Bergman's model expresses the models identifiability

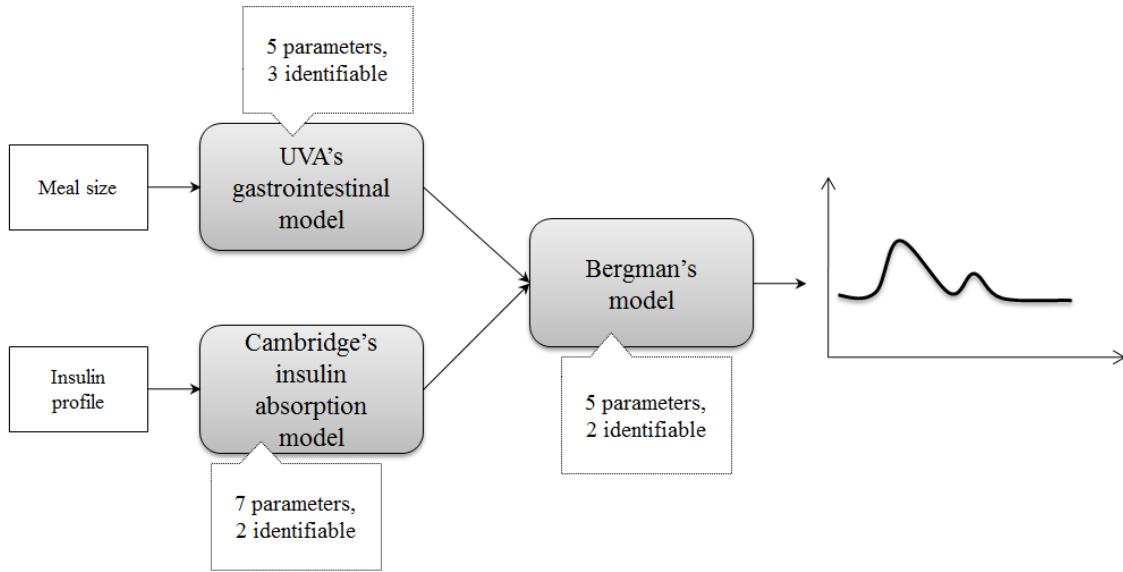


Figure 5.1: Optimal set of parameters for Bergman's model combined with UVA's gastrointestinal model and Cambridge insulin absorption

in the optimization cost index by using the optimal set of parameters. The range of days considered for experimentation is from one to four days. At the end of the experiment design for all the models under study, the best option was chosen from within of all the experiments designed, including the different options of parameter sets, models, and length of the experiments.

In the next lines, experiments designed comprising one, two, three and four days of monitoring are displayed.

The results of an experiment designed for *one* day of monitoring consisting on the five following hours to a meal are:  $\tau = -31, 59$  minutes, Meal = 100 gr CHO and Bolus = 10 IU. In the case of only being able of monitoring one day, the optimum experiment to be followed requires the diabetic patient to advance the bolus approximately half hour, and then eat a large meal (100 grams CHO). It is worth remembering that 100 grams of carbohydrates is the top value given to the optimizer for the experiment design. Also, 10 is the top boundary for the bolus variable in the optimization, and it is the corresponding value of insulin for an standard treatment of insulin boluses for a 100 grams of carbohydrates meal. This means that, in addition to maximizing identifiability of the model, the experiment design keeps the patient under standard glycaemic control.

This is a direct response to the output restrictions to the simulation model that force the patient's glucose to be within healthy range. The glucose profile associated to the experiment described can be seen in Figure 5.2.

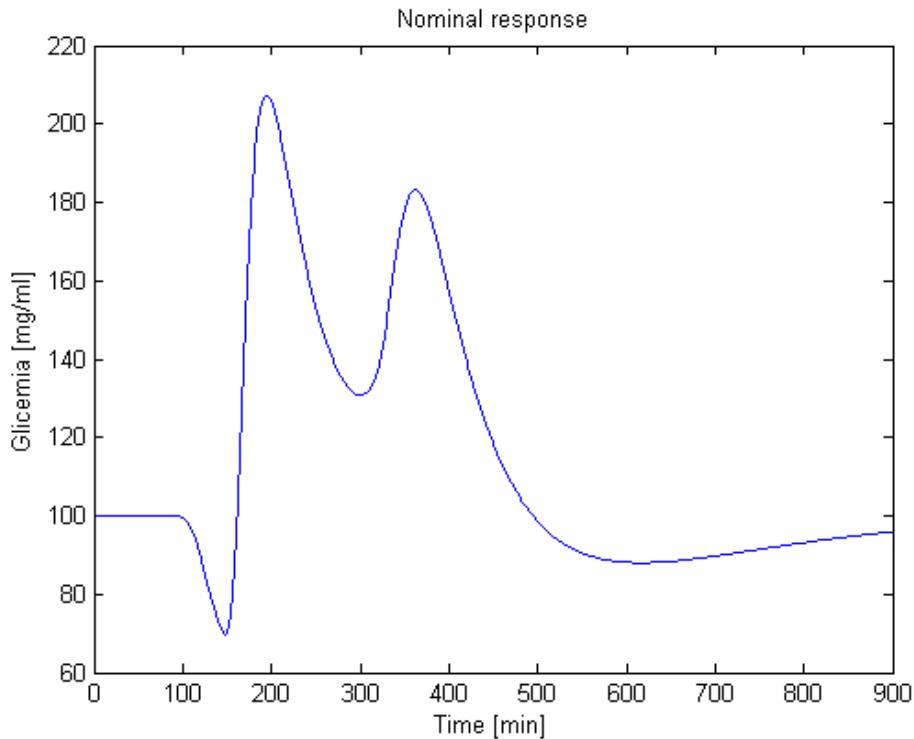


Figure 5.2: Bergman's model response to the proposed experiment for 1 day

The glucose profile shows a very quick response to the advanced insulin bolus, dropping quickly to a level of approximately 70 mg/dl. This level is the safety threshold indicated as a constraint to the output in the optimization algorithm in order to keep the patient under safe conditions. This simulation shows how the optimizer is exciting the output variable as much as possible, and always keeping the (virtual) person safe.

The conditions of the experiment designed for *two* days monitoring are:

- Day 1.  $\tau = 25, 23$  minutes, Meal = 86, 81 gr CHO and Bolus = 10 Insulin Units
- Day 2.  $\tau = -29, 58$  minutes, Meal = 100 gr CHO and Bolus = 10 Insulin Units

In this new experiment two different therapies are applied to the patient. The second day is a replicate of the conditions of the experiment design corresponding to only one day. In the first day of identification, a bolus of 10 insulin units is delayed approximately 25 minutes, for a meal of 86,81 grams of carbohydrates. In this case the proportion of insulin units to grams of carbohydrates of the meal is not maintained, thus, not a standard glycaemic control is performed on the patient (including the effect of the delay). The glucose profile for the designed two-day experiment is shown in Figure 5.3.

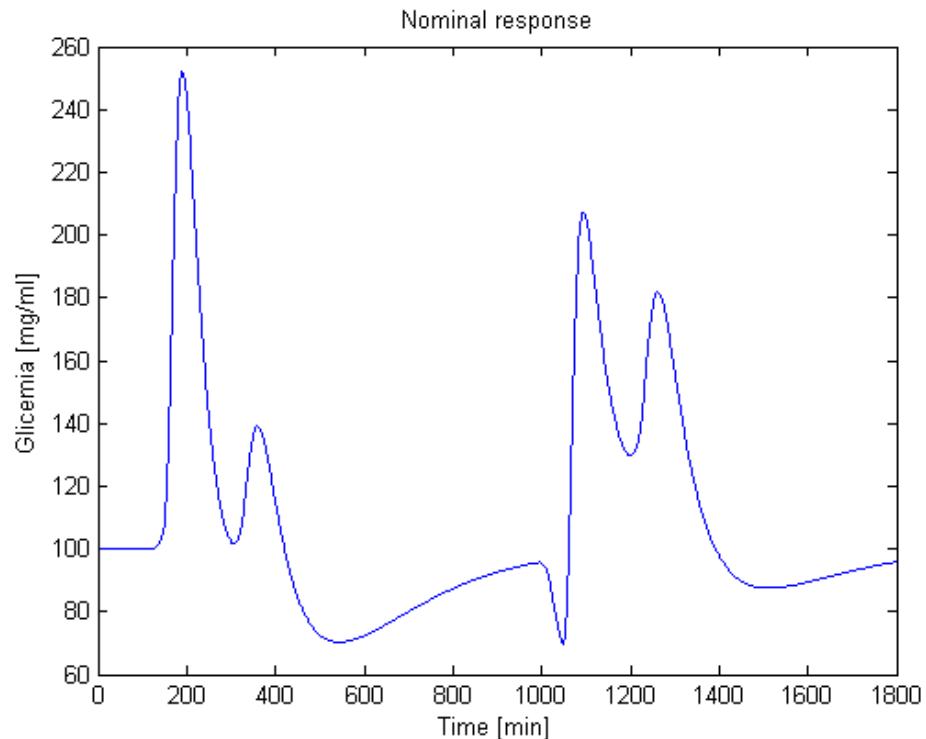


Figure 5.3: Bergman's model response to the proposed experiment for 2 days

Observing the first day response of the model to the proposed experiment, the fact of non-proportionality of insulin to carbohydrates makes more sense. The model moves from 250 mg/dl down to 70 mg/dl, which are the top and bottom safety thresholds in the experiment design optimization algorithm. The optimizer is exciting the measurable variable (blood glucose) through all the feasibility glucose range in order to maximize identifiability. Given that in the first day, due to the delay of the insulin bolus, there is no falling of the blood glucose to the lowest level, it has to drop to that level after the meal, that is why the insulin bolus is still relatively big with respect to the size of the meal.

Given those values, the maximum amount of carbohydrates for that meal is 86, 81.

It is interesting to understand why the two days differ in the relative bolus administration time. In the first day, no insulin excitation is applied in approximately 30 minutes from meal time and all the information extracted from that period is contributing to increase the identifiability of the gastrointestinal input model, the UVA model. In the other day, the opposite situation is happening: only the insulin subcutaneous model is experiencing excitation in the first period of the simulation, so information from glucose is affecting only the parameters of that model. The stronger the variation of the states of the model in that period, the greater the identifiability will be.

The conditions of the experiment designed for *three* days monitoring are:

- Day 1.  $\tau = 26$ , 43 minutes, Meal = 85, 6 gr CHO and Bolus = 10 Insulin Units
- Day 2.  $\tau = -30$ , 44 minutes, Meal = 100 gr CHO and Bolus = 10 Insulin Units
- Day 3.  $\tau = -30$ , 71 minutes, Meal = 100 gr CHO and Bolus = 10 Insulin Units

In this case, the advancing of the bolus is repeated two days. Given that the meals do not have to be consecutive, or that one day does not have any relation to the others, the order of the experiment days is not relevant for the result of the experiment. The glucose profile for the designed three-day experiment is shown in Figure 5.4

The second and third days are virtually the same experiment, and the first day displays again the extreme excitation of glucose between safety boundaries. It can be appreciated once again how there is a separation of dynamics between the different input models.

The conditions of the experiment designed for *four* days monitoring are:

- Day 1.  $\tau = 25$ , 44 minutes, Meal = 86, 81 gr CHO and Bolus = 10 Insulin Units
- Day 2.  $\tau = 14$ , 36 minutes, Meal = 99, 64 gr CHO and Bolus = 10 Insulin Units
- Day 3.  $\tau = -29$ , 92 minutes, Meal = 100 gr CHO and Bolus = 10 Insulin Units
- Day 4.  $\tau = -30$ , 88 minutes, Meal = 100 gr CHO and Bolus = 10 Insulin Units

In this final case the two last days are repeated again, displaying an advance of the insulin bolus, and a new delayed bolus profile appears, but in this case with a big meal.

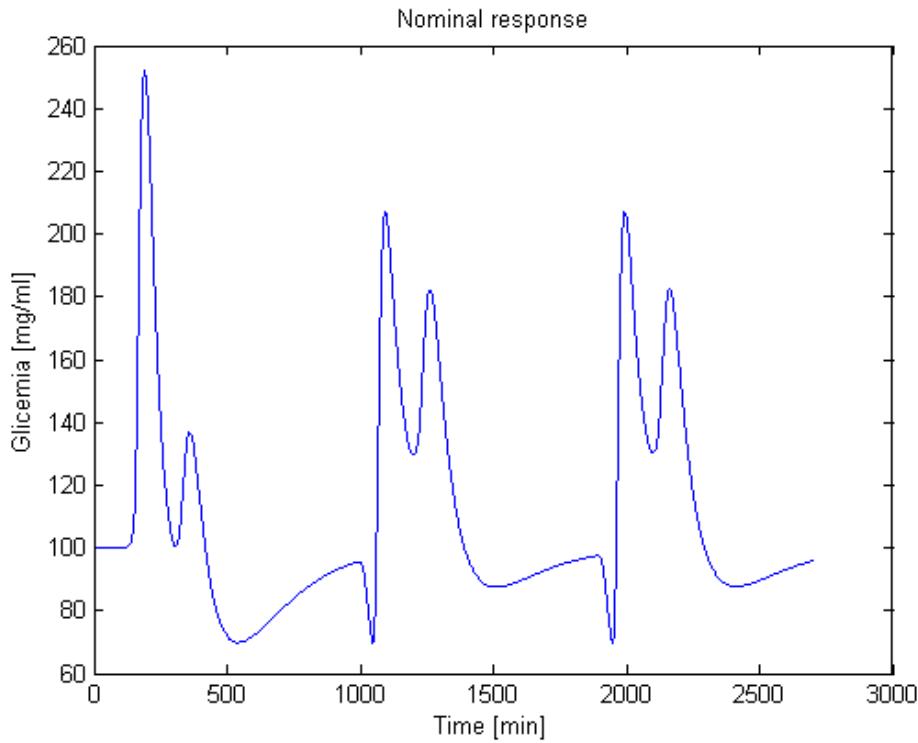


Figure 5.4: Bergman's model response to the proposed experiment for 3 days

The delay on the bolus is smaller so that the meal can be bigger, contributing to the improvement of identifiability of UVA's gastrointestinal model. The glucose profile for the designed four-days experiment is shown in Figure 5.5.

As it can be observed, there is not much difference in the first two days of identification because the difference in the meal size is relatively small.

The most relevant observation of this experiment design with the Bergman model is that separating dynamics of the input models is the key to improve the identification of the whole system. It is better for a model general identifiability to advance the bolus rather than having it delayed following the cases of one day experiment design (advancing bolus) and three-day experiment design (advancing bolus two days out of three). However, a better scenario comprises both situations, advancing and delaying the bolus. A minimum of a two-day experiment is then required for basic identification of the system.

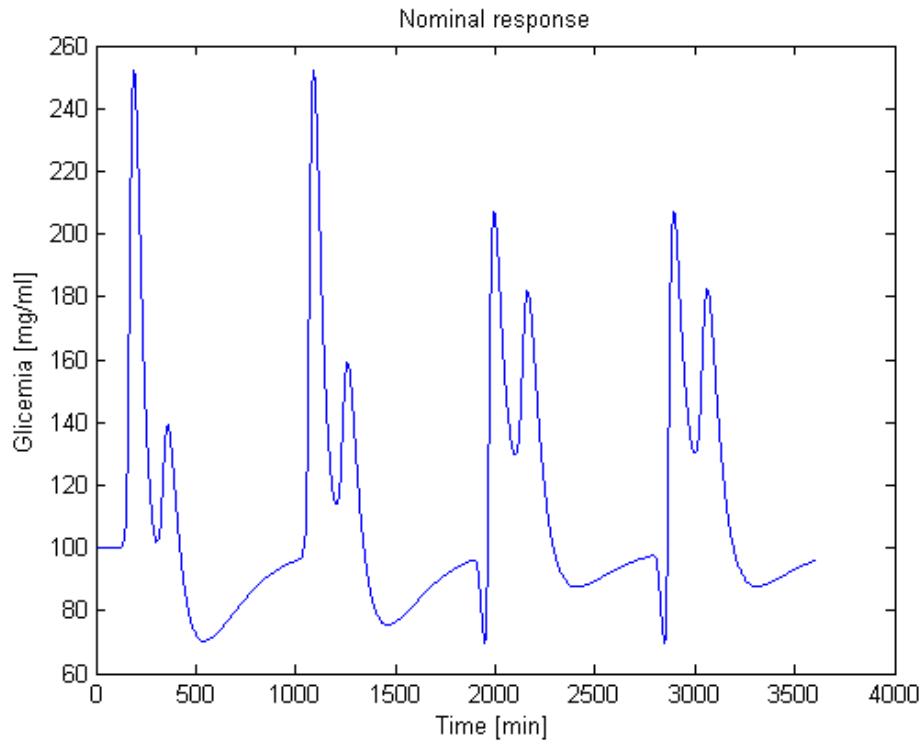


Figure 5.5: Bergman's model response to the proposed experiment for 4 days

Experiment's length is selected attending to: 1) identifiability conditions, 2) experiment parameters and 3) practical implementation aspects. It has been stated that a minimum of two days for monitoring seems reasonable for identification purposes due to the two different profiles of experiment. Naturally, as more data (more days of experiment) is available, the probability of a satisfactory identification increases. However, the practical implementation aspect limits the number of days in a experiment. In Figure 5.6 evolution of identifiability of each parameter of the optimal set in Bergman's model can be observed for all the cases exposed before and after the experiment design.

An abrupt decrease in the coefficient of variation of all parameters occurs between one and two days of monitoring, reinforcing the mentioned minimum of identifiability in two days, both for the experiments with and without optimal design. If more days are added to the experiment identifiability keeps increasing less rapidly. Very similar trends in the identifiability rise are observed for both the original experiments and for the optimal designed experiments. The identifiability levels, as measured by coefficient

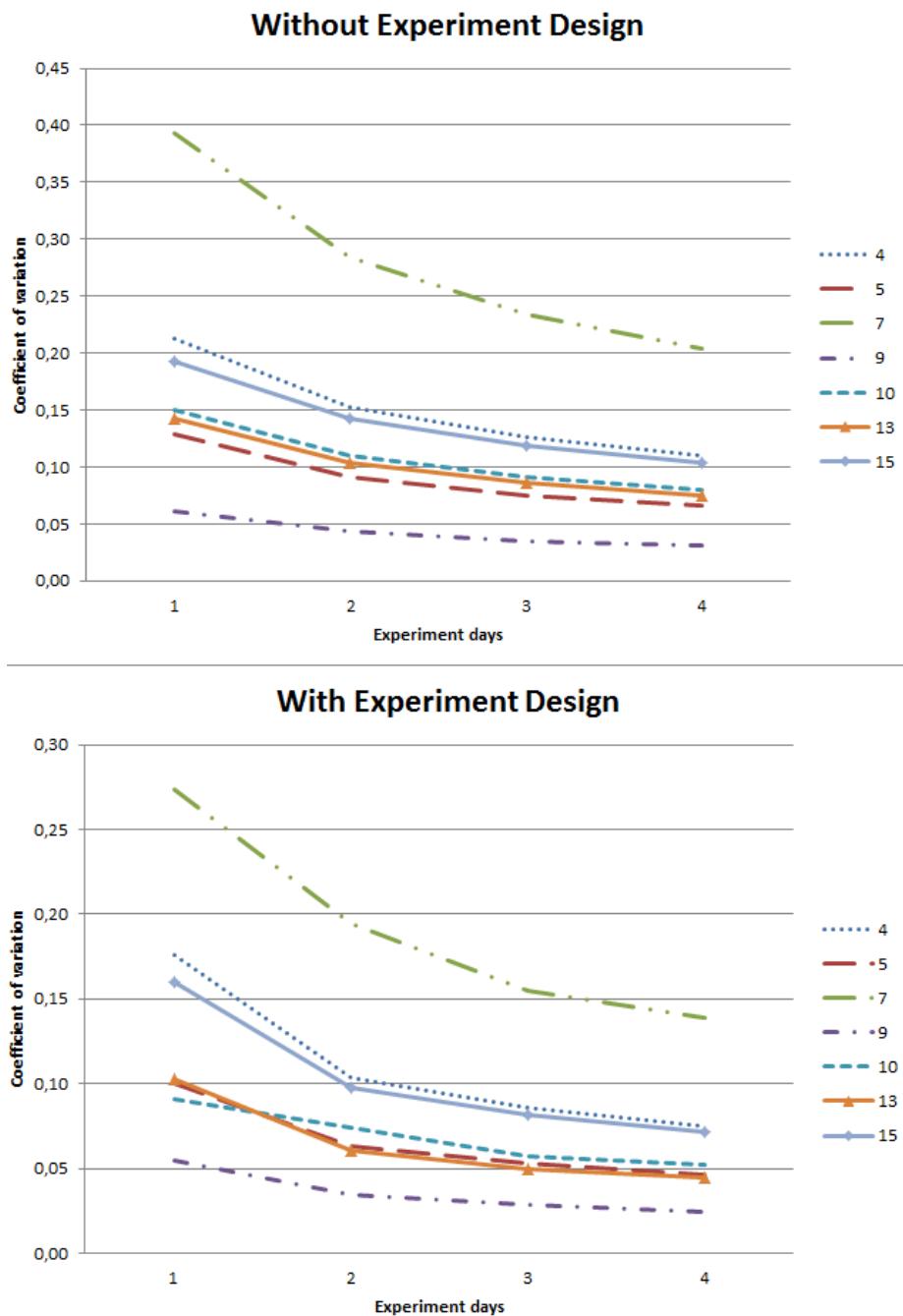


Figure 5.6: Coefficient of variation for Bergman's model's parameters. Top graph shows the evolution of the parameters' identifiability with experiment length. Bottom panel displays the identifiability of the same parameters applying the designed experiments. It is clearly shown how identifiability increases (CVs decrease) with experiment length

of variation of the parameters are much lower for any experiment length when choosing the optimal design for the experiment. A three-day experiment was chosen thinking on the limitations introduced by the clinical implementation of the experiment. It is desired to have as many identification days as validation days, which makes the real monitoring period twice as long as the experiment designed. Given that the average life period of a continuous glucose monitor is approximately one week, the three days experiment was chosen as monitoring duration, for maximization of identifiability. Results have to be compared designed experiments based on other models in order to make the experiments as model-independent as possible. Comparison between the coefficients of variation of a three-day identification experiment is displayed in Figure 5.7.

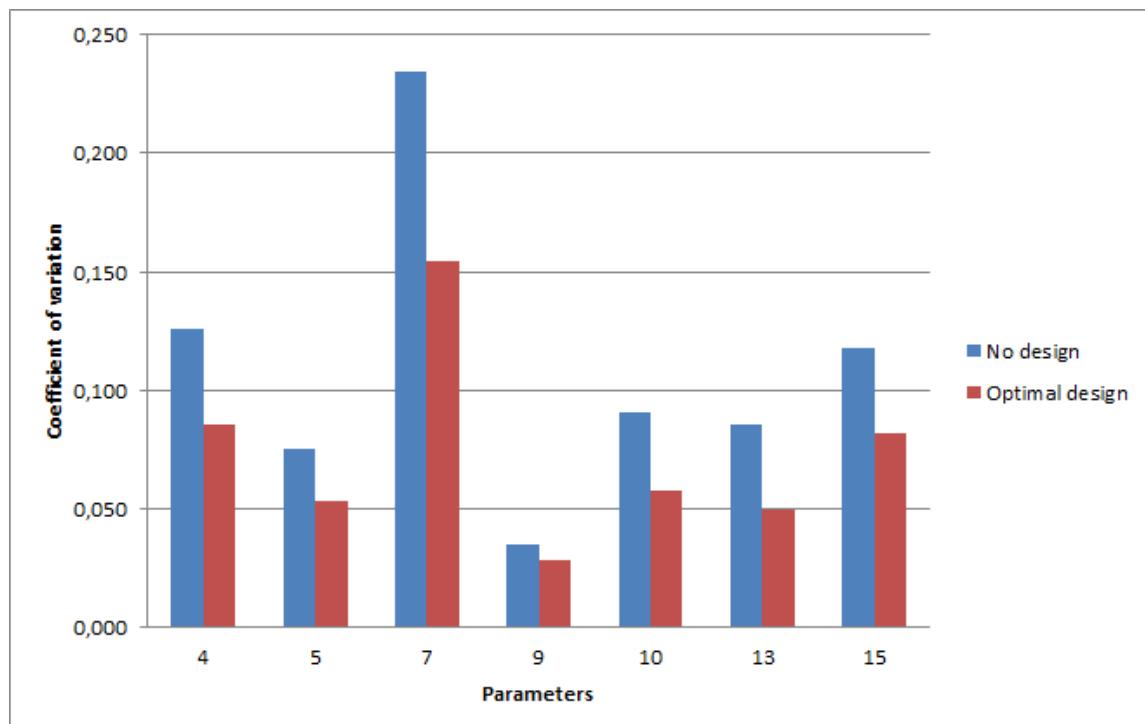


Figure 5.7: A comparison of identifiability for each parameter with and without the experiment design for the case of a three meals monitoring

A remarkable fall in the coefficient of variation is observed for all the parameters, and not just for the general identifiability of the model. In the original sensitivity analysis of Bergman's model all the parameters were selected to be identifiable, assuming identifiability of the model when all its parameters presented CVs below 30%. For Bergman's model the limitant parameter (highest CV) was parameter 7,  $K_{max}$ .

Using experiment design we reduce the variability of this parameter to half of the identifiability threshold, not sacrificing identifiability in any of the other parameters.

## 5.2 Experiments designed with modified Panunzi's model

Identifiability of the modified version of Panunzi's endogenous model, which was described in Section 1.4 will be analyzed next. Modified Panunzi's model and the two input models already used for Bergman's system comprise a total of 16 parameters, 4 of which correspond to the endogenous model. The 16 parameters are described in Table 5.3.

Model	Number	Parameter	Meaning
Endogenous	1	$K_{xgl}$	Insulin sensitivity
	2	$V_g$	Distribution volume of glucose
	3	$T_{gh}$	Hepatic balance
	4	$k_i$	Insulin action delay
Gastric	5	$K_{abs}$	Glucose absorption rate in the gut
	6	$K_{max}$	Maximum gastric emptying rate
	7	$K_{min}$	Minimum gastric emptying rate
	8	$b$	Involved in the gastric emptying
	9	$c$	Involved in the gastric emptying
Insulin	10	$V_i$	Distribution volume of insulin
	11	$k$	Proportion of insulin in the slow channel
	12	$k_{a1}$	Transfer rate in the slow channel
	13	$k_{a2}$	Transfer rate in the fast channel
	14	$k_e$	Insulin elimination
	15	$V_{MAX,LD}$	Michaelis-Menten parameter
	16	$k_{M,LD}$	Michaelis-Menten parameter

Table 5.3: Parameters to be identified in Modified Panunzi's model

Many of the parameters involved in the sensitivity analysis for Bergman's endogenous model are repeated for this case, but the numeration is different due to the different number of parameters in the endogenous model. The results of the identifiability study are shown in Table 5.4.

Number	Parameter	Iteration				
		1	2	3	4	5
1	$K_{xgl}$	1,14	0,98	0,55	0,53	0,31
2	$V_g$	0,46	0,43	0,26	0,24	0,23
3	$T_{gh}$	0,51	0,18	0,18	0,18	0,06
4	$k_i$	4,91	-	-	-	-
5	$K_{abs}$	1,60	1,12	0,67	-	-
6	$K_{max}$	0,78	0,66	0,53	0,24	0,24
7	$K_{min}$	0,87	0,68	0,33	0,33	0,30
8	$b$	0,41	0,20	0,18	0,17	0,13
9	$c$	0,66	0,51	0,34	0,20	0,18
10	$V$	-	-	-	-	-
11	$k$	0,91	0,81	0,36	0,29	0,27
12	$k_{a1}$	0,45	0,43	0,26	0,25	0,20
13	$k_{a2}$	3,77	2,57	-	-	-
14	$k_e$	0,88	0,71	0,49	0,49	0,30
15	$V_{MAX,LD}$	1,92	0,92	0,66	0,62	-
16	$k_{M,LD}$	-	-	-	-	-

Table 5.4: Modified Panunzi's model parameter coefficients of variation (CV) for a three meals identification

Following the same methodology than in the Bergman's model analysis, only two parameters were not identifiable because of FIM singularity. These two parameters (10 and 16) show in table 5.4 with their row in the first iteration blank. Parameter 16 is fixed in the identification because of its low sensitivity, and parameter 10 is fixed to its nominal value due to a very strong correlation (proportional relation) to parameter 1. Six iterations on the sensitivity analysis were needed for the modified Panunzi's model in order to obtain a feasible set of parameters for identification, to produce a 10 parameter set with good identifiability.

The set of parameters in the last column of table 5.4 is considered the optimum set of parameters for identification even though 3 of the parameters of that set present CVs in the acceptable limit of identifiability (30%). Those parameters ( $K_{xgl}$ ,  $K_{min}$  and  $k_e$ ) in the proposed limit of identifiability are considered of great importance for the identification of individual patients in the postprandial period.  $K_{xgl}$  represents the insulin sensitivity of the patient, a parameter that is known to present high inter and intra-patient variability, thus needed to be subject to identification.  $K_{min}$  stands for one of the parameters defining gastric emptying, also very variable between and within patients.  $k_e$  is the insulin

elimination rate, and strongly affects the insulin dynamics and overall sensitivity of the patient. All these parameters were sensitive to identification in diabetic patients, and were decided to be included in the optimal set, following the proposed methodology. The summary of identifiable parameters is explained in Figure 5.8.

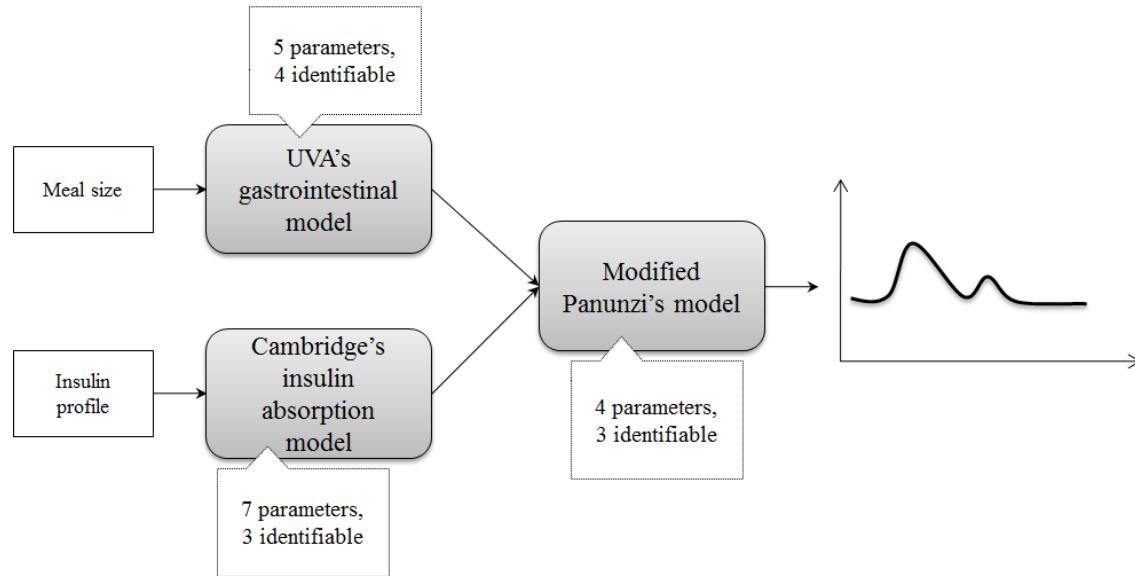


Figure 5.8: Optimal set of parameters for the modefied Panunzi's model combined with UVA's gastrointestinal model and Cambridge insulin absorption

These 10 parameters will be used in the experiment design for the calculation of the index of total identifiability of the model. All identifiable parameters are distributed amongst all the submodels of the modified Panunzi's system. The same approach of multiple day identifications used for the optimal experiment design of Bergman's model is used in here. In the next lines, experiments designed comprising one, two, three and four days of monitoring are displayed.

The results of an experiment designed for *one* day of monitoring consisting on the five following hours to a meal are:  $\tau = -17, 95$  minutes, Meal = 100 gr CHO and Bolus = 10 IU. The same profile as in the Bergman's model experiment design is observed in the modified Panunzi's experiment for the one-day monitoring. The insulin bolus is not as advanced as in Bergman's model design, but the meal and bolus size are again in the maximum value of the constrained optimization. Separation of dynamics of the insulin absorption system is the best option for the one-day monitoring scenario. The response

of the new model to the experiment designed is shown in Figure 5.9.

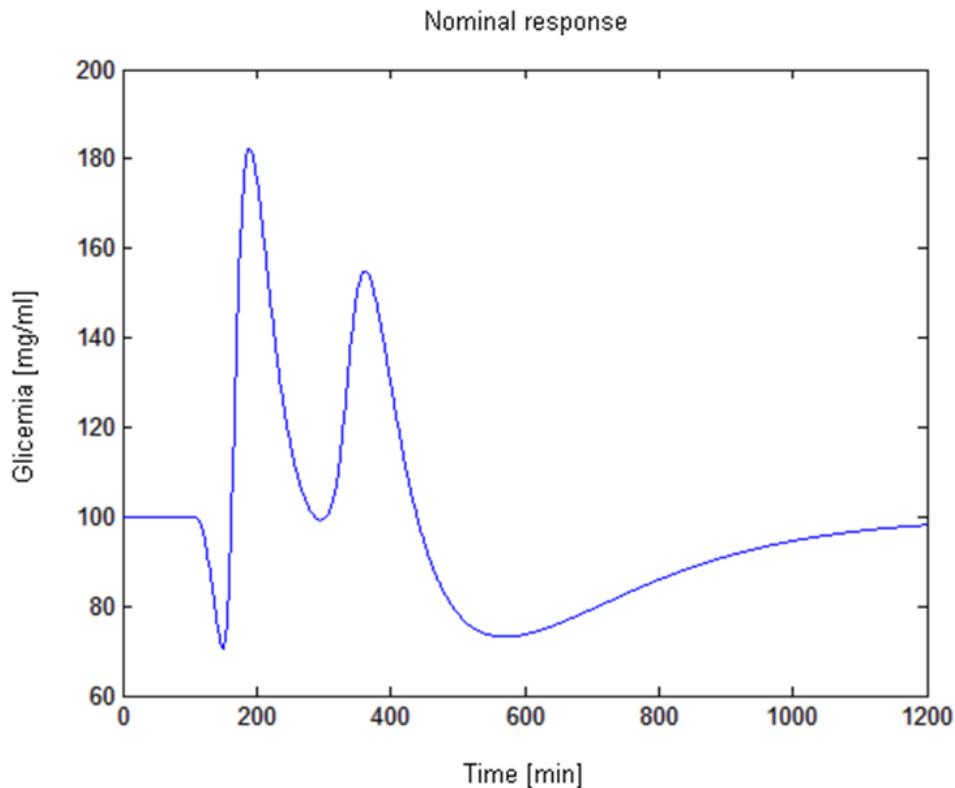


Figure 5.9: Modified Panunzi's model response to the proposed experiment for 1 day

Comparing modified Panunzi's model response in a single day experiment with the Bergman's model response (Figure 5.2), the glucose profile is virtually identical. This fact may lead to two conclusions: 1) Using simple endogenous models, the dynamics of the input models direct the behavior of glucose concentration; 2) the optimization of identifiability is not strongly dependent on the endogenous models used. Indeed the model's parameters are different in each model, but the signal from where information is being extracted is the same, and regarding that signal from a theoretical point of view, the optimal glucose profile in terms of containing information, must be very rich in information for any other model too. That is why it is expected of optimal experiment glucose outcomes to be similar for the modified Panunzi's model to the results for the other experiments already designed.

The conditions of experiment designed *two* day monitoring are:

- Day 1.  $\tau = 157$ , 80 minutes, Meal = 44, 25 gr CHO and Bolus = 3, 07 Insulin Units
- Day 2.  $\tau = -17$ , 92 minutes, Meal = 100 gr CHO and Bolus = 9, 99 Insulin Units

The second day in this experiment is the same as the single day experiment, in which the bolus is administered in advance to the meal time, thus only the insulin system is being excited in that period. In the first day, a new situation is proposed, similar to the case of Bergman's design in which the bolus was delayed, but in this case, the delay is much bigger. However, delaying the insulin bolus up to two hours and a half can be quite dangerous for the patient because of risks of hyperglycemia, that is why the meal size is much smaller for this large delay case, in order to keep the glucose profile within safety constraints. In Figure 5.10 the response of the model to the experiment proposed can be observed.

In the case of the first day the small amount of carbohydrates given to the patient (44,25 grams) has the objective of rising the blood glucose up to the superior limit of safety of the patient. The insulin bolus is given when the blood glucose gets to the maximum established, and the amount of insulin given is so that it will force the blood glucose down to the lower safety constraint. Again, the experiment design is exciting the glucose signal used for identification from one boundary to the other, in order to maximize the amount of information contained in the glucose signal. Bergman's model was probably not able rise so fast up to the limit of hyperglycemia due to the so called term of glucose efficiency (as explained in Section 1.4) that makes glucose to drop automatically when it rises above the stability point. As this term is not present in the modified Panunzi's model, glucose dynamics are much faster and permit stronger excitations from the inputs, which is more realistic.

The conditions of experiment designed *three* day monitoring are:

- Day 1.  $\tau = -17$ , 83 minutes, Meal = 100 gr CHO and Bolus = 10 Insulin Units
- Day 2.  $\tau = -18$ , 42 minutes, Meal = 100 gr CHO and Bolus = 10 Insulin Units
- Day 3.  $\tau = 165$ , 74 minutes, Meal = 44, 38 gr CHO and Bolus = 3, 01 Insulin Units

In this case the advanced bolus administration is given in two out of the three days of the experiment, while the other day a small meal is given, with the insulin bolus

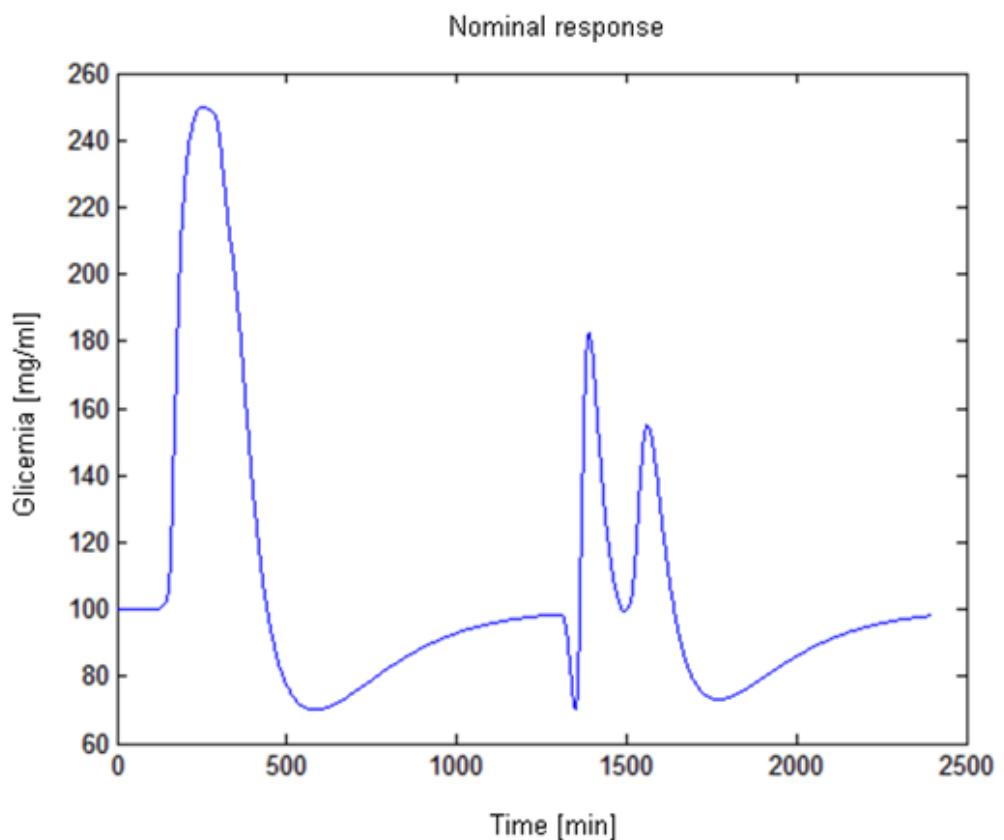


Figure 5.10: Modified Panunzi's model response to the proposed experiment for a 2 days monitoring

delayed 165 minutes, just like in the two days experiment. The fact that the “bolus given in advance” scenario is repeated proves that there is more information to be extracted from the excitation of the insulin absorption model than in the rest of the system. This conclusion is repeated for both model’s experiment design. The simulated glucose profile of the model is shown in Figure 5.11.

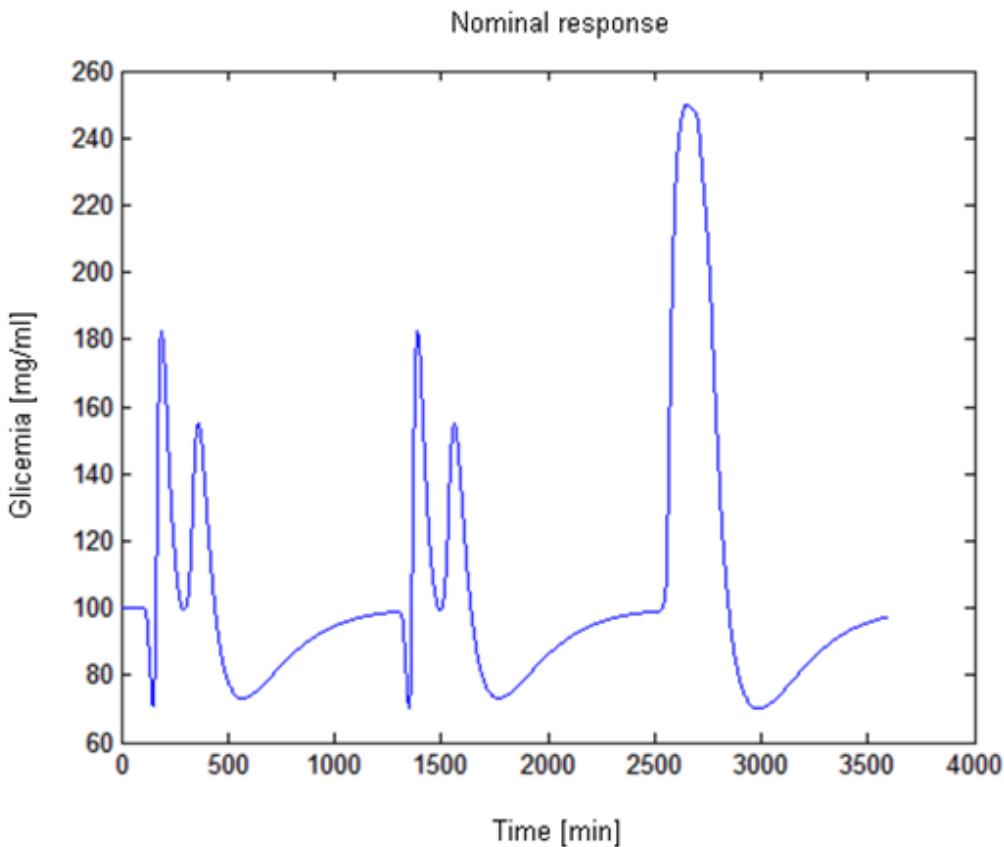


Figure 5.11: Modified Panunzi’s model response to the proposed experiment for a 3 days monitoring

This is a very good example of how the different experiments designed for each independent day are not dependent on the order of those days. Sometimes the case of the small meal and delayed bolus happens in the first day, some others at the end, and it can also be in the middle of the 3 days. Also, it is confirmed that increasing the number of experiments causes the richest information postprandial period to repeat in order to maximize the identifiability. No further synergies between different days where

observed, and every day is independent in the optimization algorithm.

The conditions of experiment designed for the final case of *four* day monitoring are:  
The conditions of experiment designed *three* day monitoring are:

- Day 1.  $\tau = 167, 06$  minutes, Meal = 44, 25 gr CHO and Bolus = 3 Insulin Units
- Day 2.  $\tau = -17, 22$  minutes, Meal = 100 gr CHO and Bolus = 10 Insulin Units
- Day 3.  $\tau = -18, 41$  minutes, Meal = 100 gr CHO and Bolus = 10 Insulin Units
- Day 4.  $\tau = 86, 19$  minutes, Meal = 47, 07 gr CHO and Bolus = 4, 13 Insulin Units

The conditions for day four are a little bit different than the previous ones, but it hardly defines a new type of therapy for a day of identification. It is a variation of the profile with a small meal and delayed insulin bolus shown in day one, considering a smaller delay so that the insulin bolus has not only a role of correction of the level of blood glucose, but also tries to counteract the rise of blood glucose that is still coming from the gut. That is observed in Figure 5.12, where the time of administration of the bolus does not wait for the blood glucose to stop rising and get stabilized, and instead, it forces it down before the absorption is over.

As a summary, with the optimal set of parameters of the modified Panunzi's model, two different profiles of glucose are determinant for the good identification of the model.

- The first therapy has already been seen in Bergman's model experiment design, and it consists of an advance of the bolus of insulin to the meal time, improving the identifiability of the insulin subsystem.
- The second therapy consists on a big delay (2-3 hours) in the administration of the bolus, while giving a small amount of carbohydrates (40-50 grams) to avoid grave hyperglycemia. With this therapy, separation of insulin and glucose dynamic from the meal is obtained, while maintaining plasma glucose in a safe range.

In Figure 5.13 the evolution with the number of days of the parameters identifiability is shown for the optimal set of parameters.

The graphs are very similar to those seen in the analysis of Bergman's model. The influence of the number of days, especially between days one and two, is very important

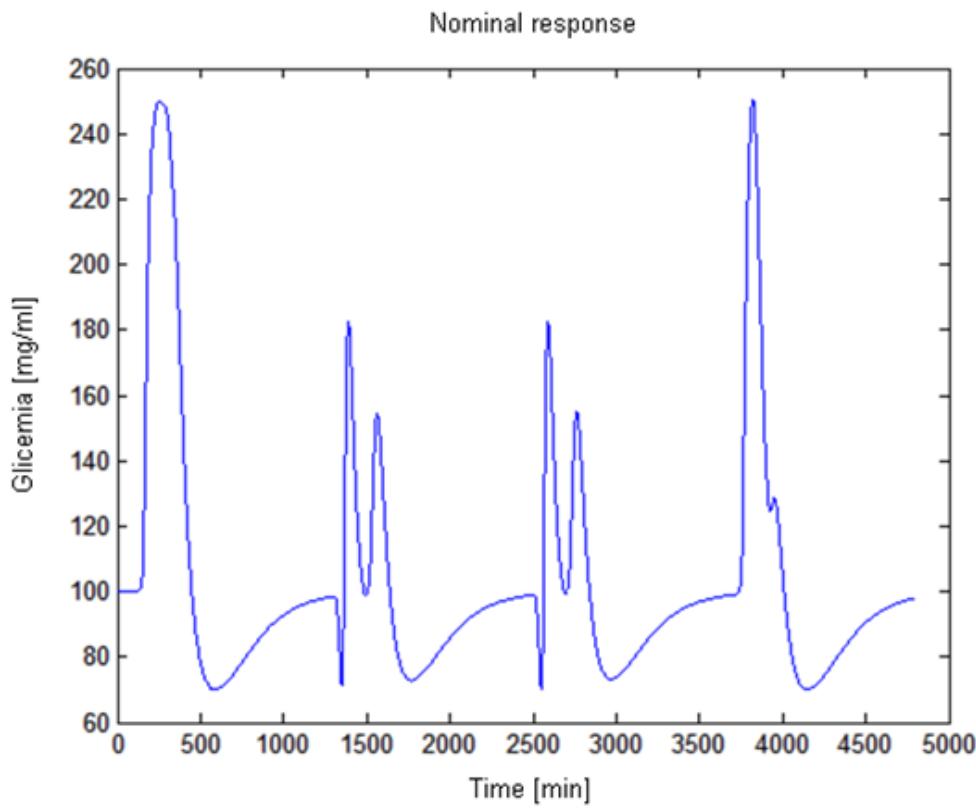


Figure 5.12: Modified Panunzi's model response to the proposed experiment for a 4 days monitoring period

both in the case of experiment and no experiment design. It is worth noting the difference on scale between the top and bottom pannel. While without experiment design the parameter's CVs start in some cases as high as 0,8, applying optimization of identifiability this starting variability is reduced to approximately 0,6. Again, a minimum of two days is required for optimal identifiability of the whole system, because of the great drop of CVs of all the parameters in the system, but also because of the inclusion of both therapies proposed which guarantee the separation of dynamics of both input systems. For the same reasons exposed in Bergman's model analysis, a three days experiment is preferred, and it concurs with the findings of the experiment design using modified Panunzi's model.

The influence of the experiment design in the identifiability of the model can clearly

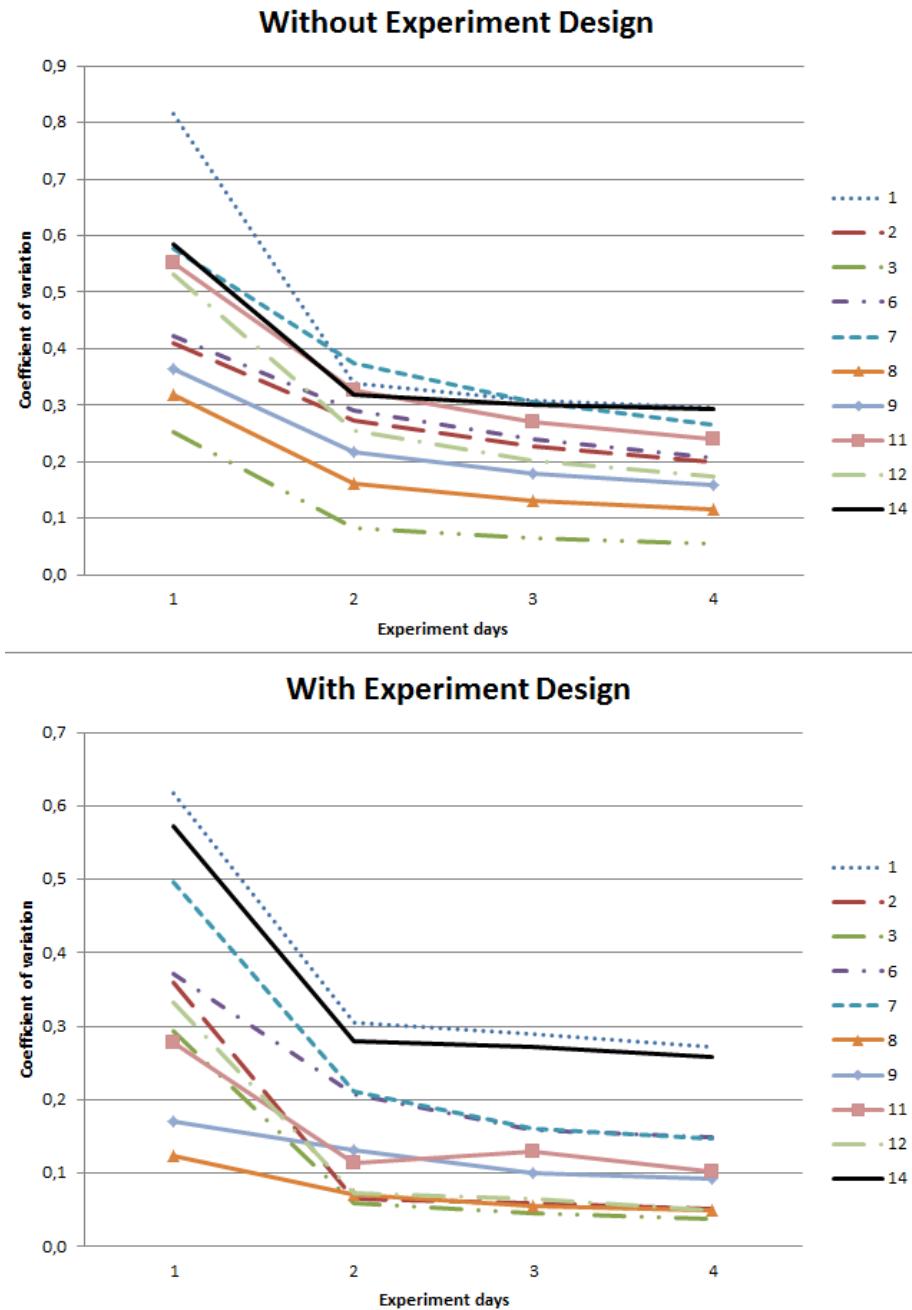


Figure 5.13: Modified Panunzi's optimal set of parameters and their coefficients of variation as they get smaller with the number of monitoring days. Top graph shows the evolution of the parameters' identifiability with experiment length. Bottom panel displays the identifiability of the same parameters applying the designed experiments.

be seen in Figure 5.14, where it is shown that experiment design improves identifiability dramatically. In this case, identifiability of almost all the parameters rises, getting their coefficients of variation divided by two or three. Special cases are the parameters 1 and 14, whose identifiability rises to a smaller extent. This stiff identifiability might be explained by the strong correlation between these two parameters since both of them are very related to the plasma insulin dynamics and its influence in blood glucose (parameter 1,  $K_{xgl}$ , is the insulin sensitivity and parameter 14,  $k_e$ , is the plasma insulin elimination). On the contrary, parameter 2 identification improves dramatically with experiment design.

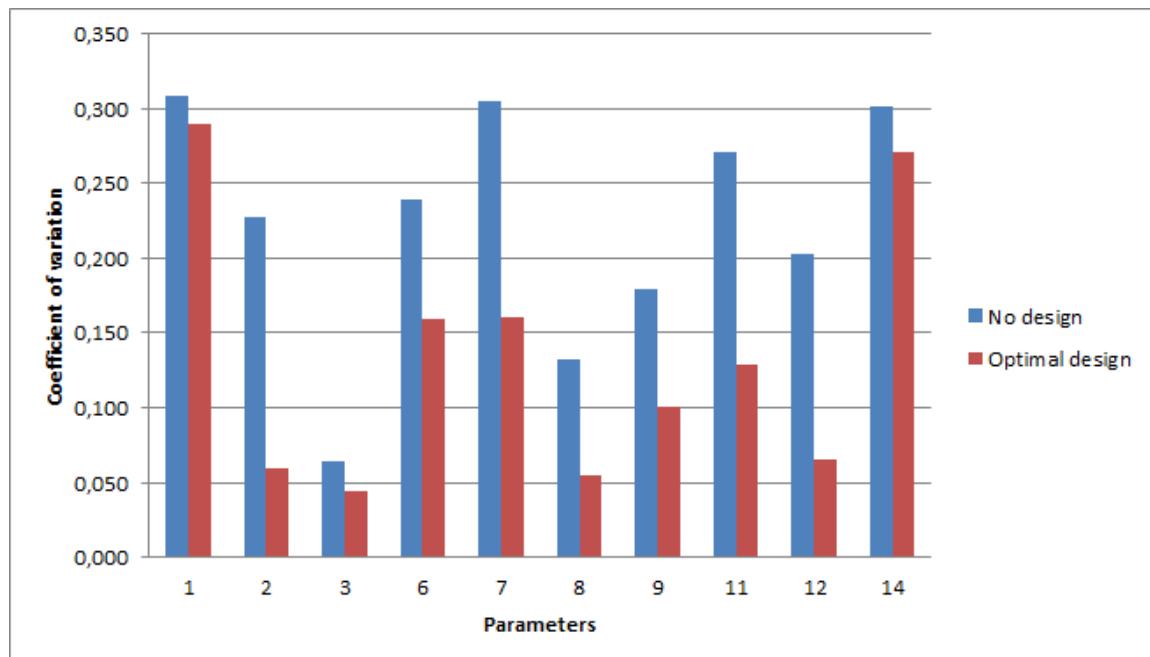


Figure 5.14: Identifiability of the optimal set of parameters of modified Panunzi's model before and after the experiment design for the case of a three meals monitoring

### 5.3 Discussion and clinical protocol

Identifiability has been drastically improved in both models, but in theory, the experimental design should not end here. The optimal design requires interaction with the experimental setup in order to keep improving the conditions of the identification. For example, the assumption of nominal values in the parameters for the analysis and

experiment design may be questionable, and every identification performed on real data must be subject to an identifiability analysis. With the collection of new experimental data, we may be able to design further experiments based on a new setpoint on the parameter space.

Up to this point, only two models were tested following the exposed methodology, but the experiment design is open for more models to be challenged in their identifiability. When looking at the identifiability analysis, it was accepted that a parameter was identifiable if its CV was below 0.3. After the optimal experiment design, the optimal set of parameters has had the CVs of all its parameters reduced in many cases far below the 0.3 threshold. Naturally, the question arises: “what if more parameters were included in the *optimal* set?”. ‘It is very much possible that those parameters CV drop below the identifiability threshold after the optimization of identifiability. However, it is impossible to know this *a priori*, and a iterative approach must be used to check this hypothesis. Unfortunately, optimal experiment design is a very heavy computation task, and embedding it in an iterative paradigm seems unreasonable, and thus this hypothesis is discarded.

The results of the experiment design do not have to be taken literally in practice. The assumptions made when using each model with its nominal parameters have great importance on the results of the experiment design. From a practical point of view, experimenting with several different models allows us to draw some general qualitative conclusions. That is why the design has been repeated several times in this thesis. Qualitative and common characteristics of the experiments designed are extracted in the next lines, trying to get the most important features needed in a monitoring for the good posterior identification of a patient, regardless of the model used.

The experiment design aims to separate the dynamics of the sub-models for a better identification. There are two different single day experiments present: in one case insulin dose is given without meal perturbation, limited by the hypoglycemia constraint, in order to obtain information of the subcutaneous insulin model. The other situation is exactly the opposite: a meal is given while the insulin dose is delayed limited by the hyperglycemia constraint. During this time, the intestinal absorption model is the only one acting on the output signal. Optimal experiment design forces the model to move in all the range of operation set in order to preserve patient’s safety, thus extracting more information out of the data. Maximum excitation of the model is required, while separating the perturbations to the output signal for maximum identifiability.

The application of the designed experiment in ambulatory conditions is not

straightforward. In the optimal experiment design the (virtual) patient is supposed to be stabilized into euglycemic range. In real experiments, even during the fasting state, patients may not be under steady state conditions, presenting glucose concentration measurements in any glycemic range. Under these circumstances, application of the optimal experiment design to the clinical setting requires some cautions. Indeed, if the patient is hyperglycemic at the beginning of the experiment, delaying more than two hours the insulin dose can be dangerous for the patient's health, with extremely high glucose values. On the other hand, if the patient is hypoglycemic (or prone to hypoglycemia) in the preprandial period, administrating an insulin dose without carbohydrates ingestion can lead to severe hypoglycemia.

Therefore, an adaptation of the optimal experiment design to the initial metabolic state of the patient is proposed. If the patient starts the monitoring with a blood glucose level over 150 mg/dL or in the 100-150 mg/dL range with an increasing trend (hyperglycemic risk), the insulin bolus is given followed 30 minutes later by a 100 grams meal. If the glucose level is below 100 mg/dL or in the 100-150 mg/dL range with a decreasing trend (hypoglycemic risk), a 40 or 60 (patient's choice) grams meal is given and the insulin dose administration is delayed 120 minutes. In both situations the dose to be administered is the one recommended to the patient, depending on the size of the meal and his/her usual insulin-to-carbohydrate ratio. The patient can choose among three menus with the same relative nutritional composition. With this protocol the separation of dynamics is achieved, while minimizing risks for the patients. This protocol is summarized in table 5.5.

<b>Pre-prandial glycemia</b>	<b>Action</b>
Greater than 150mg/dL or between 100 – 150 and rising	Infusion of an insulin bolus 30 minutes before of 100 grams meal
Lesser than 100mg/dL or between 100 – 150 and decreasing	40 grams meal and bolus infusion 2 hours later

Table 5.5: Clinical protocol for application on ambulatory experiments

The protocol was approved by the Ethical Committee of the Clinic University Hospital of Valencia and home-made monitoring was performed for validation of the identification procedure. Patients were monitored during two weeks, with a wash-up week in between, and instructed to follow the protocol defined by the experiment design. In a week, three days are used for identification and three for validation, as it was scheduled due to restriction of the lifetime of the sensor utilized. The data acquired using this protocol was used by Rossetti *et al.* to test new computer generated insulin treatments for diabetic

patients [74].

# Chapter 6

## CGM Statistical Modeling and Validation

Identifiability is highly dependant on the quality of the measurements obtained. In this chapter I will move from the improvement on data acquisition methodology towards the analysis and improvement of data acquisition tools. Sensor induced errors are one of the main problems in diabetes experimentation nowadays. Continuous glucose monitoring (CGM) devices estimate plasma glucose from interstitial glucose measurements every 1-5 minutes. They are potentially much more informative than the traditional and sparse method based on capillary self-monitoring of blood glucose. However, CGM devices are only approved as adjunctive to capillary self-monitoring due to their low accuracy.

Despite its relevance to the interpretation of glycemic time-series and to the implementation of a future AP, few studies have investigated the error characteristics of current CGM devices, especially in the postprandial state. Breton et al. [6] developed an error model for CGM based on a posteriori recalibrated data from the FreeStyle NavigatorTM (Abbott Diabetes Care, Alameda, USA). However, in contrast to real-time calibration algorithms, a posteriori recalibration implies the use of all reference BG points to minimize the sensor readings-reference glucose mismatch. In addition, Facchinetto et al. [24] demonstrated using simulation studies how small errors in data recalibration can affect the derived statistical properties for the error. Therefore, error characterization of real-time CGM devices in the postprandial state is still an open issue and an area for additional investigation.

In this chapter I performed an analysis and modeling of the postprandial error of two commercial real-time CGM devices, the Dexcom® SEVEN® PLUS and the Medtronic® Paradigm® Veo™, that were simultaneously worn by subjects with type 1 diabetes following a set of four standardized mixed meals. Statistical properties of the signal error were analyzed and a new simulation model was generated aiming for its integration into in silico studies.

## 6.1 Data and methodology

Twelve subjects with type 1 diabetes treated with continuous subcutaneous insulin infusion (CSII) (male/female 3/9, age  $41.8 \pm 7.3$  years, diabetes duration  $20 \pm 10$  years, HbA1c  $8.0 \pm 0.6\%$ ) were studied in the postprandial state under controlled conditions, on four different occasions. Patients were monitored twice when eating a standardized meal with 40 g CHO content and twice, in addition, when they received a 100 g CHO meal. On each occasion, CGM was performed simultaneously with a SEVEN® PLUS (Dexcom, San Diego, USA) device and a Paradigm® Veo™ (Medtronic, Northridge, USA) device.

Both devices were inserted the day before the experiment, in order to allow the devices to stabilize before the experiments. During the first day the monitors were calibrated according to the manufacturer instructions using the same capillary measurements. The day of the experiment, the devices were calibrated using the same capillary measurement obtained in fasting state.

The patients arrived at 9:00 AM to the clinic. Pre-prandial plasma glucose was standardized around 100 mg/dL by means of a feedback insulin infusion. At 13:00 h, patients ate the standardized meal and received the corresponding insulin bolus using a subcutaneous insulin infusion system. Plasma glucose levels were measured for 5 hours after the meal, every 5 minutes the first two hours after the meal and every 10 minutes afterwards, thus obtaining more frequent measurements during the meal rise phase. Plasma blood glucose levels were measured on arterialized venous blood using a reference method (YSI 2300 STAT Plus Glucose analyzer, YSI Yellow Springs, Ohio, USA), and synchronously compared with CGM's subcutaneous measurements (5 min frequency). Only forty-seven datasets were available for analysis when the SEVEN® PLUS device was used, because one dataset was discarded due to a flat sensor measurement at 40 mg/dL during the whole duration of the experiment. In the case of the Paradigm® Veo™ device, only 42 datasets were available because the first 6 experiments were not

monitored with this device.

YSI data is assumed to be very accurate, especially if compared with any CGM device. Nowotny et al. [60] evaluated the accuracy of the YSI 2300 Glucose Analyzer among other devices, showing 2.2% accuracy. This is higher than previous reference methods, and also better than the rest of the analyzers evaluated.

In order to obtain a time series of the same sample size, YSI data were linearly interpolated in the late postprandial period in order to obtain a 5 minutes periodic consistent dataset, and be able to compare it with the CGM data. Although interpolation of reference data can introduce undesired dynamics when modeling, in this case interpolation only occurs in the late postprandial period, where dynamics of both the glucose and the monitor were slower. In addition, linear interpolation were used in literature before, in the context of glucose modeling of sensor signals [6, 23] where reference measurements were sparse (15 minutes).

The signal analyzed was the CGM error calculated as:

$$E(k) = CGM(k) - YSI(k) \quad (6.1)$$

where  $E(k)$  is the error at time  $k$ ,  $CGM(k)$  is the CGM signal and  $YSI(k)$  the reference blood glucose, both at the same time  $k$ . Some issues were taken into account in the analysis. First, the existence of a physiological delay between subcutaneous and plasma glucose, which could be patient-specific [42]. Second, filters included in the CGM devices may introduce additional delays (signal-processing delays). Both types of delays represent a confounder factor of CGM accuracy and should be compensated prior to further study of the error. Third, stationarity of the error time-series should be investigated since non-stationarity may yield misleading probability distributions. Indeed, when fitting probability distribution and time-correlation of a signal, stationarity of the process is assumed. In the event of having a time varying stochastic process, non-linear transformations must be applied to the signal so that it becomes as time unvarying as possible [5]. The choice of the non-linear transformation to be used usually depends on the problem and the data available.

The analysis described in Figure 6.1 was followed for both CGM devices:

1. The delay of the CGM signal with respect to reference glucose was computed as the lag with maximum CGM versus YSI cross-correlation for each patient.

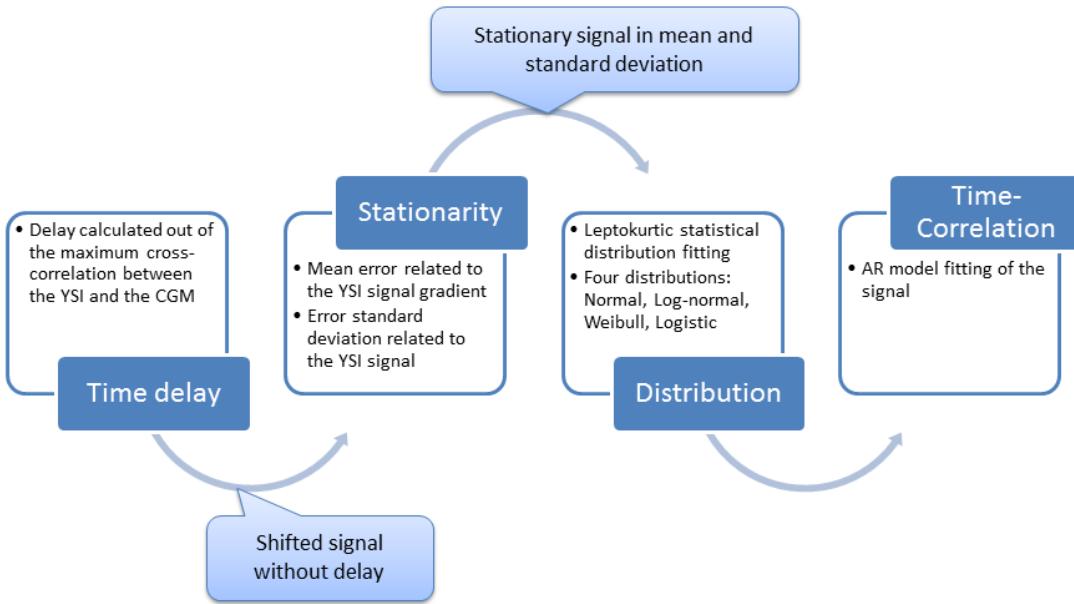


Figure 6.1: Summary of stepwise statistical analysis process.

A probability distribution function was fitted to the delay's histogram using a classical maximum likelihood (ML) estimator from Matlab Statistics Toolbox. Error time-series were then shifted by the corresponding delay for compensation prior to next analysis thus “synchronizing” CGM and reference glucose signals.

2. Stationarity of the shifted error time-series was analyzed. A signal was considered to be stationary if time invariance of the statistical moments held. Mean and standard deviation of the error signal were calculated across the population of sensors in every instant of the postprandial period in order to obtain a time dependent signal. Time dependence of the mean and the standard deviation was analyzed. Besides graphical inspection, datasets were fitted to autoregressive (AR) models and the presence of a unit root was investigated (random walk process in the case of first order models). A stationary process should have no unit roots.
3. Finally, the sample probability distribution for the error was analyzed and an AR model fitted to reproduce time correlation and for simulation purposes. A set of candidate probability distribution functions was defined and ML estimators in Matlab Statistics Toolbox were used for data fitting for every distribution. A simple quadratic error index was used to measure the fit for comparison purposes.

When performing a linear regression or just a simple correlation between variables, correlations were tested to be statistically significant using a t-statistic testing the null hypothesis of non-correlation

## 6.2 CGM Modelling

Both CGM devices were first compared to the reference measurements in order to get a general error magnitude of each error. For this comparison, the MARD (Mean Absolute Relative Deviation) was computed, following the next formula:

$$MARD := \frac{\sum_{i=1}^n |(CGM_i - YSI_i)/YSI_i| \cdot 100}{n} \quad (6.2)$$

where  $n$  is the number of samples,  $CGM_i$  stands for the sensor glucose reading at time  $i$  and  $YSI_i$  is the reference glucose at time  $i$ . The SEVEN® PLUS monitor showed a MARD of 17.28%, and the VEO monitor 20.12%, for the whole dataset.

Clinical accuracy for both datasets is illustrated in Figure 6.2 using two complementary error grid analysis (EGA) methods: the Clarke-EGA [13] and the glucose-rate grid proposed by Kovatchev et al. [49]. Regarding the Clarke grid, both monitors show very similar behavior. The SEVEN® PLUS monitor placed 73.04% of the data pairs in zone A, and 26.41% in zone B, while Veo system had 66.23% of the data pairs in zone A, and 33.25% in zone B. Both monitors had less of 1% of the monitoring data within the zones C, D or E. Regarding the glucose rate grid, the SEVEN® PLUS monitor had 80.96% of data points in zone A, 14.89% in zone B, 2.41% in zone C, 1.42% in zone D and 0.32% in zone E. Veo monitor showed 83.37% of the data points within the zone A, 13.53% in zone B, 0.67% in zone C, 2.10% in zone D, and 0.31% in zone E.

A representative CGM trace from each device is shown in Figure 6.3. The YSI signals shown correspond to different patients following the detailed experiments. In the SEVEN® PLUS monitor sample, a rapidly increasing glucose concentration after the meal is observed, followed by a slow decreasing glucose in the late postprandial period. The VEO monitor sample shows a rapidly increasing glucose response right after the meal and a rebound increase in the late postprandial period. It can be appreciated that both monitors have trends similar to the YSI signal they are estimating.

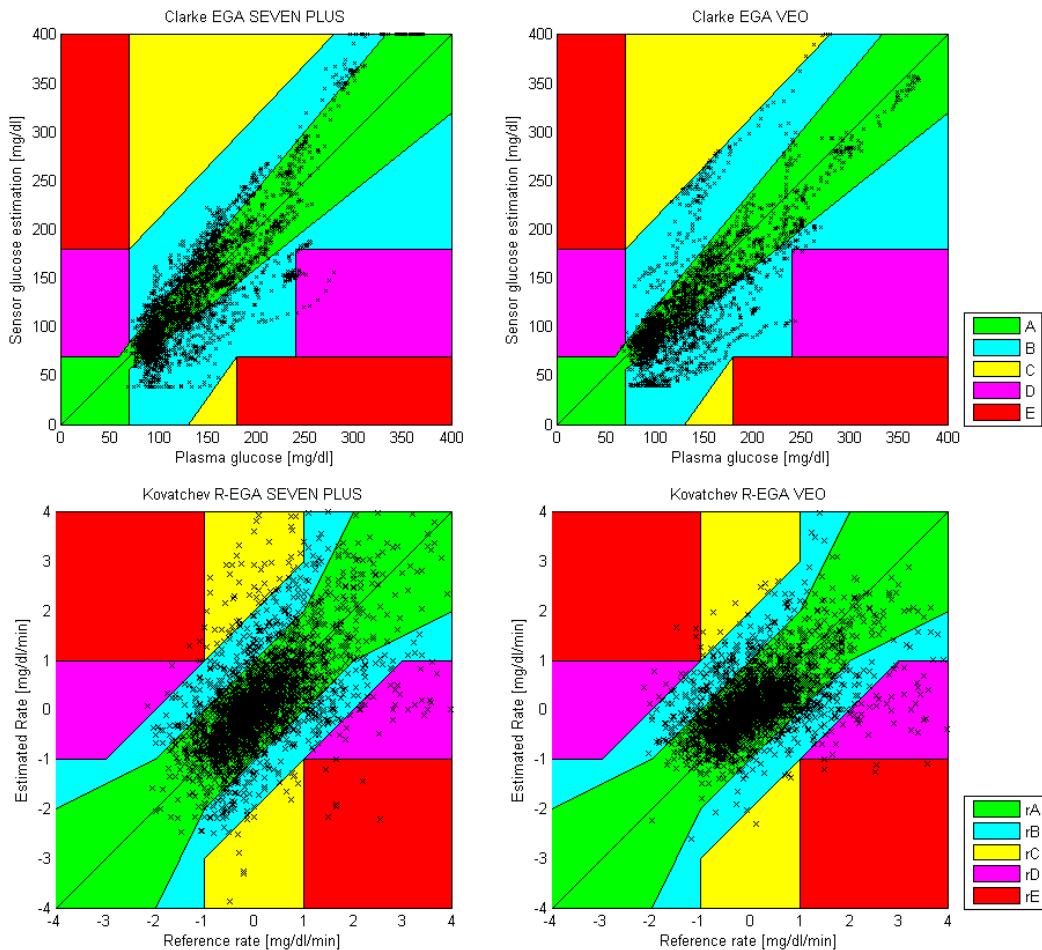


Figure 6.2: Clarke error grid (EGA) and rate error grid from continuous glucose-EGA plots for both monitors. Top row shows Clarke EGA plots, observed glucose reference measurements *versus* the prediction of the monitors, SEVEN® PLUS on the left and VEO on the right. On the bottom row, the rate error grid from continuous glucose-EGA is displayed, for the SEVEN® PLUS on the left and VEO on the right.

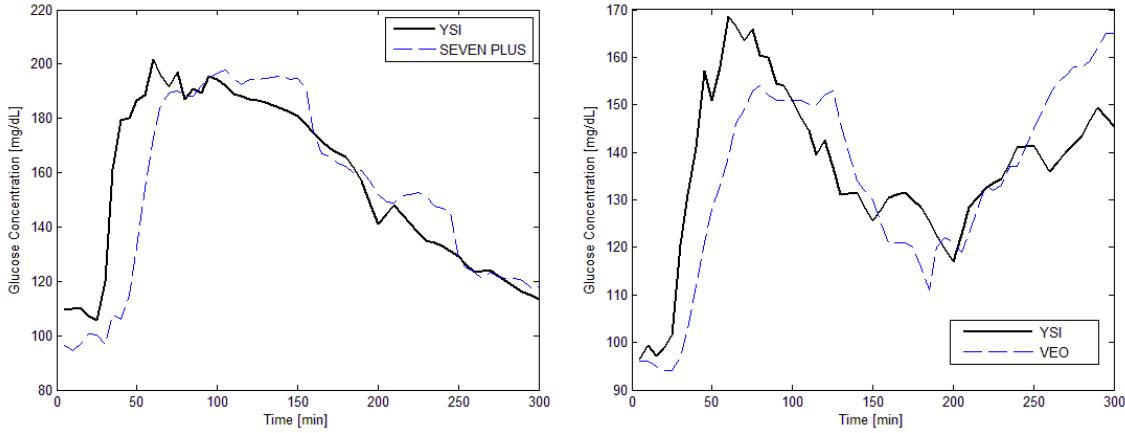


Figure 6.3: Representative sample of two different postprandial periods for the two monitors. The left pannel shows the SEVEN® PLUS monitor sample and the right pannel shows the VEO monitor sample.

### 6.2.1 Analysis of delay

The delay histogram for all the available datasets is shown in Figure 6.4. This delay was consistent throughout the entire postprandial period, and no significant difference on the calculated delays was achieved by considering only the first two hours (transient) of the entire postprandial period. Observing the time correlation of the datasets, in the SEVEN® PLUS monitor, 28 showed no significant delay between the two signals (i.e. delay was less than five minutes due to the CGM resolution), and 19 for the Paradigm® Veo™ .

The histogram analysis hinted that the delay follows an exponential probability distribution:

$$f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}} \quad (6.3)$$

where  $\mu$  is the exponential parameter. The exponential parameter was adjusted for each monitor. For the SEVEN® PLUS monitor the parameter was  $\tau_{sevenplus} = 1.08$  with a 95% confidence interval [0.82, 1.46] and for the Paradigm® Veo™  $\tau_{veo} = 1.69$  with a 95% confidence interval [1.28, 2.35], resulting in the curve fitting shown in red in Figure 6.4.

The observed delay for the SEVEN® PLUS device was consistent with previous works [86] for the SEVEN® system, where an average value of 5 minutes is reported as the lag for maximum cross-correlation. However, no probability distribution for the delay

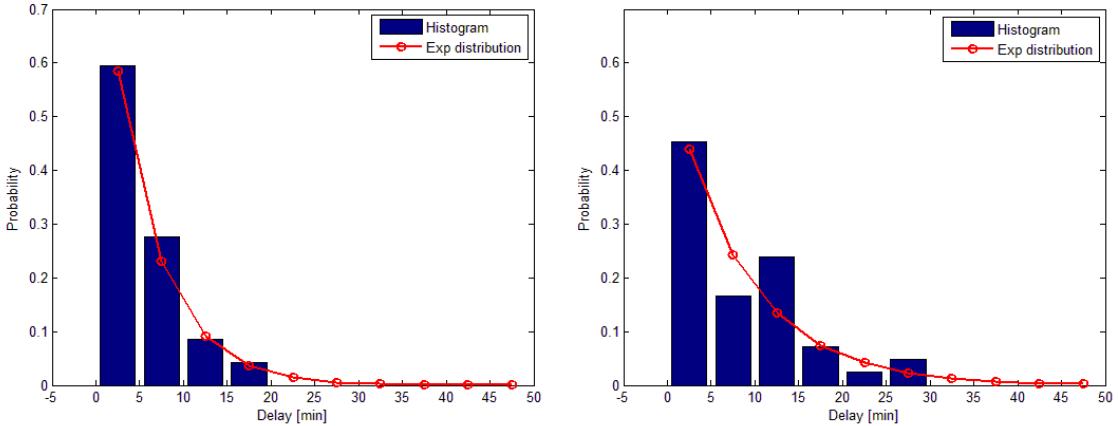


Figure 6.4: Delay-time histograms for the SEVEN® PLUS monitor (left) and the Paradigm® Veo™ (right).

was reported for comparison. Compared with the Dexcom system, the Paradigm® Veo™ showed a significant ( $p=0.007$ ) larger average delay (8 minutes). Our results were slightly larger than those reported by Keenan et al. [43], but neither these values nor the SEVEN® Plus values were comparable due to the difference in computation methods. Therefore, measurement of the delay should be standardized in order to enable comparisons between different studies. Additionally, the Paradigm® Veo™ exhibited a slower decay in the fitted exponential distribution, indicating higher variability of the delay value up to 30 minutes.

### 6.2.2 Analysis of Stationarity

The statistical moments of the data analyzed after delay compensation were not time invariant, as shown in Figure 6.5. For both monitors, the mean and standard deviation of the error were non-stationary. Our goal was to find a transformation of the signal in order to make it as close to stationary as possible. This transformation had to be invertible for simulation purposes. This means that the transformation had to be based on available information in the simulation context, e.g., plasma glucose from a glucoregulatory model, or equivalently, the YSI signal from the experimental data.

In general, stationarity of the error time-series is assumed when fitting probability distributions to a stochastic signal. However, it can be observed that these assumptions were not correct when interpreting the error of CGM in the postprandial state. Mean

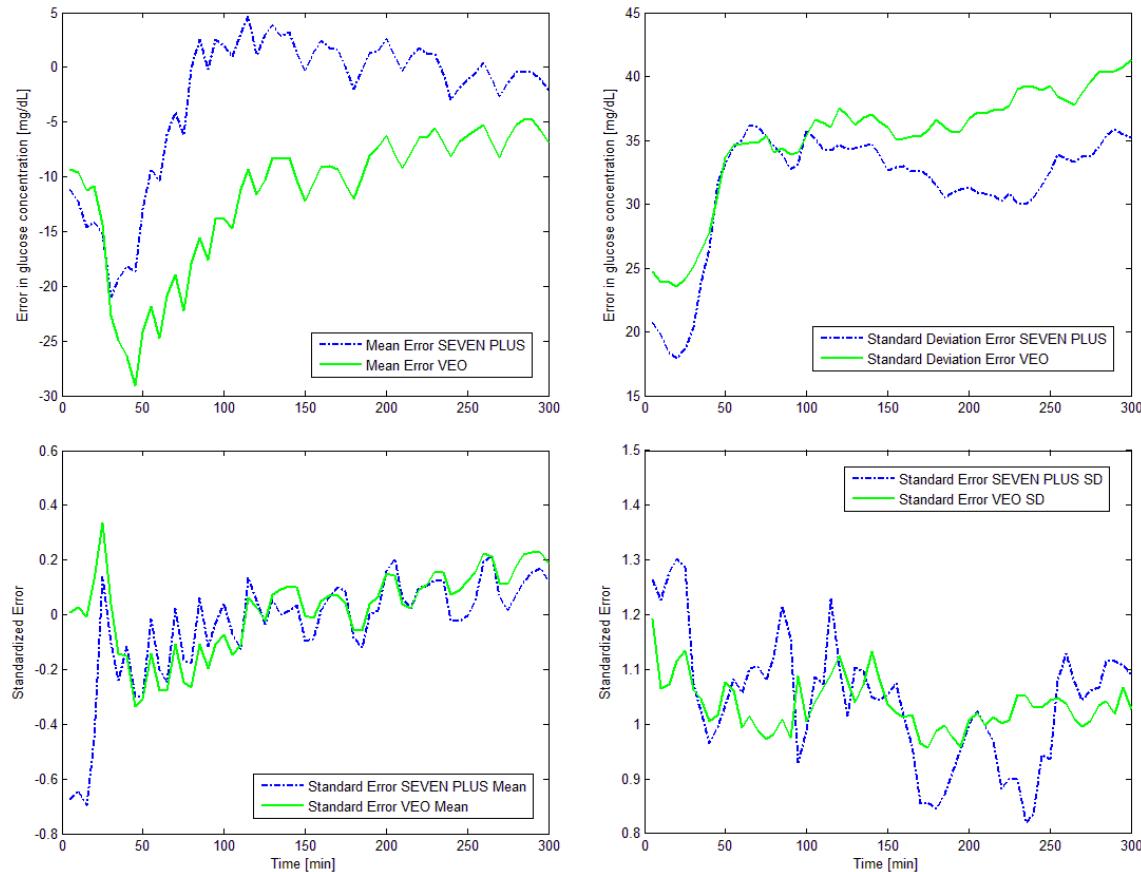


Figure 6.5: Top panels show the meal error and standard deviation of the error after delay compensation for the SEVEN® PLUS and the Paradigm® Veo™ monitors. Bottom panels show the same signals after the standardization process. The left figures show the mean error for both monitors while on the right the standard deviation of the error signals is represented. Notice the different scales on all four plots.

and standard deviation after delay compensation were both time-varying, especially for the first two hours after a meal intake. This may be related to the performance of the real-time calibration algorithm after the high rise of glucose following the intake up to its peak value.

Despite the standardization of the initial conditions by means of an insulin feedback phase, both monitors had an initial mean underestimation of glucose of 10 mg/dL. Surprisingly this value was equal to both devices. It is unknown to the author whether it may correspond to a setting of the manufacturers. An initial standard deviation of 20-25 mg/dL was observed yielding significant errors.

Both monitors were worn simultaneously by the patient and calibrated with the same calibration points obtained in fasting state the same day of the experiment. Calibration points have been recognized as crucial factors influencing the accuracy of CGM readings [7, 85]. In this case, the calibration performed was done 4-5 hours before the experiment in fasting state, and it did not seem to induce differences between both devices.

After meal intake, both devices were unable to follow the high rise of glucose, increasing the initial underestimation up to a peak value (approximately 20 mg/dL for SEVEN® Plus and 30 mg/dL for the Paradigm® Veo™) about 50 minutes later. The particular sensor delay should be additionally considered for a correct interpretation. During this time, the error standard deviation increased steadily, having a similar peak time as the mean error, with a peak value of 35 mg/dL for both monitors.

After this initial period, the glucose underestimation decreased. The CGM glucose estimation recovered slowly towards the glucose reference value. A faster recovery was observed by the SEVEN® PLUS. It is noteworthy that approximately zero mean error was observed for the SEVEN® PLUS monitor after peak time, in contrast to VEO signal, which appears to be showing a bias in the mean error. This bias converged slowly to zero, but the 5 hours postprandial monitoring window we used here was insufficient for capturing the whole stabilization of the mean error in the VEO monitor.

If variability of CGM readings were reduced, our results suggest that CGM would allow for a good representation of postprandial glucose. Regarding the standard deviation, a plateau at 35 mg/dL was reached for the SEVEN® PLUS, while a slight increasing trend was obtained for the Paradigm® Veo™. Thus, the meal event represented a big challenge for both devices producing a significant variability, which is still a major issue in CGM performance. Certainly, sources of this variability should be investigated in future studies to confirm whether it is due to variability of the physiological delay, the

intensity signal produced by the sensor or the calibration algorithm itself.

A positive correlation between the YSI signal and the standard deviation of the error was found. This is an indication of the variability of the sensor to capture the postprandial peak. In addition, a negative correlation of the YSI rate of change and the mean of the error reflects that sensor readings were falling behind the reference glucose in the rising trend (the error becomes more negative in average), in spite of delay compensation.

The second moment of the error signal closely resembled the YSI blood glucose measurements for both monitoring systems. The correlation coefficient for the SEVEN® PLUS monitor was  $r_{sevenplus} = 0.93$  ( $range_{95\%} = [0.90, 0.96]$ ) and for the Paradigm® Veo™  $r_{veo} = 0.82$  ( $range_{95\%} = [0.72, 0.89]$ ), both significant ( $p < 0.005$ ). The regression lines for these correlations were:

$$\text{SEVEN PLUS: } STD(k) = 0.3142 \cdot YSI(k) - 12.9056 \quad (6.4)$$

$$\text{VEO: } STD(k) = 0.2711 \cdot YSI(k) - 3.7679 \quad (6.5)$$

where  $STD(k)$  is the standard deviation of the error at time  $k$ .

As for the mean of the error signal, the correlation with the raw YSI signal did not show a significant correlation coefficient. The correlation with the gradient of the YSI signal was better, with a coefficient of  $r_{sevenplus} = -0.79$  ( $range_{95\%} = [-0.87, -0.68]$ ) and  $r_{veo} = -0.54$  ( $range_{95\%} = [-0.70, -0.33]$ ), both being statistically significant ( $p < 0.005$ ). This correlation was not as evident as the one with the second moment, but it was much better than the correlation found to any other signal derived from the YSI. Regarding to the correlations found we can guess that stationarity achieved with these transformations was better in the SEVEN® PLUS monitor than in the Paradigm® Veo™. The regression line for the correlation of the first moment was:

$$\text{SEVEN PLUS: } Mean(k) = -2.0046 \cdot dYSI(k)/dt - 1.0628 \quad (6.6)$$

$$\text{VEO: } Mean(k) = -1.323 \cdot dYSI(k)/dt - 10.517 \quad (6.7)$$

where  $Mean(k)$  is the average error at time  $k$ . Again, as it happened with the delays, these regressions were consistent throughout the entire postprandial period. Considering these regressions as good characterizations of the statistical moments of the signals, the error signal can be transformed as follows, so that it becomes (quasi-)stationary:

$$\bar{E}(k) = \frac{E(k) - \text{Mean}(k)}{\text{STD}(k)} \quad (6.8)$$

In case of perfect estimation of the mean and standard deviation at time  $k$ , the transformed error signal  $\bar{E}(k)$  will have zero mean and unity standard deviation throughout the postprandial period.

Stationarity of the signal was improved after this transformation, as shown by the roots of the AR models fitted before and after the “standardization”. The error signals were tested to fit AR models, looking for the lowest-order model with random uncorrelated residuals. The transformed statistical moments along time can also be seen in Figure 6.5, where a clear improvement on stationarity of both signals can be appreciated comparing top and bottom panels.

The SEVEN® PLUS time-series showed a good fitting to a first order AR model before and after the transformation. The model root was  $z = 0.9892$  before the transformation, and  $z = 0.9249$  after. The transformation moves the dynamics of the data away from those of a random walk (unit root), thus making it more stationary. The Paradigm® Veo™ time-series adjusted well to a first order AR model before the transformation and to a third order model after. The model root before the transformation was  $z = 0.9977$ , reaffirming the non-stationarity of the process. After the transformation, the model root closer to unity was  $z = 0.9846$ , showing some improvement towards stationarity, but not as much as in the SEVEN® PLUS. Correlations were weaker in the case of the Paradigm® Veo™ which translated into a worse compensation for non-stationarity.

### 6.2.3 Distribution fitting

The datasets used for distribution and autocorrelation fitting were already transformed following the delay compensation and the standardization processes explained before. Kurtosis was high for both datasets, so leptokurtic distributions were used for fitting. Asymmetry was very low in the data, but some non-symmetric distributions were also

tested. Four different statistical distributions were finally considered:

$$\text{Weibull: } f(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} \quad (6.9)$$

$$\text{Normal: } f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \quad (6.10)$$

$$\text{Log-Normal: } f(x) = \frac{1}{x\sigma\sqrt{2\pi\sigma^2}} e^{-(\log x - \mu)^2/2\sigma^2} \quad (6.11)$$

$$\text{Logistic: } f(x) = \frac{1}{1 + e^{\frac{x-\mu}{s}}} \quad (6.12)$$

Five Paradigm® Veo™ sensors showed unusual behaviors with very big mean absolute errors (over 50 mg/dL), even though their calibrations were correct. These offsets produced a multimodal histogram (see Figure 6.6). Given that this behavior was only observed in 5 sensors, those monitoring periods were not used to fit a probability distribution, remaining a total of 37 datasets. Given that the SEVEN® PLUS did not register faulty monitoring periods, the Paradigm® Veo™ seems to be more prone to abnormal sensor behaviors. This should be confirmed with larger studies.

The Log-Normal and Weibull distributions are only defined for positive values. To cope with this problem, data were transformed by adding its absolute minimum value (equal to 4.3069 for the SEVEN® PLUS and 2.4439 for the Paradigm® Veo™), thus making all errors positive and not centered in zero. The estimates are shown in Table 6.1. The best estimator was given at the normal distribution for the Paradigm® Veo™ ( $\mu_{Veo} = 2.435$  and  $s_{veo} = 0.5415$ ) and the logistic distribution for the SEVEN® PLUS ( $\mu_{sevenplus} = 4.297$  and  $s_{sevenplus} = 0.587$ ). Probability plots for the logistic and normal distribution are shown in Figure 6.6. However, the difference between the fit of the logistic distribution and the normal distribution for the SEVEN® PLUS was almost negligible ( $1.92 \times 10^{-3}$  vs  $2.02 \times 10^{-3}$ ). Normality assumption after all the transformations seems thus sensible.

Finally, a first order AR model explains appropriately the error data for the SEVEN® PLUS, which is consistent with Breton et al. [6]. However, a third order AR model was needed for the Paradigm® Veo™. This may be due to distinct filtering algorithms for the raw signal since both monitors use linear regression based calibration algorithms. Based on the previous analysis a model for each monitor was derived. This model incorporated time-varying statistical characteristics of the error for their integration into in silico tests for controller validations. These models are detailed in the next chapter, along with the full CGM simulation model of each monitor.

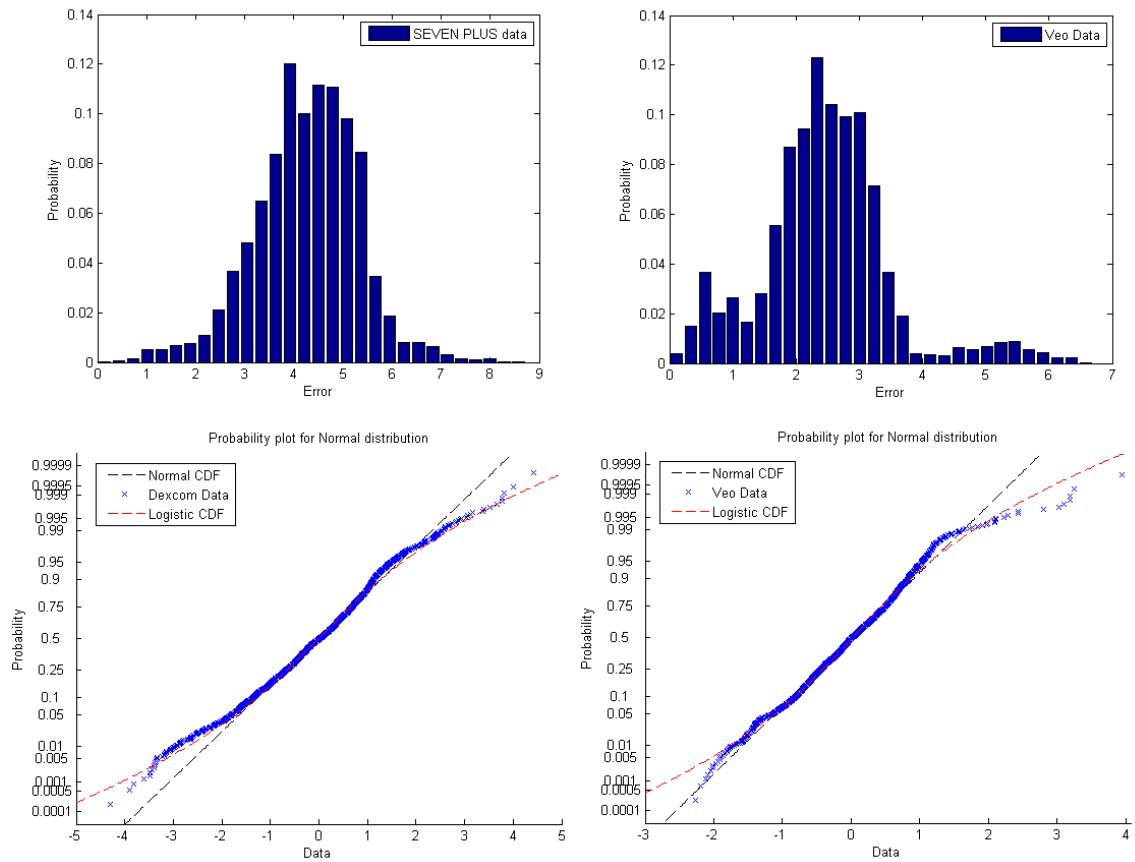


Figure 6.6: Histograms of the standardized error (top) and scaled cumulative distribution (bottom) plots for the SEVEN® PLUS monitor (left) and the Paradigm® Veo™ (right).

Distribution	Parameters SEVEN® PLUS		Fit $\times 10^{-3}$	Parameters Paradigm® Veo™		Fit $\times 10^{-3}$
	$\mu$	$\sigma$		$\mu$	$\sigma$	
Normal	4.27	1.057	2.02	2.282	0.693	1.39
Weibull	4.652	4.329	2.49	2.515	3.412	2.56
Logistic	4.297	0.587	1.92	2.289	0.387	1.84
Log-normal	1.402	0.764	42.6	0.751	0.866	42.61

Table 6.1: Parameters of the distribution fitting results for the SEVEN® PLUS and the Paradigm® Veo™ monitors

### 6.3 Validation

Following the error signal analysis, simulation models were built for the SEVEN® PLUS and Paradigm® Veo™ consisting on the following steps:

1. A time series was generated from the fitted AR model.
2. Transformations applied to get (quasi-)stationarity were inversely applied reproducing time varying characteristics of error mean and variance.
3. A realization for the delay following the identified probability distribution was computed and applied.

The SEVEN® PLUS transformed error time-series followed a first order AR model

$$\bar{E}(k) = \alpha_1 \cdot \bar{E}(k-1) + \beta \cdot w(k), \quad \bar{E}(0) = w(0) \quad (6.13)$$

where  $\bar{E}(k)$  is the correlated standardized error (output of the AR system),  $w(k)$  is a white noise signal,  $\alpha_1 = 0.9249$  is the AR parameter, and  $\beta = 0.3756$  is the correction of the variance factor, chosen in order to get the desired characteristics of the normal distribution for the data after filtering. The Paradigm® Veo™ time-series followed a third-

order model:

$$\begin{aligned}\bar{E}(k) &= \alpha_1 \cdot \bar{E}(k-1) + \alpha_2 \cdot \bar{E}(k-2) + \alpha_3 \cdot \bar{E}(k-3) + \beta \cdot w(k) \\ \bar{E}(0) &= w(0), \quad \bar{E}(1) = w(1)\end{aligned}\tag{6.14}$$

with :  $\alpha_1 = 0.9471$ ,  $\alpha_2 = -0.1936$ ,  $\alpha_3 = 0.2271$  and  $\beta = 0.2198$ .

The white noise used as input of the AR models follows the normal distribution fitted to the data in the previous chapter. Normality was assumed for both monitors. For the SEVEN® PLUS the mean was  $\mu_{SEVENPLUS} = 4.27$  and the standard deviation  $\sigma_{SEVENPLUS} = 1.057$ , considering a shift of 4.3069. For the Paradigm® Veo™ the parameters were:  $\mu_{Veo} = 2.282$   $sigma_{Veo} = 0.693$ , and a shift of 2.4439.

In order to reproduce the time-variance of the error mean and standard deviation, we had to invert the standardization applied to the CGM error data during the analysis process, i.e.:

$$E(k) = STD(k) \cdot \bar{E}(k) + Mean(k)\tag{6.15}$$

$STD(k)$  and  $Mean(k)$  were calculated as:

$$\text{SEVEN PLUS: } STD(k) = 0.3142 \cdot \widetilde{PG}(k) - 12.9056\tag{6.16}$$

$$\text{VEO: } STD(k) = 0.2711 \cdot \widetilde{PG}(k) - 3.7679\tag{6.17}$$

$$\text{SEVEN PLUS: } Mean(k) = -2.0046 \cdot \frac{d\widetilde{PG}(k)}{dt} - 1.0628\tag{6.18}$$

$$\text{VEO: } Mean(k) = -1.323 \cdot \frac{d\widetilde{PG}(k)}{dt} - 10.517\tag{6.19}$$

where  $\widetilde{PG}(k)$  is the plasma glucose given by the virtual patient simulator, or in case of using real patient data, the plasma glucose measurements. The CGM readings were then calculated by adding the error to the simulator glucose output, and shifting  $d$  steps, where  $d$  is a realization of an exponential distribution  $Exp(\mu)$ :

$$\begin{aligned}CGM(k) &= \widetilde{PG}(k-d) + E(k-d), \quad d = Exp(\mu) \\ \mu_{sevenplus} &= 1.08, \quad \mu_{veo} = 1.69\end{aligned}\tag{6.20}$$

An illustration of several simulation runs for both models is shown in Figure 6.7. Average MARD for the simulated SEVEN® PLUS was 18.72%, and for the simulated Paradigm® Veo™ was 18.38%.

For validation purposes a virtual dataset a hundred times larger than the original dataset was created (the model was run a hundred times for each YSI trace), to better simulate the fitted distributions in their entire domain. The YSI samples used for the calculation non-stationarity of the error were extracted from the same dataset used for the model fitting. For this validation, the Paradigm® Veo™ data used for comparison only comprehends predictions from the 37 sensors than showed no unusual behavior. In Figure 6.8 validation of the signal delay is shown, comparing the probability distribution of the real dataset versus the simulated data. Figure 6.9 shows the comparison of the statistical properties of the data and the simulation, as well as illustrates the time variation of the simulated data.

Computer simulations of the adjusted model for both monitors show very similar delay distribution as the original data, as shown in Figure 8. The histograms of both simulated models are much more uniform than the original data, following closely the original exponential distribution that generated them.

Validation of the statistical moments and non-stationarity of the error can be extracted from Figure 6.9. Time variant mean and standard deviation were obtained for the simulated data, with similar dynamics than the original data. Mean signal was almost identical for the SEVEN® PLUS device, but it showed an offset for the late postprandial period in the Paradigm® Veo™ simulation (approx. 15 mg/dL). This was expected since the correlations for the Medtronic monitor were much lower than those of the Dexcom device. Standard deviation time variation was also very similar to that of the original error, presenting small offsets at the end of the postprandial period (approx. 5 mg/dL for the SEVEN® PLUS and 7 mg/dL for the Paradigm® Veo™). These offsets did not contribute negatively on the clinical accuracy of the simulated error, since the overall MARD was similar for the simulated dataset and the real data, especially for the SEVEN® PLUS. Indeed, simulated MARD for the Dexcom monitor was 18.72%, compared to the 17.28% of the real monitor. For the Paradigm® Veo™ the simulated MARD was 18.38%, being somewhat higher than the observed error from only 37 sensors, which was 14.98%.

In conclusion, both models were successfully validated reproducing the same statistical properties than the original data. Importantly, although the models have been tested only against postprandial data, additional inconveniences are not expected when analyzing data in the fasting state since postprandial control is considered by far the

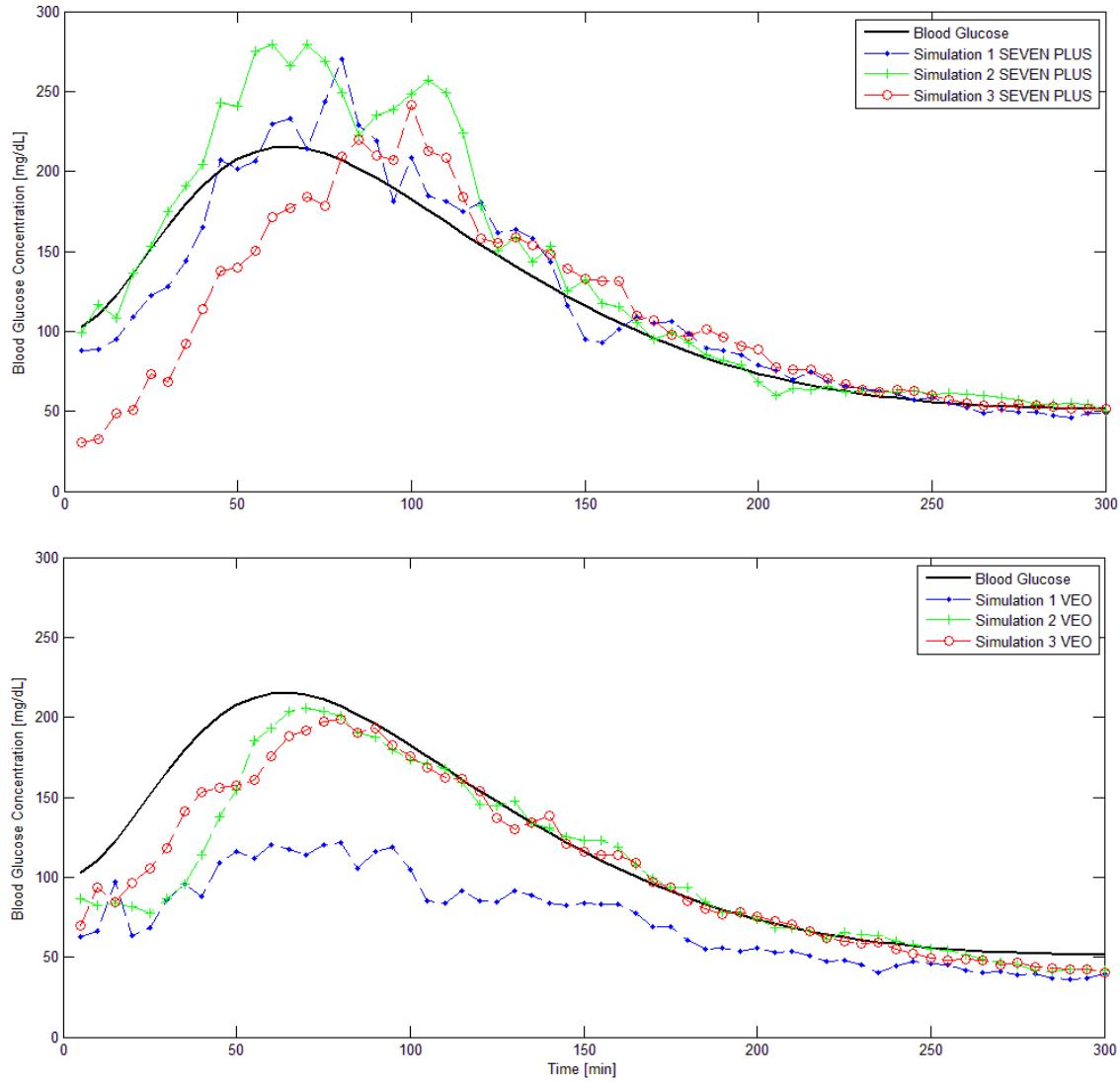


Figure 6.7: Three simulation of each monitor with the proposed models for the postprandial period of a virtual diabetic patient. Top panel shows the SEVEN® PLUS response while bottom panel corresponds to the Paradigm® Veo™ simulations.

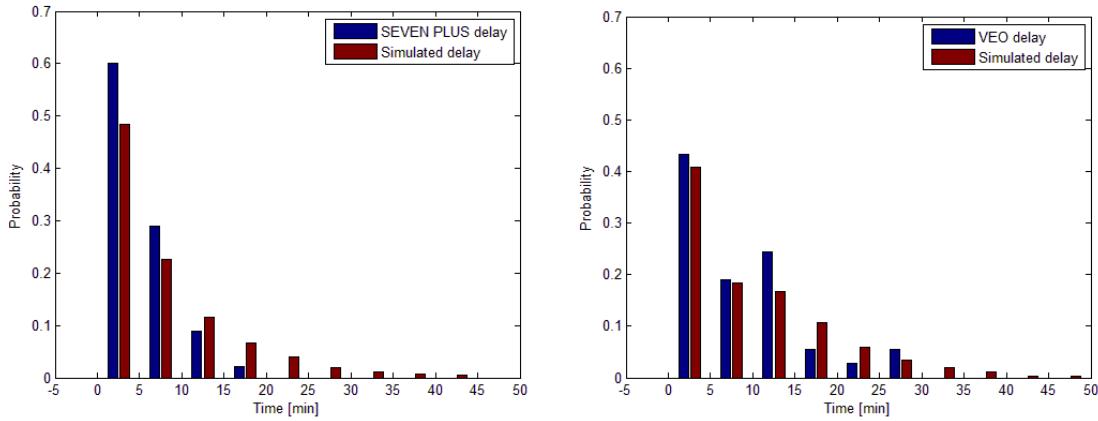


Figure 6.8: Real and simulated delay distribution for both monitors. The left panel shows the delay of the SEVEN® PLUS monitor, against the simulation distribution using the proposed model. In the right panel the distributions for the Paradigm® Veo™ is shown.

major challenge in developing the artificial pancreas. Meal intake was a big challenge for both devices, which were unable to follow the high rise of glucose. The high observed variability in both cases threatened postprandial performance. Compared to the SEVEN® PLUS device, the Paradigm® Veo™ device seemed to exhibit longer delays and higher probability of abnormal sensor behaviors in postprandial conditions. Certainly, further improvements in calibration algorithms are needed to reduce variability in CGM.

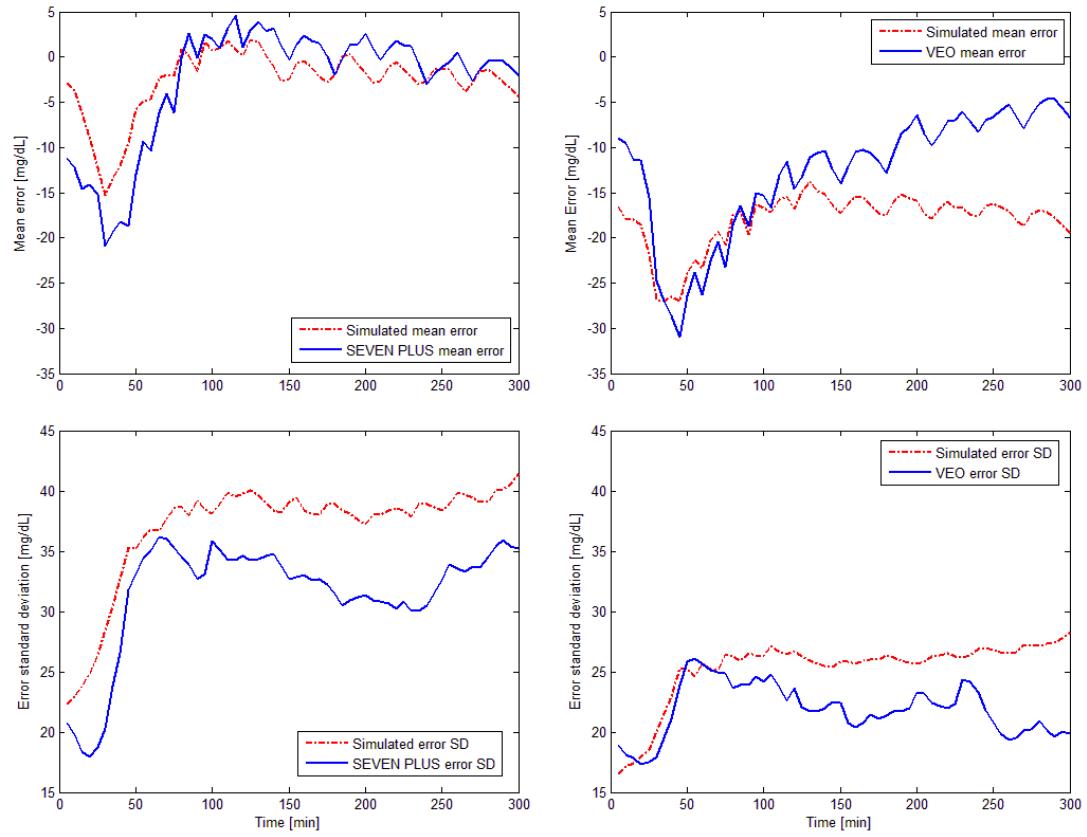


Figure 6.9: Top row shows time variation of the mean error for both monitors and for the simulated error. In the bottom row the standard deviation of the error and the simulation is plotted against time. Left column shows the SEVEN® PLUS data and model, and the right column illustrates the statistical behavior of the Paradigm® Veo™ and its representing model.

# Chapter 7

## Conclusions

This part of the thesis consisted of a preface of the identification problem in the diabetes context. Foundations for both virtual and experimental identification have been established, and the main contribution of this thesis is now at hand.

Many of the proposed models for diabetic patients suffer from low identifiability. A modification of a literature model was proposed in chapter 1.4, but the goal of this model was not to increase current identifiability of the diabetes models but to fit better the physiological properties of the endogenous glucose system of a diabetic patient. Identifiability can be improved in models which have identifiability issues by tuning the conditions of the experiment for better data acquisition properties.

In chapter 5 a classic optimal experiment design approach was followed for identifiability enhancement. To cope with the inherent problem of model dependance on the experiment results, the procedure was repeated on two different models: the Bergman minimal model and the modified Panunzi's model. The models chosen were amongst the simplest of all the models in literature in order to help in the computation process, which is extremely heavy for the optimal experiment methodology. From the outputs of the experiment design for both models qualitative results were extracted and merged into diabetes experiments guidelines for identifiability optimization.

Three days of monitoring were considered optimal for the current state of CGM. Monitoring sensors are often replaced every week, with recalibration occurring usually after sensor replacement. Considering three days of monitoring for identification, and three or four days for validation, the full lifespan of a sensor is used and more coherent

results are expected. Two different profiles were used for the identification days: Advance of the meal with respect to the insulin delivery, and delay of the meal. The meal was smaller when advanced and larger when delayed, for patient's safety. The insulin delivery before meal profile was found to enrich the data more than the other profile, being this case repeated twice in the three days experiment.

In the ambulatory environment, the experiments were slightly modified for easier application. The preprandial glucose level and trend were taken into consideration when deciding which of the profiles to apply. Two scenarios were considered:

- If preprandial glucose was greater than  $150mg/dL$  or between 100–150 and rising, the insulin bolus was delivered in advance, followed by a 100 CHO grams meal.
- If preprandial glucose was lesser than  $100mg/dL$  or between 100 – 150 and decreasing, the insulin bolus was delayed 2 hours to a meal of 40 CHO grams.

This protocol was used for gathering CGM data from 12 diabetic patients in real-life conditions in an experiment in the Clinical University Hospital of Valencia [74]. Data obtained from those 12 patients in an in-clinic experiment is used later on in this thesis for identification purposes. The same data was used for the development of a simulation model for CGM devices.

Two CGM devices were analysed and modelled in chapter 6. The model proposed is stochastic in nature, and can be used for simulation purposes and for testing controllers. Three different characteristics were used for the construction of the model:

- Delay of the CGM signal to blood glucose.
- Stationarity of the error of the CGM.
- Autocorrelation and probability distribution of the error.

Delay along the population was found to follow an exponential distribution. Contrary to the general opinion, the error signal of the CGM was found to be non-stationary, and as such much more difficult to model. Very simple correlations were found between the error signal and the blood glucose profile and its derivative. Quasi-stationarity was achieved when using those correlations to transform the error signal. Finally, simple

AR models and probability distributions were used for characterization of the stationary signal.

Identification trials on the diabetes context are historically found challenging. Virtual data identification, for example using virtual patient's data from any of the simulators described, is of great importance as an introduction to the real experiments. I believe the CGM model developed in this part of the thesis helps in the completion of virtual identifications for the diabetes context. Compared to other models in literature, the proposed model tackles with the stationarity issue in a way that no other has done to date. As for the delay and distribution characterization, the findings of this thesis are in accordance with the results of other publications.

## **Part III**

# **Interval identification**

# Chapter 8

## *In silico identification*

All the mathematical models for type 1 diabetes reviewed in this thesis suffer from problems that lead to not mimicking the real patient's physiology. The logic step to take before jumping into identification on real experimental data consists in identifying data that closely resemble the dynamics of the models used. Virtual patient's identification is then defined as an identification procedure analog to the experimental data fitting, where the virtual data is generated with the same model used for identification. Uncertainty can be then artificially added to the virtual patient's data to challenge the identification method. The aim of this experiment is to test the methodology of identification in the diabetes context.

Interval models were used for consideration of uncertainty within the diabetic patient, and for the identification fitting. Virtual patients were simulated with inherent variations in their physiologic parameters in order to simulate intrapatient variability. Interval bounding of all the possible outputs of the patients has to be performed for the complete characterization of the patient's data, but overestimation of the uncertainty has to be considered and penalized when using interval models for identification. From this point of view, we devised another point of view to interval identification. In this thesis, dual objective minimization is used to accomplish interval identification of diabetic patients. the two objectives to be optimized are: 1) the interval data bounding, 2) the glucose interval width. Both objectives are competing in the sense that maximizing the interval output of the model surely will bound more data samples, but the prediction capabilities of the model are worsened.

Multiobjective optimization was then applied to both simulated gold reference data and simulated CGM data of the same patient's, and the results compared.

## 8.1 Multiobjective Optimization Set-up

A cohort of 14 virtual patients was generated by means of the Cambridge's model. Three postprandial periods were simulated for each patient, according to the optimal experiment design described in chapter 5: on Day 1 and Day 3 a meal with 100 grams of carbohydrates (CHO) was ingested and the insulin bolus was advanced 30 minutes; on Day 2 the patient ate a 40-gram CHO meal and delayed the bolus for 120 minutes. For all simulated days the patient was considered at euglycemia before the meal intake (model initial conditions). This experimental set-up was shown to be beneficial for model identification due to the separation of insulin and meal dynamics.

The virtual patients were considered to have intra-patient variability (time-varying model parameters). Input errors were also considered for the insulin infusion rate and the estimation of carbohydrate intake. The parameters considered time-varying are listed next:

- $S_{iT}$ : insulin sensitivity on glucose transport from blood to interstitium.
- $S_{iD}$ : insulin sensitivity on glucose utilization.
- $S_{iE}$ : insulin sensitivity on endogenous glucose production.
- $k_e$ : insulin elimination rate.
- $k_{12}$ : rate of glucose transport from interstitium to blood.

The following parameters were treated as patient-dependent and time-invariant, due to constraints of the interval simulator:

- $t_{maxG}$ : time constant for glucose absorption in the gut.
- $t_{maxI}$ : time constant for insulin absorption.

Additionally, the following input errors were introduced:

- pump: a random time-varying error for the insulin infusion rate from the insulin pump.
- meal estimation: a repeated bias plus a random time-varying error for the carbohydrates estimation given by the patient.

Variability in the meal absorption is characterized by uncertainty in the meal estimation and variability in insulin pharmacokinetics is characterized by the parameter  $k_e$  and uncertainty in the insulin pump infusion. For the sake of simplicity, the time-varying parameters and errors considered were assumed constant throughout a postprandial period. However, they were changed from one day to another of monitoring following a random process with mean equal to the nominal value of the parameter ( $\theta$  for the errors) and a standard deviation of 10% of the nominal value. As demonstrated by Calm *et al.* [9] using optimal interval simulation methods, the consideration of 10% uncertainty in the model parameters may produce glucose trajectories differing in 100 mg/dL, so it is considered a sensible value for the reproduction of variability.

When interval parameters are considered, the simulation of the model must produce a glucose envelope representing the set of possible glucose trajectories that the patient may exhibit according to the defined intra-patient variability. To this end, the interval simulators for the Hovorka model developed in [8] and [19] were used in this work.

Measurement errors induced by a CGM device during home-monitoring were simulated. In this work the model presented in chapter 6 was used for simulation of the real-time CGM Dexcom<sup>®</sup> SEVEN<sup>®</sup> PLUS (Dexcom<sup>®</sup>, San Diego, CA), due to more accurate modeling results produced.

Classic parametric identification results in a single point in the parameter space, being insufficient for the case of a time-varying model based on poor prediction capabilities. Bounding all glucose measurements by means of an interval model, with time-varying parameters considered as intervals, may help identifying both the patient and the intra-patient variability, increasing the robustness of the derived therapies or control algorithms. However, in the case of home monitoring where large measurement errors may appear (especially due to CGM) some relaxation may be needed to avoid large intervals for parameters that loosely fit the physiological variability.

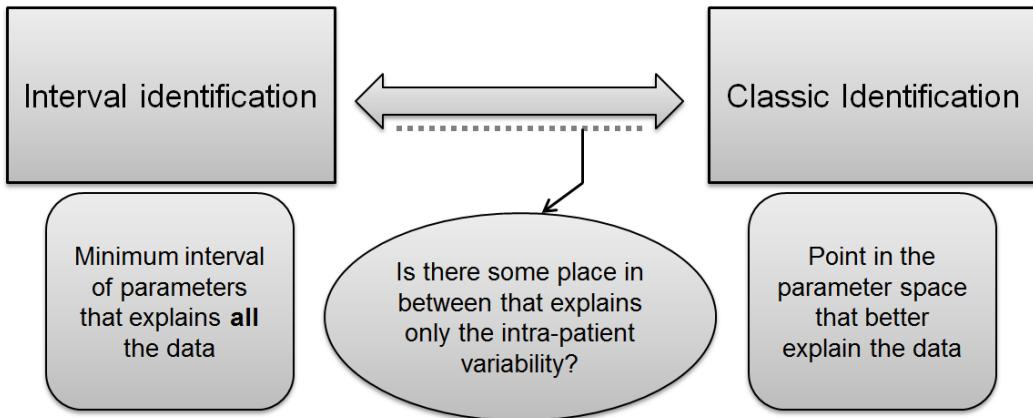


Figure 8.1: Ranging from classic to interval identification.

In Figure 8.1 a possibility of hybrid identification ranging from classic to interval is hinted. Given the fact that neither the classic identification point of view is valid for the current diabetic patient models (due to parameter variability), and also considering that the pure interval bounding is inefficient due to large measuring errors, a compromise is required for the appropriate characterization of real patient variability. The hypothetical trade-off point in between the two identification paradigms has never been investigated in literature (to my knowledge). Should this point exist, the compromise errors that are derived from this methodology will be characterized in the following lines.

Multi-objective identification is the process of simultaneously optimizing two or more conflicting objectives subject to certain constraints. The solution of this kind of problems is not a single optimal point in the objectives search space, but a family of solutions called a Pareto Front (PF). Each individual in the family is non-dominated by the other individuals, i.e. they cannot be replaced by other point in the objective space for improving an objective, without worsening another one. Evolutionary algorithms are popular solvers for this kind of problems. The  $\epsilon$ -MOGA evolutionary algorithm [31] was used in this work.

Optimization requires the objectives to be defined as minimization indexes. The first objective considered for minimization is the squared  $z$ -norm of the fitting error vector  $e := (e_1, \dots, e_n)$ , defined as:

$$J_1 := \|e\| = \sum_{i=1}^n e_i^2, \quad e_i := \begin{cases} 0 & \text{if } G_i \in [\hat{G}_i, \bar{G}_i] \\ G_i - \bar{G}_i & \text{if } G_i > \hat{G}_i \\ \hat{G}_i - G_i & \text{if } G_i < \hat{G}_i \end{cases} \quad (8.1)$$

where  $i$  denotes the sample index,  $n$  is the number of measurements,  $G_i$  the  $i$ -th glucose measurement and  $[\hat{G}_i, \bar{G}_i]$  the glucose interval predicted for time sample  $i$  given by the interval model. This is computed by means of a guaranteed interval simulator as proposed in [8]. Thus, if  $|\bar{G}_i - \hat{G}_i| = 0$  (null interval width) then  $J_1$  corresponds to the classic sum-of-squares cost index. If  $J_1 = 0$ , the envelope given by  $[\hat{G}_i, \bar{G}_i]$ ,  $i = 1 \dots n$ , encloses all the measurements.

The other objective, by definition of multi-objective optimization, must be conflicting with the first one in the minimization. Given that when working with intervals the error is minimized by simply increasing the size of the interval (trivial solution at maximum width), the other objective to be pursued must be a minimization of the interval width. The following index is then considered:

$$J_2 = \max_{1=1 \dots n} |\bar{G}_i - \hat{G}_i| \quad (8.2)$$

Maximum width was chosen for the minimization because variability in the width of an interval model response can be very high in the postprandial period. Minimizing average widths for example may lead to very wide glucose intervals in the early postprandial period which translate in wide interval parameters. Limiting the maximum width instead bounds the interval more consistently throughout all the postprandial period.

Both objectives can be shown in a cartesian plot, in order to visually characterize the PF. An example of a PF is shown in Figure 8.2. The extremes of the PF correspond to the classic (right) and interval approach (left). In between, every point in the PF is a relaxation of the interval approach.

The selection of the “best” solution in the PF will depend on the nature of the measurements and the desired degree of robustness. When using reliable measurements

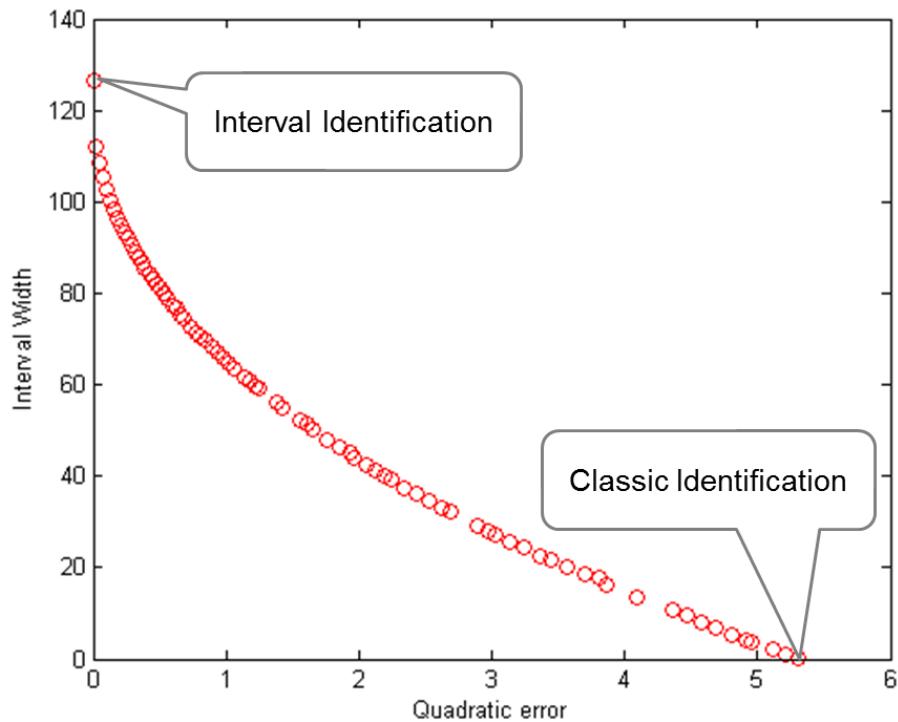


Figure 8.2: Example of a Pareto Front exploring all optimal possibilities from a zero-width problem (classic identification) to a minimum-width problem with zero error (interval approach).

(e.g. reference BG measured with a YSI 2300 STAT Plus™ Glucose Analyzer), and assuming there is no model mismatch, it is desired that all the data points are predicted by the glucose envelope generated by the model, i.e.  $J_1 = 0$ . This is equivalent to choosing the pure interval solution of the problem.

As already said, identifying with CGM data is much more challenging due to the accuracy limitation of these devices, but home monitoring is undoubtedly more feasible for clinical practice. The optimization method works the same way independently of the data source, but choosing one point or another in the PF does not. The zero-error solution only makes sense when identifying from reference BG since it encloses robustly all the data available, which are reliable. When CGM is used, data cannot be trusted as much as reference BG, so a tradeoff solution has to be chosen.

A relaxation is thus needed in order to identify an interval model representing actual patient's variability. Two solutions can be devised:

- Choosing an “acceptable” error, and finding the closest PF individual for that error.
- Estimating an appropriate maximum width for the glucose envelope, and choosing a PF individual accordingly.

The second approach was used here. Indeed, the optimal width of the pure interval solution for the BG-based identification may be a good approximation of intra-patient variability. Thus, a method is needed to estimate this optimal width from CGM data. This is explored further in later sections of this chapter.

The identification algorithm considered that every uncertain parameter was characterized by its lower and upper bound. The final parameter identification vector results:

$$\theta = [\underline{S}_{iT}, \overline{S}_{iT}, \underline{S}_{iD}, \overline{S}_{iD}, \underline{S}_{iE}, \overline{S}_{iE}, \underline{k}_{12}, \overline{k}_{12}, \underline{k}_e, \overline{k}_e, \underline{pump}, \overline{pump}, \underline{meal}, \overline{meal}, t_{maxI}, t_{maxG}] \quad (8.3)$$

where  $\underline{S}_{iT}$ ,  $\overline{S}_{iT}$  are the upper and lower bounds of the parameter  $S_{iT}$ . Similar notation is used for the rest of the interval parameters.  $pump$  and  $meal$  stand for errors in pump infusion and meal estimation respectively.  $t_{maxI}$  and  $t_{maxG}$  are the only real-valued parameters.

Even though identification in a domiciliary environment can only be performed reasonably with a CGM, reference BG is the variable that the model attempts to predict, and as such it is the signal to be used for the validation of the identifications performed. In order to calculate the prediction abilities of the identified model, the Mean Absolute Relative Error referred to a glucose envelope ( $MARD_I$ ) was used:

$$MARD_I[\%] := \frac{\sum_{i=1}^n \left| \frac{d_H(BG_i, [\hat{G}_i, \bar{G}_i])}{BG_i} \right| * 100}{n} \quad (8.4)$$

$$d_H(BG_i, [\hat{G}_i, \bar{G}_i]) := \begin{cases} 0 & \text{if } BG_i \in [\hat{G}_i, \bar{G}_i] \\ BG_i - \bar{G}_i & \text{if } BG_i > \bar{G}_i \\ \hat{G}_i - BG_i & \text{if } BG_i < \hat{G}_i \end{cases} \quad (8.5)$$

where  $BG_i$  is the reference BG measurement at time  $i$ ,  $n$  is the number of data points, and  $d_H$  is the Hausdorff distance from the point  $BG_i$  to the interval  $[\underline{\hat{G}}_i, \bar{\hat{G}}_i]$  representing the glucose envelope at time  $i$  given by the interval model.

All statistical significances in this chapter were calculated using Student's t-test for the correlation values.

## 8.2 Identification from reference glucose

An example of identification for one representative patient is shown in Figure 8.3. This patient is an illustrative example, but little differences are found from one patient to another. For this patient two simulations, corresponding to both ends of the PF in Figure 8.2, plus the BG data are plotted. As expected, the output of the interval model robustly encloses all the simulated days.

As for the rest of the patients, the identified interval models covered completely the data for all the postprandial periods. As expected, the method does not encounter any difficulty on reproducing noise-free data. However, identification from BG reference data is very limited due to the requirement of in-patient studies. Further investigation is required on the presence of uncertainty in the measurements introduced by CGM.

The pure interval extreme of the PF is in the following considered the real variability of each patient, and the width associated (error is zero for all cases) is considered the optimal width to be identified in each case.

## 8.3 Identification from CGM

Figure 8.4 shows the BG-based and CGM-based Pareto Fronts for the sample patient. For the same error, all the models identified from the CGM have wider intervals than the ones from reference BG. This is also true in the rest of the patients, especially as it gets closer to the pure interval solution (zero error).

The top left solution in the PF for CGM identification has a much larger width than any other individual in the solution. Such a large width is necessary to cover all the

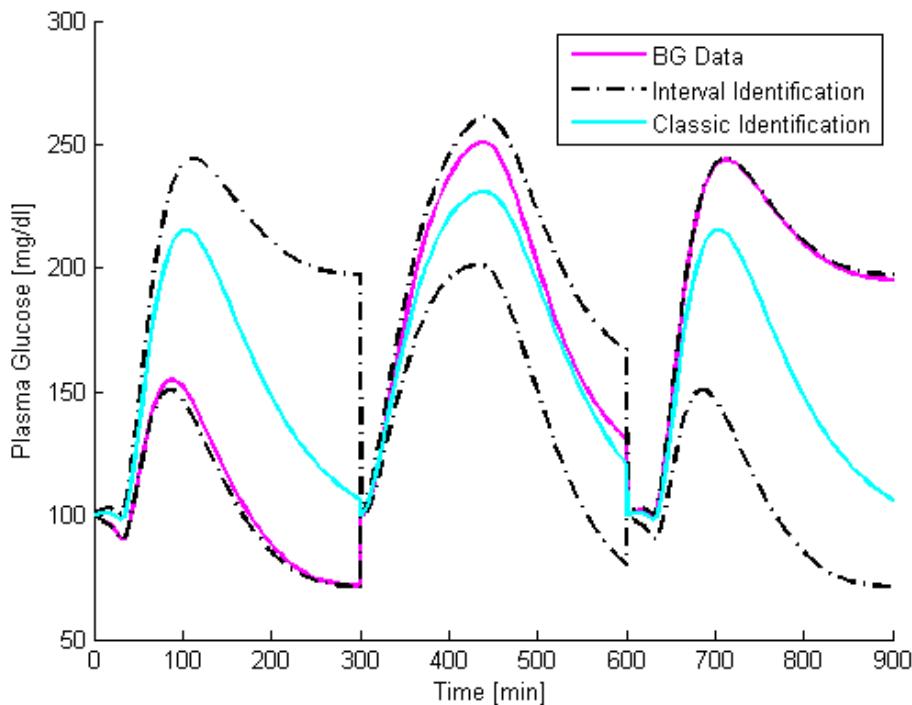


Figure 8.3: Comparison of classic and interval identification. Three 5-hour postprandial periods in consecutive days are shown. The first and third days correspond to identical meals and insulin infusion. Black dotted lines represent the interval solution, enclosing all the monitoring days. The cyan line (classic identification) behaves as an average glucose profile for the three days.

CGM excursions especially due to the error present in the first time instants of every postprandial period, where the glucose envelope given by the model is very thin, but the CGM measurements may be very noisy. From a practical point of view this solution can be discarded.

The CGM-based identification results for both ends of the PF are shown in Figure 8.5. As with the glucose reference fitting, classic identification shows only an average behavior. The interval model covers all the data provided by the CGM but patient's actual intra-patient variability is overestimated since it is confounded by CGM measurement error.

As stated earlier, a method for estimation of the “optimal width” (width for pure

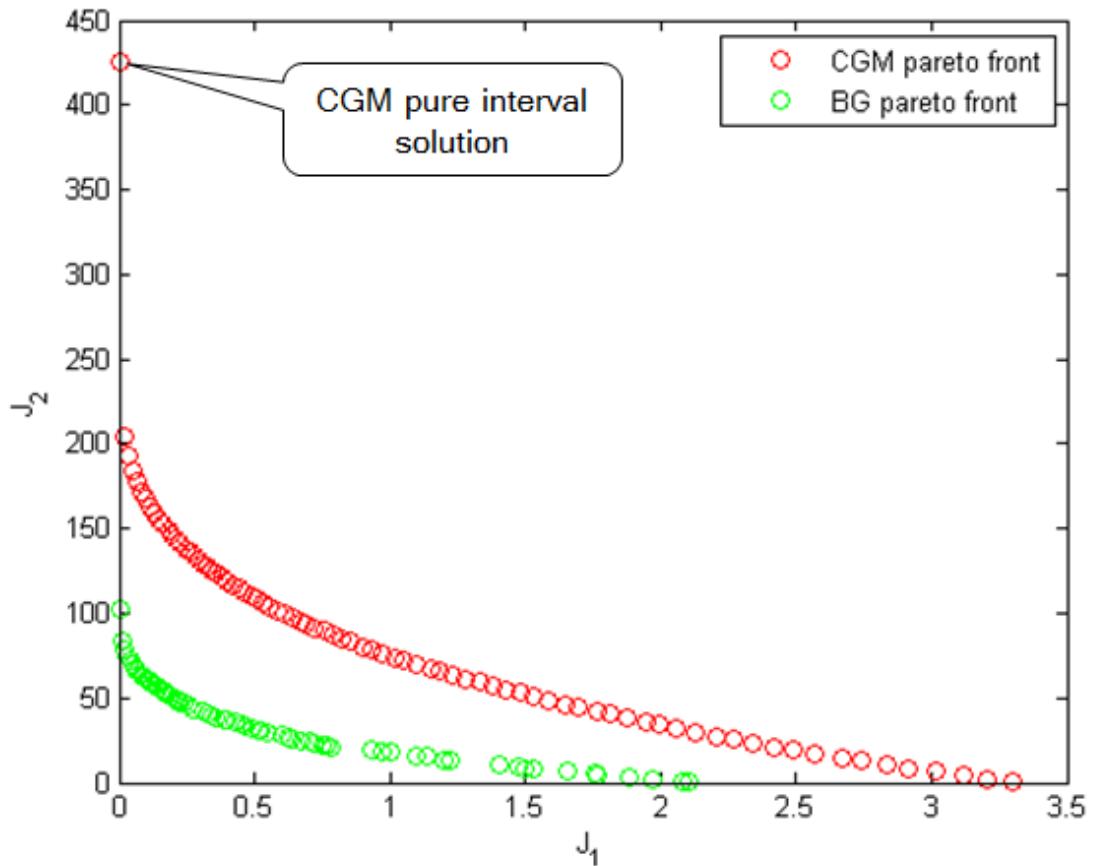


Figure 8.4: Both PF for the identification of a patient, using CGM and BG data. CGM identification shows larger width and error for most patients. CGM pure interval identification requires too wide prediction envelopes.

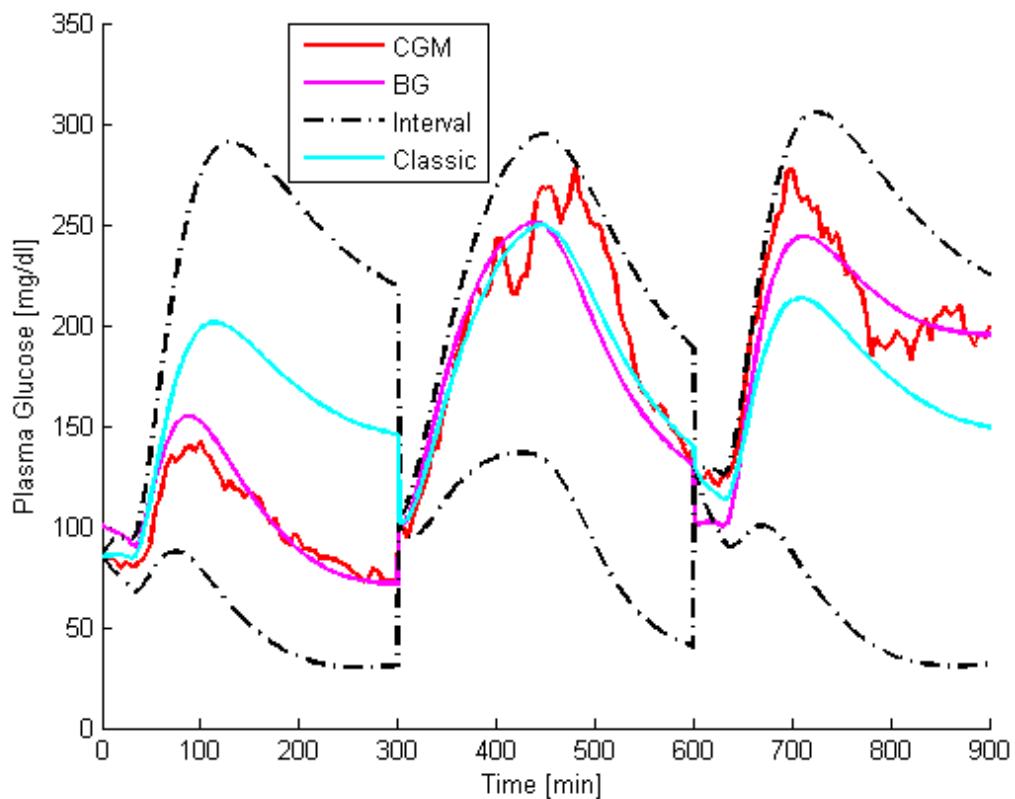


Figure 8.5: Comparison of classic and interval identification from CGM. Black dotted lines are too wide to be considered useful. Three 5-hour postprandial periods in consecutive days are shown. The first and third days correspond to the same meal and insulin infusion to the patient.

interval identification on BG data) for each patient is needed. Correlation between the optimal width and the so-called “experimental width” obtained from the overlapping of responses for similar days was investigated. Experimental data was used to further understand correlations between BG and CGM widths in real patients. The possibility of estimating the BG-based experimental width from the CGM signal, and thus the optimal width for the identification, was investigated with the analysis of clinical data from 12 patients who underwent clinical mixed meal studies as described in detail in [74]. Postprandial variability was analyzed after two predetermined meals containing 40 g or 100 g of carbohydrates with normalized initial conditions through an insulin feedback procedure. Each patient was evaluated in 4 different occasions, twice with each meal. Frequent YSI measurements were taken simultaneously to a Dexcom® SEVEN® PLUS monitor. The BG-based optimal maximum width is highly correlated to the maximum width observed by overlapping the BG signal (experimental width) for days 1 and 3 where the patient ate the same meal ( $r=0.848, p<0.005$ ). This is also true for the average width for all the days of the postprandial period ( $r=0.797, p<0.005$ ).

In practice neither the optimal width nor the possibility of overlapping several BG postprandial periods is available. To cope with this, experimental correlations between BG and CGM width were investigated using the experimental dataset introduced before. Table 8.1 shows the results obtained demonstrating the existence of high correlation between the CGM-based and the BG-based width computations from the experimental cohort. Table 8.1 also shows the correlation among the BG-based and CGM-based widths for an *in silico* study with the virtual patients’ cohort reproducing the clinical study design in [74]. Correlations are very similar, reinforcing the CGM model used and the findings of chapter 6.

		CGM	
		Max width	Avg width
BG	Max width	0.67 (0.72)	0.75 (0.66)
	Avg width	0.79 (0.87)	0.86 (0.88)

Table 8.1: Correlations of postprandial BG-based vs. CGM-based widths for the experimental patients’ cohort. The *in silico* counterparts of the same correlations are shown within parenthesis.

Table 8.2 shows the correlations of the BG-based optimal width (both maximum and average) with the experimental width from the CGM. These correlations were found on the virtual patients’ cohort. Out of the three days of monitoring, maximum and average widths of CGM overlapping signals are studied as estimators for the optimal width

found in the identification from reference glucose. Max3 and Avg3 are the maximum and average width from the three days of the experiment, respectively. Max2 and Avg2 correspond to the maximum and average width from 2 equal days (days 1 and 3), respectively.

	CGM experimental width			
	Max3	Avg3	Max2	Avg2
Optimal maximum width	0.565	0.781	0.775	<b>0.832</b>
Optimal average width	0.588	<b>0.8</b>	0.707	0.78

Table 8.2: Correlations of optimal widths in the virtual patients vs. CGM model-based experimental widths. The best estimators (best correlations) of optimal widths are highlighted.

All correlations shown in Table 8.2 are significant ( $p < 0.005$ ). The best correlations are obtained for the average CGM experimental width, either for 3 or 2 days. This better correlation is due to the influence of the averaging computation in the CGM noise, which is alleviated throughout the postprandial period. The corresponding regression lines to both estimators corresponding to the average of CGM signals are:

$$OptimalMaxWidth = 0.6617 * Avg2 + 60.4569 \quad (8.6)$$

$$OptimalAvgWidth = 0.3979 * Avg3 + 27.2764 \quad (8.7)$$

Finally, equation 8.6 or 8.7 will define the individual in the CGM-based PF to be chosen as optimal identification. This is equivalent to launching a constrained optimization problem with  $J_1$  as cost index and fixed glucose envelope maximum width given by 8.6 or fixed average width given by 8.7. This sort of optimization should then include non-linear constraints on the output of the optimization model, which may be difficult depending on the optimization algorithm.

As an illustration, the identification results for the sample patient are shown in Figure 8.6, where Avg2 and Avg3 estimators are compared. The use of CGM experimental widths over a number of days for predicting the “optimal” width of a particular patient is then justified. However, the predictive capability of the identified model is still to be checked.

Results of the BG prediction errors are shown in Table 8.3 for the virtual patients’ cohort. The mean error for the samples not enclosed by the glucose envelope ( $d_H > 0$ )

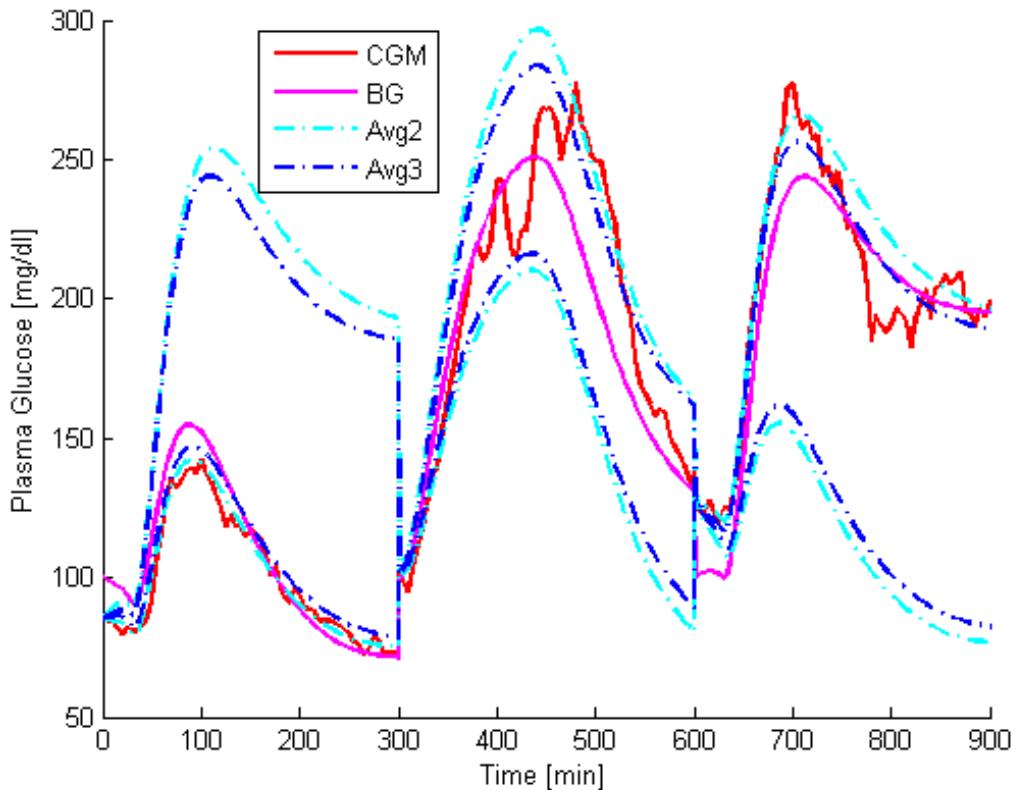


Figure 8.6: Comparison af Avg2 and Avg3 estimators for the optimal width in the identification from CGM data. There is little difference between estimators, but there is a big difference between the use of optimal width estimation and pure interval identification from CGM data (see figure 8.5). Three 5-hour postprandial periods in consecutive days are shown. The first and third days correspond to the same meal and insulin infusion to the patient.

is also computed for comparison purposes, and denoted as  $ErrOut$ . In order, the table shows the median for the population of the estimated width, the number of samples enclosed in the glucose envelope, the relative error considering only the samples outside the glucose envelope, and the  $MARD_I$  (%). The two width estimators that showed better correlations, as highlighted in Table 8.2, were tested.

	Avg2			
	$MaxWidth[mg/dL]$	$Predicted[\%]$	$ErrOut[\%]$	$MARD_I[\%]$
Median	83.3	67.8	7.5	3.3
Min/Max	67.6/163.3	34.3/97.9	1.9/38.3	0.1/11.9
Avg3				
Median	61.4	52.8	8.3	4.6
Min/Max	42.4/96.4	35/97.9	2.4/30.9	0.1/15.6

Table 8.3: Widths and prediction errors for the selected (best) optimal width estimators. Top row shows the performance of the identification if the maximum glucose envelope width for each patient is predicted from the width of two identical days. Bottom row displays the case when the average glucose envelope width for each patient is estimated from the width of the 3-day CGM registries.

A long-term simulation of a 15-day monitoring period for a sample patient was analyzed in order to test the method’s strength when coping with longer monitoring periods and outliers. The same procedure described before for virtual patient data generation was used for the generation of this dataset. In this case a total of 7 meals with 40-gram CHO content and 8 meals with 100-gram CHO content were given as input to the model. Variability of the uncertain parameters remained at 10% of the nominal value of the parameters.

Visual representation of four out of the 15 simulated days are shown in Figure 8.7. Optimal width (computed from BG-based model identification) is compared against Avg3 width estimator, consisting in equation 8.7 applied to the average width of the 15-day simulation. The four days selected represent “worst-case” results and illustrate the performance of the method against large CGM measurement error (days 1, 3 and 15) and outlier patient behavior (day 11). For the rest of days similar results than those shown in Figure 8.6 were obtained.

Reviewing the characteristics of the outlier day from the 15-day simulation, a special combination of glucose lowering power is hinted, which seems not to be present in any other day of the experiment. To check this, Figure 8.8 shows the variation of the most

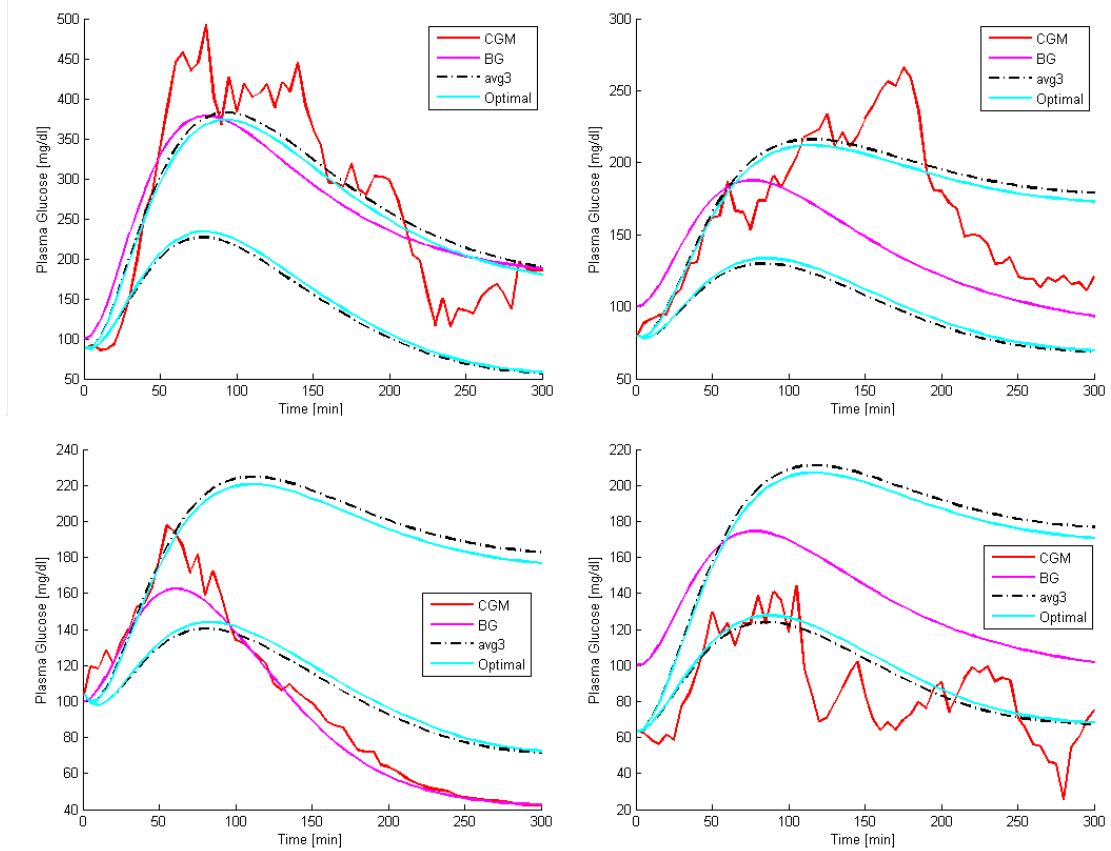


Figure 8.7: Comparison of Avg3 width estimator and the optimal iwdth for the 15-day identification experiment. Top right (Day 1), top left (Day 3) and bottom right (Day 15) panels show cases of CGM monitoring errors with satisfactory BG prediction from the model. The bottom left panel (Day 11) shows the only case where BG was not enclosed by the predicted glucose envelope for most part of the postprandial period due to the outlier patient behavior (extreme parameter values).

relevant parameters for each simulated day in the 15-day experiment. Only the insulin-related parameters are displayed for the sake of clarity even though parameters were subject to variability as stated earlier. In the figure, variation is calculated as relative to the nominal parameter, and it is considered positive if it induces a decrement of glucose concentration (e.g. larger insulin sensitivities), and negative if it causes glucose to increase (e.g. larger insulin elimination rate). Displaying the parameters deviation in a stacked form as in Figure 8.8 provides helpful information on the final influence of variability on the model's output for a particular day. However, variation magnitude does not directly correlate to the magnitude of the effect since not all the model parameters have the same sensitivity.

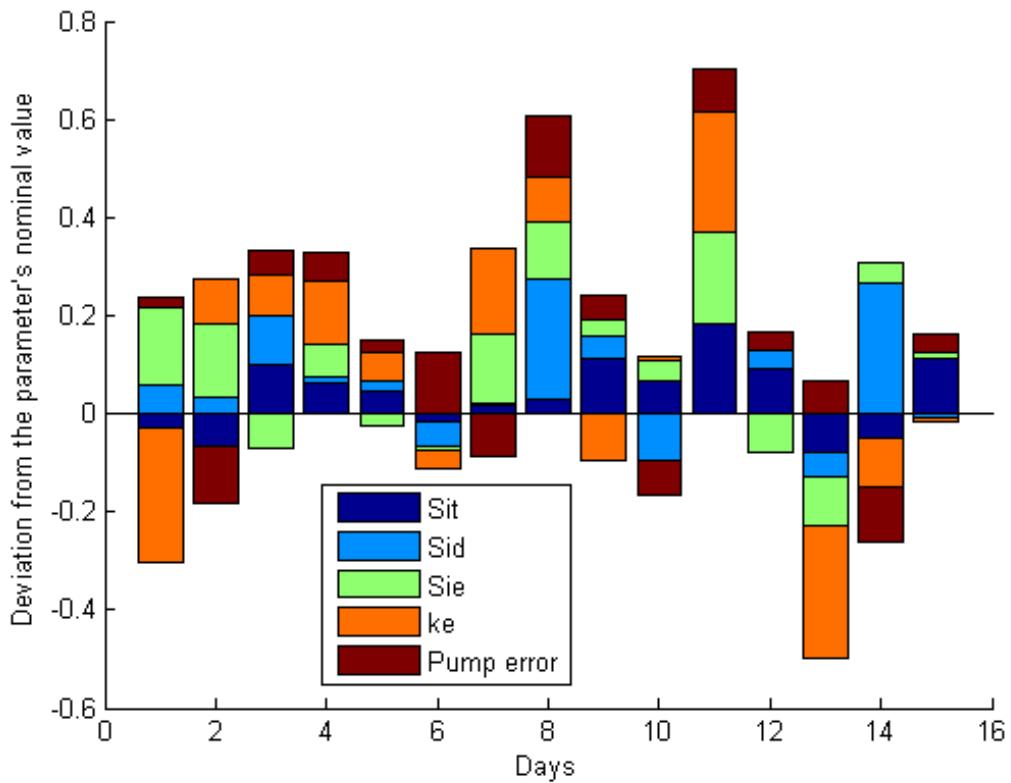


Figure 8.8: Depiction of the parameter variation within the 15-days experiment. For the sake of simplicity, only the insulin-related parameters are shown here. Parameter deviation is positive if it produces a decrement of glucose levels and negative otherwise.

## 8.4 Discussion

Identification from virtual data supposes a much lighter challenge than experimental identification, but very interesting findings are extracted from this chapter, considered as the basic previous work before encountering the problems of experimental data management.

The identification from reference BG data was successful, as illustrated in Figure 8.3. The identified interval model was able to capture patient's variability with minimum width when no measurement error is present in the data. Day 1 and day 3 correspond to the same meal, but patient's behavior is very different, as typically observed in the clinical experience. Nonetheless, the model was able to capture this variability. It is understood that the use of interval models may help in deriving more robust glycemic control strategies.

Width estimation prior identification is required for CGM-based identifications. Very strong correlations were found between the experimental widths from reference glucose and CGM data for real type 1 diabetic patients. Therefor the estimation of the optimal width from CGM measurements seems feasible. The maximum correlation was obtained among average widths, as illustrated in Table 8.1. Very similar correlations were calculated for the virtual patient cohort, encouraging the use of the CGM model proposed. As Table 8.2 shows, the average widths of the overlapped patient's CGM postprandial traces were better correlated and thus they were better estimators of the optimal width of each patient. This may be due to the fact that averaging CGM widths can filter outliers in the CGM signal width, while maximum width values do not. Noisy outliers in the CGM signal width do not correlate well with BG widths, which are considered to be noise-free. Note that Avg2 computes the average from two identical days while the optimal average width corresponds to the whole identification period so they are not directly comparable. Finally, the optimal width estimation was computed from virtual CGM data from successive days using linear regression models.

Overlapping of two identical days (Max2 and Avg2) yielded much better estimations of the patient's optimal width, as data in Table 8.2 shows. This can be explained by the fact that the glucose variability observed in two days with the same meal and insulin therapy is a more accurate description of the patient's physiological variability, since there is no uncertainty in the model inputs. In future data acquisition experiments, repeating of identical days should be a priority. In reality though, obtaining two identical monitoring days for a patient can be very difficult. So, two different optimal width estimation

methods were provided giving more flexibility independently of the data available.

The validation of the model and its prediction ability was satisfactory, as shown in Table 8.3. The model was able to predict about 60% of the reference samples in the postprandial period, with a slightly better performance of Avg2 estimator. The non-predicted samples showed relative errors below 10% in both cases. Regarding  $MARD_I$  an error below 5% was obtained. Thus, the proposed method allows identifying satisfactorily interval models characterizing intra-patient variability from CGM data.

Regarding the 15-day simulation study, all the simulated days were well predicted within the model's boundaries except day 11 (see Figure 8.7), despite the error introduced by considering the CGM measurement as initial condition at each postprandial period. Day 11 gathers several factors that make it very difficult to predict. As shown in Figure 8.8, at day 11 all the parameters pulling glucose concentration down to hypoglycemic range concur with maximum values, i.e., it can be considered an outlier behavior. This is likely to happen in rather large experiments, where extreme variability at particular days may be found. Unfortunately, outlier behavior is unpredictable, and cannot be separated in the identification process from CGM malfunctions, which are much more common.

As shown in Figure 8.7, days 1, 3 and 15 are a good representation of the ability of the model to filter CGM malfunctions. On days 1 and 3 the monitor glucose readings were larger than reference glucose, at the risk of overestimating the glucose envelope width during the identification process. However, optimal width estimators are able to cope with this problem successfully predicting the actual BG values without width overestimation. Similarly, BG at day 15 was greatly underestimated by the CGM for the whole postprandial period. Nonetheless, this did not affect the predicted glucose envelope, which successfully enclosed BG except for the effect of the error in the initial conditions. It is deduced that the method is robust when confronted to outliers. Good prediction in most days of monitoring is achieved even in presence of extreme patient variability (day 11) or CGM malfunctions (days 1 and 15).

Ultimately, the selection of the PF individual in the identification process will determine the degree of filtering of outlier behaviors and CGM malfunctions. Accurate prediction of outlier behaviors will imply larger glucose envelope widths in detriment of filtering of CGM malfunctions and general performance. Thus a compromise solution is required. An alarm system can be devised when the CGM signal runs out of the predicted envelope dangerously towards hypoglycemia requiring a confirmation capillary glucose measurement. Such alarm system may prevent undesired hypoglycemic events and it may enhance CGM accuracy if the capillary measurements are used to recalibrate the

monitor.

Apart from the validation of the interval models, I believe that Pareto Fronts can provide insightful information for model identification under variability, both in an engineering environment as well as in the clinic. Physicians can use width index in the PF as a direct estimation of patient variabilitym which currently is calculated heuristically

# Chapter 9

## Pilot Experimental identification

Feasibility of uncertainty identification from virtual data has already been established, but validation on clinical data is required for the implementation of model-based controllers, which depend on the accuracy of the individual model obtained for each patient. However, model individualization has been proven difficult for data-based models [77, 64, 25] or physiology-based models [62]. Furthermore, suitability of classical metrics such as Mean Square Error for model evaluation has recently been questioned in the context of diabetes [20] where clinical implications of prediction errors must be considered.

Genetic algorithms for multiobjective optimization have been proven useful, but the computation times of this sort of optimizations are very long. In the virtual identification paradigm described in chapter 8, where only one optimization was needed for each patient the use of multiobjective optimization was affordable and given that these techniques provide much more information of the optimization indexes than a single objective optimization, its use was encouraged. In this chapter however, several optimizations are needed for each patient in order to complete a full cross-validation study. This type of identification setup is too computation demanding for the multiobjective method with so far, and a composite single index optimization is explored.

There are two main barriers to individual patient's model identification: error/noise sources in the measurement devices (especially with the use of CGM devices for ambulatory data acquisition), and uncertainty/variability in the patient's behavior because of circadian rhythms and other non-modeled dynamics, such as alterations in the

endocrine system, changes in daily life, stress, illness, etc. In this chapter, the feasibility of interval model identification for the characterization of intra-patient variability is investigated in clinical data. The minimization of a composite cost index comprising: (1) the glucose envelope width predicted by the interval model, and (2) a Hausdorff-distance-based prediction error with respect to the envelope, is proposed. A cross-validation study is performed for the evaluation of the method using clinical data from 12 patients with type 1 diabetes who underwent four in-clinic mixed meal tests with standardized initial conditions. CGM uncertainty is ignored as this work is considered to be a pilot identification study, hence the complexity of the experiment is reduced to a minimum. CGM influence on the identification from clinical data is reviewed later on this thesis.

## 9.1 Optimization and index definition

The dataset used for identification has been quickly introduced before in this thesis, and further details can be found in [74]. This data consisted in four monitored postprandial periods for twelve subjects with type 1 diabetes under CSII. On two occasions the patients received a mixed meal containing 40 g of CHO. On the other two occasions they ate a meal with the same relative macronutrients composition but with greater CHO content (100 g). For each meal, either a standard bolus or a computer-generated bolus-basal combination was administered following randomization. Pre-prandial plasma glucose was set around 100 mg/dL by means of a manual feedback intravenous insulin infusion. Hypoglycemia was avoided by using intravenous glucose infusion in case the patient's glucose levels were decreasing rapidly towards hypoglycemic levels. Plasma glucose was measured for 5 hours after the meal, every 5 minutes the first two hours after the meal and every 10 minutes afterwards, using a reference method (YSI 2300 STAT Plus Glucose analyzer, Yellow Springs Instruments, Ohio, USA). Plasma insulin was also measured periodically (every 15 minutes the first two hours, and every 30 minutes afterwards) along all the duration of the experiment.

Due to the different sampling periods of the measurements, cubic spline interpolation was applied in order to get sample-per-minute data on all variables. Due to the high accuracy of YSI measurements [60] uncertainty modeling effort can be focused only in model inaccuracies and within-patient variability.

Identification of interval models has been traditionally addressed under the framework of bounded-error identification (see chapter 3.2.1 and references therein).

However, when large intra-patient variability is present no parameter values consistent with the error will generally be found. Robust predictions for therapeutical decisions can be achieved if the interval model is able to bound the patient's response, i.e., the experimental measurements should be included in the output envelope predicted by the model at each time instant  $i$

$$y^*(t_i) \in y(t_i; \mathbb{P}), \quad \forall i \in I \quad (9.1)$$

where  $y(t_i; \mathbb{P}) = [\underline{y}(t_i; \mathbb{P}), \bar{y}(t_i; \mathbb{P})]$  stands for the interval prediction at time instant  $i$  for the to-be-identified parameter set  $\mathbb{P}$  and  $I = \{1, \dots, n\}$  is the index set of the available measurements. In practice, a relaxation of the above problem may be needed, allowing for small errors with respect to the inclusion envelope due to noise in the measurements and compensation for non-modeled dynamics.

As in the previous chapter, the interval version of the Cambridge model [8, 19] was used. They provide a tight guaranteed envelope  $[\underline{G}(t; \mathbb{P}), \bar{G}(t; \mathbb{P})]$  for uncertain model inputs and model parameters  $\mathbb{P}$ . Plasma insulin was available for all the experiments in the dataset, so the insulin absorption model was not used in this identification in order to keep the pilot study simple. This reduces the identification and simulation of results to only 2 sub-models: the Cambridge endogenous model and the Cambridge glucose absorption model.

Interval parameter values  $\mathbb{P}$  must be found so that the model output envelope bounds the measurements. Denoting as  $G^*(t_i)$  the glucose measurement at time instant  $i$ , equation 9.1 can be rephrased as:

$$G^*(t_i) \in [\underline{G}(t; \mathbb{P}), \bar{G}(t; \mathbb{P})], \quad \forall i \in I \quad (9.2)$$

Small errors with respect to the inclusion envelope will be allowed due to noise in the measurements and compensation for non-modeled dynamics, as already stated. Thus, in this chapter, optimization of a composite index is proposed for model identification. It comprises two components:

- Envelope width - The glucose envelope width must be minimized in order to avoid trivial solutions with large predictions bands. It is computed as the maximum difference throughout the postprandial period of the upper and the lower bound of

the interval model evaluation at time  $t_i$ :

$$J_w(\mathbb{P}) := \max_{i \in I} (\text{width}([\underline{G}(t; \mathbb{P}), \bar{G}(t; \mathbb{P})])) = \max_{i \in I} (\bar{G}(t; \mathbb{P}) - \underline{G}(t; \mathbb{P})) \quad (9.3)$$

- Prediction error - The number of measurements outside the predicted glucose envelope must be minimized and so has to be the error with respect to it. The error was computed here as the sum of squares of the Hausdorff distance  $d_H$  between the samples and the predicted envelope:

$$J_e(\mathbb{P}) := \sum_{i=1}^n d_H^2(G^*(t_i), [\underline{G}(t; \mathbb{P}), \bar{G}(t; \mathbb{P})]) \quad (9.4)$$

$$d_H(G^*(t_i), [\underline{G}(t; \mathbb{P}), \bar{G}(t; \mathbb{P})]) := \begin{cases} 0 & \text{if } G^*(t_i) \in [\underline{G}(t; \mathbb{P}), \bar{G}(t; \mathbb{P})] \\ G^*(t_i) - \bar{G}(t; \mathbb{P}) & \text{if } G^*(t_i) > \bar{G}(t; \mathbb{P}) \\ \underline{G}(t; \mathbb{P}) - G^*(t_i) & \text{if } G^*(t_i) < \bar{G}(t; \mathbb{P}) \end{cases} \quad (9.5)$$

The cost index to be minimized is thus defined as:

$$J_{we}(\mathbb{P}) := J_w(\mathbb{P}) + \gamma \cdot J_e(\mathbb{P}) \quad (9.6)$$

where  $\gamma$  is a weighting factor between the minimization of the envelope width and the fitting error. A very small value for  $\gamma$  yields to very small intervals for the identified model parameters  $\mathbb{P}$  with loose data fitting, while large values of  $\gamma$  provide good coverage of the data with large intervals for  $\mathbb{P}$ . The weight  $\gamma$  has to be tuned a priori and it will define the degree of relaxation given to the optimization problem. In this chapter, a battery of pilot identifications was run using several different weighting factors  $\gamma$ . Then, the authors decided to set  $\gamma=100$  considering that the fitting results displayed a good compromise between the envelope width and data compliance.

Minimization of the cost index was performed with the global optimization algorithm Evolution Strategy with Covariance Matrix Adaptation (CMAES) (chapter 3.3.2). The optimization to be performed is a non-linear single objective minimization with linear restrictions in the parameters (intervals). CMAES performs very fast optimizations on a single objective with any number of optimization parameters. Unfortunately, the build released by Hansen et al. [29] does not implicitly consider restrictions neither in the

outputs nor in the inputs, so they have to be integrated in the cost index. Interval parameters consist of two independent parameters, upper and lower bounds, where obviously all the lower bounds must be smaller than their respective upper bounds. These greater-or-equal restrictions were checked first in the evaluation of the optimization index. If any of the lower bounds were superior to its upper bound, the index was declared invalid (set to NaN), and penalized greatly in the search [29]. The final cost index is shown next:

$$J(\mathbb{P}) := \begin{cases} \text{NaN} & \text{if } \bar{p}_i < \underline{p}_i \\ J_{we}(\mathbb{P}) & \text{otherwise} \end{cases} \quad (9.7)$$

where  $\bar{p}_i$  and  $\underline{p}_i$  represent the upper and lower bounds of the parameter  $p_i$ .

Meal absorption is a highly complex physiological process as demonstrated by dual and triple tracer studies. However, to date, only relatively simple models fitted for specific meals are available. Besides, estimation of carbohydrates intake by the patient is a big source of uncertainty. It is thus expected a great extent of uncertainty in model parameters  $Bio$ ,  $t_{maxG}$  and model input  $D_g$ .

As the experimental data were taken under controlled conditions, it can be justified here to consider no misestimation of the grams of CHO ingested (real-valued  $D_g$ ) eliminating a big confounder of uncertainty.

Besides, two different time constants  $t_{maxG40}$  and  $t_{maxG100}$  were considered for the meals of 40 and 100 grams of carbohydrates, respectively, to account for possible non-linearity with respect to meal size. However, since only one meal of 40 g or 100 g would remain for identification considering that one day is left for validation purposes, variability in  $t_{maxG40}$  and  $t_{maxG100}$  cannot be expressed adequately by the available data, thus making their identification as intervals questionable. For this reason they were also considered real-valued. Uncertainty in the gastrointestinal system was then considered as an interval scale factor  $\alpha$  multiplying the glucose absorption rate  $G_{ex}(t)$ . This means that an envelope is produced from the rate of appearance resulting from an identified gastrointestinal model with real-valued parameters. Simulation studies have demonstrated that this aggregation of uncertainty into a single parameter  $\alpha$  may be considered equivalent to parametric uncertainty in  $Bio$  and  $t_{maxG}$  and similar data fitting capabilities are expected. Indeed, looking at Figure 9.1, little difference between the simulation bands of either parameter is observed. Furthermore, this allows a significant

reduction in the number of interval parameters in the identification problem, and thus in the computational cost.

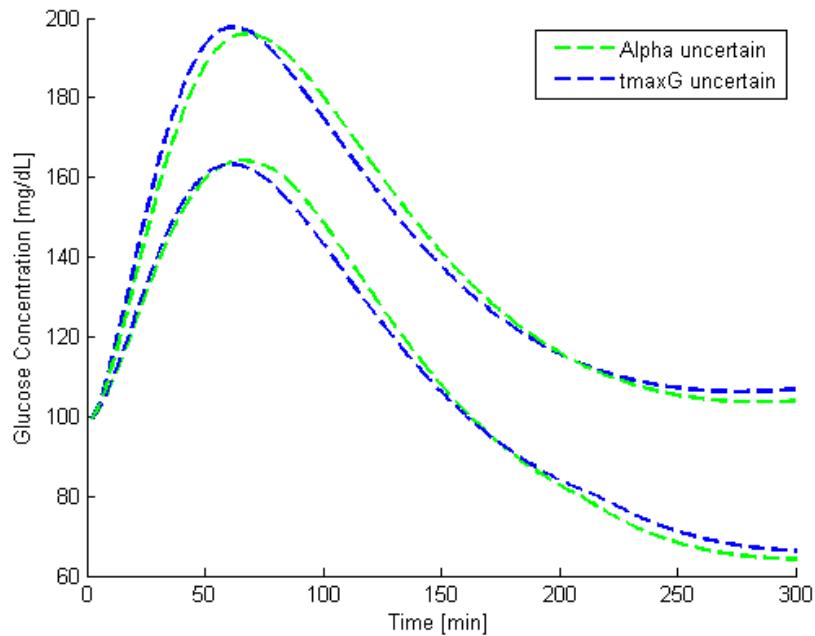


Figure 9.1: Simulation of a postprandial period considering a 15% uncertainty either in the parameter  $t_{maxG}$  (in blue) or in the aggregated parameter  $\alpha$  (green).

Description of uncertainty in the insulin and endogenous subsystems was considered in the insulin sensitivity parameters  $S_{IT}$ ,  $S_{ID}$  and  $S_{IE}$ , and the glucose transport among compartments  $k_{12}$ .

Identification of physiological models in diabetes often suffers from identifiability issues, especially when no tracer data is available and meal intake and insulin delivery is concomitant, as in this case. In order to test the ability of the method to properly characterize the source of uncertainty from only insulin and glucose data, two different scenarios were considered:

1. In Scenario 1 implicit uncertainty in the gastrointestinal (interval value for  $\alpha$ ), insulin and endogenous subsystems was considered.
2. In Scenario 2 only uncertainty in the insulin and endogenous subsystems was considered ( $\alpha=1$ ).

This is illustrated in Table 9.1. Parameters marked as “Interval” were identified as uncertain (interval values), while parameters marked as “Real” were identified as real-valued parameters. Parameters not listed or, listed as “Fixed” were not identified and maintained in their nominal values throughout the identification and validation process.

Parameters	Scenario 1	Scenario 2
$S_{IT}$	Interval	Interval
$S_{ID}$	Interval	Interval
$S_{IE}$	Interval	Interval
$k_{12}$	Interval	Interval
$\alpha$	Interval	Fixed
$t_{maxG40}$	Real	Real
$t_{maxG100}$	Real	Real
$V_g$	Real	Real

Table 9.1: List of parameters to be identified in each scenario. “Interval” stands for parameters identified with uncertainty. “Real” are traditional parameters identified. “Fixed” stands for parameters not being identified.

Out of the 4 days of monitoring available, only 3 were used for identification, while one day was kept apart for validation purposes (“leave-one-day-out” cross-validation). A few pilot optimizations showed that identifications were much more accurate when the first 30 minutes of each postprandial period were excluded for fitting the data, which may be due to non-modeled dynamics.

All possible permutations for the identification-validation days were used, obtaining a **full cross-validation** study for all the patients and all the postprandial periods. Henceforth, the different permutations are denoted by the validation day number. Thus, permutation 3 uses days 1, 2 and 4 for identification, and validates with day 3. All the permutations of identification-validation days were performed twice for testing the repeatability of the results.

Six different measures were computed from the identification and validation days in order to evaluate the goodness of fit and model prediction ability, respectively:

1. Width [mg/dL]: maximum width of the glucose envelope generated by the identified interval model. It represents the patient’s uncertainty for the postprandial period

2. Predictions [%]: number of glucose measurements included inside the predicted glucose envelope.
3. MARDout [%]: relative error of the samples that were not well predicted and fell out of the glucose envelope. It complements the Predictions measure. For example, if Predictions is low but MARDout is also low, the data may follow the dynamics of the model, but with an offset that forces the data out of the prediction band.
4. MARDtot [%]: relative error for all the glucose measurements. The glucose samples correctly predicted count as error zero, according to the error definition in equation 9.5.
5. gMARDout [%]: clinically-penalized relative error of the samples out of the glucose envelope. The clinical penalization was performed following the indications detailed by del Favero in [12] (see below).
6. gMARDtot [%]: clinically-penalized relative error for all the glucose measurements. Correctly predicted samples count as error zero. The penalization used was also the one proposed by del Favero.

MARDout and MARDtot are relative errors based on the Mean Absolute Relative Difference with respect to the glucose envelope, defined as:

$$MARD := \frac{1}{N} \sum_{i=1}^N \left| \frac{d_H(G^*(t_i), [\underline{G}(t; \mathbb{P}), \bar{G}(t; \mathbb{P})])}{G^*(t_i)} \right| \cdot 100 \quad (9.8)$$

where  $d_H$  is the Hausdorff distance previously defined, and  $N$  is the number of samples considered in the computation. For MARDout  $N$  is equal to the number of samples outside the envelope, and for MARDtot  $N$  is equal to the total number of samples in the postprandial period.

gMARDout and gMARDtot are modified versions of MARDout and MARDtot so that they gain on clinical interpretability. Del Favero et al. [20] proposed a penalization on identification indexes where danger of hypoglycemia and hyperglycemia is associated with larger weights in the index. In this work, an extension of this index to deal with glucose envelopes was defined as:

$$gMARD := \frac{1}{N} \sum_{i=1}^N \left| iPen \left( G^*(t_i), [\underline{G}(t; \mathbb{P}), \bar{G}(t; \mathbb{P})] \right) \frac{d_H(G^*(t_i), [\underline{G}(t; \mathbb{P}), \bar{G}(t; \mathbb{P})])}{G^*(t_i)} \right| \cdot 100 \quad (9.9)$$

where

$$iPen \left( G^*(t_i), [\underline{G}(t; \mathbb{P}), \bar{G}(t; \mathbb{P})] \right) := \begin{cases} Pen(G^*(t_i), \underline{G}(t; \mathbb{P})) & \text{if } G^*(t_i) < \underline{G}(t; \mathbb{P}) \\ Pen(G^*(t_i), \bar{G}(t; \mathbb{P})) & \text{if } G^*(t_i) > \bar{G}(t; \mathbb{P}) \\ 1 & \text{otherwise} \end{cases} \quad (9.10)$$

and  $Pen : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  is del Favero's penalization function. As in the previous case, for gMARDout  $N$  is equal to the number of samples outside the envelope, and for gMARDtot  $N$  is equal to the total number of samples in the postprandial period.

Comparison between gMARD and MARD indexes provides information on the medical risk of inaccurate model predictions. No difference between the indexes means no risk associated to the patient by definition, since it implies a unity value for the penalization function Pen in equation 9.10.

In order to consider an identification successful, it has to present good prediction capabilities resulting in a high percentage of samples predicted and consequently a small MARD for validation data. Clinical indexes should also not differ from their classical counterparts.

Besides, in order to discard trivial solutions, small envelope widths are expected in a successful identification, since a large enough band will always predict all the samples. Nonetheless, large widths are not discouraged if the variability of the patient is also large. In order to account for this fact, an envelope fitness measure was computed as

$$f(t_i) := \max \left[ \min_d |G^*(t_i) - \bar{G}(t; \mathbb{P})|, \min_d |G^*(t_i) - \underline{G}(t; \mathbb{P})| \right] \quad (9.11)$$

$$\text{env\_fit} := \frac{1}{N} \sum_{i=1}^N f(t_i) \quad (9.12)$$

where  $G^*(t_i)$  is the measured data,  $\underline{G}(t; \mathbb{P})$  and  $\bar{G}(t; \mathbb{P})$  the lower/upper bound of the glucose envelope, all of them at time  $t_i$  for day  $d$ . The rationale of this envelope fitness measure is that the envelope width is considered non-overestimated when the upper and lower interval glucose bounds actually fit the data for some postprandial period of the experiment. In this case, it is assumed that the variability of the patient was well captured. This is depicted in Figure 9.2 for a three-days experiment. For every  $t_i$  the minimum distance from the data to the upper envelope bound among all the postprandial periods  $d$  is computed, and similarly for the lower bound, considering finally the worst case. The average with respect to time  $t_i$  of worst-case distances is then considered as an indicator of the fitness of the envelope width to data. As illustration,  $\text{env\_fit}=10$  mg/dL indicates that for any time instant  $t_i$ , there exists a day  $d$  such that the discrepancy between the data and the upper/lower bound is less than or equal to 10 mg/dL in average.

All measures listed above have to be considered simultaneously when evaluating an identification experiment. An identification with low predicted samples does not mean a bad identification result if it simultaneously presents a very low relative error. Similarly, an identification with many predicted samples (on the validation day) but poor envelope fitness (on the identification days) may correspond to a trivial solution of the identification problem.

Mean, median, standard deviation and maximum/minimum values of the above measures for all the patients are reported. All significances were calculated using non-parametric Fisher's resampling test.

## 9.2 Identification from reference glucose

The different evaluation measures for the cross-validation study are reported in Table 9.2 for identification and validation days. Evaluations on the identification days are used to test goodness of fit, while the evaluation on the validation set is a good representation of

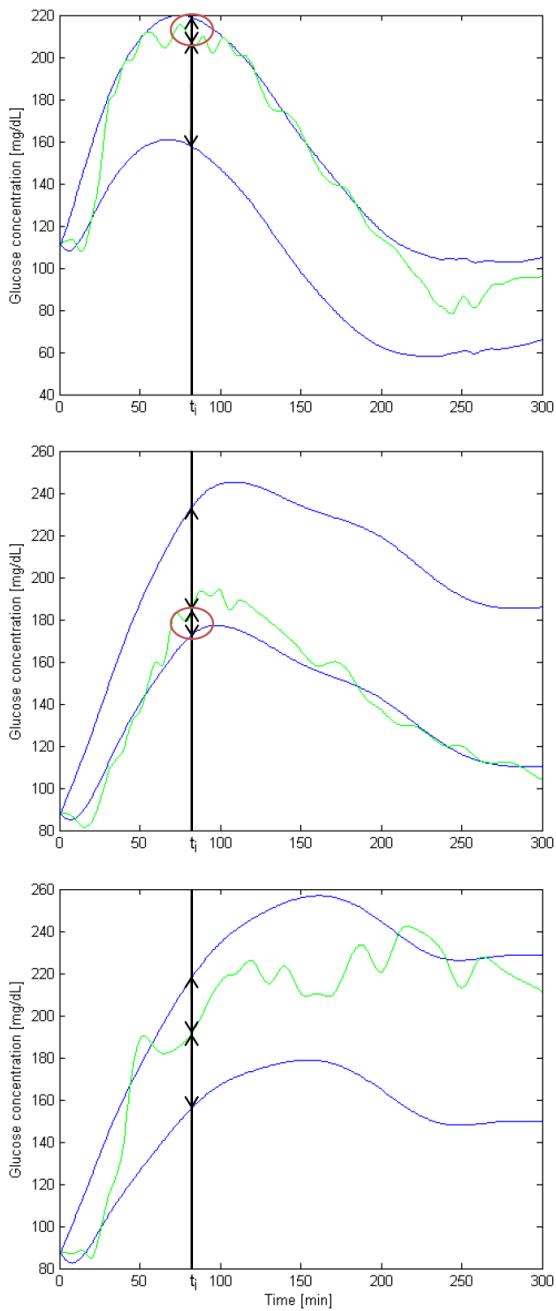


Figure 9.2: Graphic description of the envelope fitness index. Three postprandial periods are displayed for both YSI data (green) and glucose envelopes (blue). Distances of the data to the envelope are displayed, and the minimum distances to the upper and lower envelopes are highlighted. The average distance along time defines the envelope width fitness to data.

the predictability of the method and thus the overall performance of the identifications. The identified interval and real-valued parameters are reported in Table 9.3. Mean and median values for the midpoint and width of the intervals are shown, for the sake of simplicity. Median values are a good estimation of the magnitude of the identification results as they better filter outlier and are better estimators of the distribution expectancy for skewed distributions, but mean values of the evaluation measurements are useful for comparison between scenarios and datasets.

The mean value of the envelope fitness function for scenario 1 was 19.35 mg/dL (median 14.59 mg/dL) while for scenario 2 was 21.7 mg/dL (median 15.98 mg/dL). Fitness value per patient and scenario is reported in Figure 9.3. Patient 12 presents very poor envelope fitness value on both scenarios, and can be considered as an outlier.

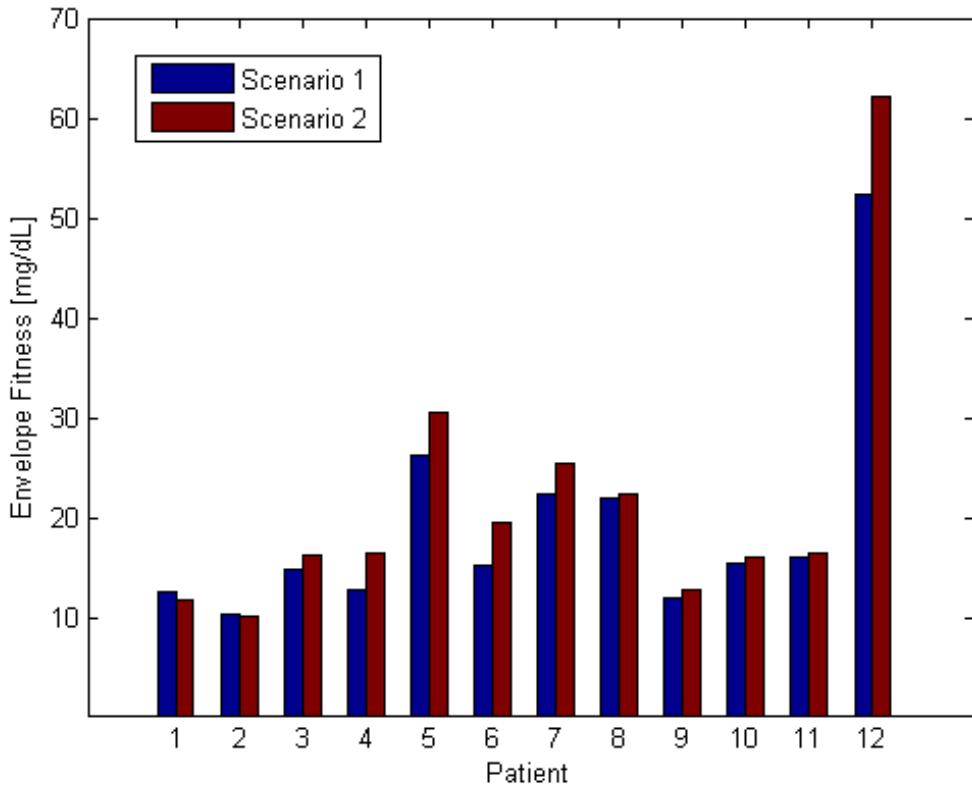


Figure 9.3: Envelope fitness for every patient and scenario.

Analyzing the results in Table 9.2 it can be observed that envelope widths for the

Scenario	Width [mg/dL]	Prediction [%]	MARDout	MARDtot	gMARDout	gMARDtot
			[%]	[%]	[%]	[%]
Identification days						
1	mean	87.7	77.9	5.79	1.22	5.92
	SD	39	7.2	6.1	1.07	6.11
	median	81.2	78.9	4	0.92	4.11
	max/min	187.9/40.4	92.2/63.6	32.14/1.37	4.84/0.24	32.16/1.38
2	mean	94.8	78.4	6.56	1.35	6.7
	SD	42.5	6.8	9.01	1.64	9.01
	median	89.9	78.1	3.9	0.85	4.08
	max/min	218.3/41.2	93.4/64.7	48.84/1.67	7.89/0.25	48.87/1.67
Validation days						
1	mean	78.8	53.5	12.23	7.61	13.46
	SD	38.4	30.1	8.83	7.88	10.62
	median	66.4	52.8	11.22	5.56	11.97
	max/min	205.1/29.3	100/0.7	32.93/0	30.41/0	48.49/0
2	mean	83.9	54.9	12.66	8.09	13.91
	SD	40.8	31.9	10.3	9.06	11.69
	median	79.1	57	9.65	5.45	12.14
	max/min	222/34.6	100/0.3	37.87/0	31.19/0	43.42/0
Validation days						
1	mean	78.8	53.5	12.23	7.61	13.46
	SD	38.4	30.1	8.83	7.88	10.62
	median	66.4	52.8	11.22	5.56	11.97
	max/min	205.1/29.3	100/0.7	32.93/0	30.41/0	48.49/0
2	mean	83.9	54.9	12.66	8.09	13.91
	SD	40.8	31.9	10.3	9.06	11.69
	median	79.1	57	9.65	5.45	12.14
	max/min	222/34.6	100/0.3	37.87/0	31.19/0	43.42/0

Table 9.2: Results for both scenarios for the 3 identification days with all the possible permutations, and the validation days.

Interval parameters		$S_{IT} \times 10^{-3}$ [mU/min · L]	$S_{ID} \times 10^{-4}$ [mU/min · L]	$S_{IE} \times 10^{-2}$ [mU/min · L]	$k_{12} \times 10^{-2}$ [min $^{-1}$ ]	$\alpha$
Scenario 1	Mid point	mean 5.05	6.08	4.43	6.43	1.12
	Width	SD 1.32	3.06	2.11	2.45	0.4
Scenario 2	Mid point	mean 5.04	0.05	0.45	1.16	0.02
	Width	SD 1.24	0.25	0.9	2.01	0.08
Nominal value		5.12	8.2	5.2	6.6	1
Real-valued parameters						
Scenario 1	mean	$t_{maxG40}$ [min] 85.1	$t_{maxG100}$ [min] 98.8			
	SD	68.5	31.4			
Scenario 2	mean	87.2	99.2			
	SD	66.7	39.8			
Nominal value		40	40			

Table 9.3: Identified parameters for both scenarios. Nominal values were extracted from [35].

three identification days were greater than those shown in the validation day (scenario 1: 87.7 vs. 78.8 mg/dL,  $p<0.005$ ; scenario 2: 94.8 vs. 83.9 mg/dL,  $p<0.005$ ). Identified widths were large, but the uncertainty registered resembled the observed intra-patient variability, given that the envelope fitness measure on both scenarios was good in average (14.59 vs. 15.98 mg/dL). This yields to an expectancy of the maximum separation of the prediction boundaries from the actual data in accordance to current criteria for glucose measuring devices accuracy. This result is valid throughout all the postprandial periods, implying that the method accurately fits the provided data even though prediction bands are wide. Besides, errors for validation days were (as expected) larger than those for identification days for both scenarios 9.2.

There were no great differences in the overall performance comparing validation results between scenarios. No statistically significant difference in the prediction error was found (7.61 vs. 8.09,  $p=0.237$ ) despite smaller envelope maximum widths in favor of scenario 1 (78.79 vs. 83.9,  $p<0.005$ ). This is so because maximum width tends to happen at the end of the postprandial period when data are more likely to be enclosed (contributing with zero error to the MARD), not having an impact in the overall prediction error. Nevertheless, the difference in the envelope maximum width (5.11 mg/dL) was not considered clinically relevant.

This equivalence between scenarios may suggest identifiability issues (as commonly observed in identification studies) and difficulties to properly detect the physiological source of uncertainty due to the nature of data used. Indeed, uncertainty sources were found in repeated patterns all across the identifications when looking at the identified parameters (see Table 9.3). In this regard, Table 9.4 shows the number of identifications for each scenario that showed uncertainty in each of the interval parameters considered. If both limits of the interval parameter were identified to be the same, that parameter was considered real-valued and no uncertainty was assumed to exist in that case.

	$S_{IT}$	$S_{ID}$	$S_{IE}$	$k_{12}$	$\alpha$
Scenario 1	95.8%	4.2%	33.3%	41.7%	6.3%
Scenario 2	93.8%	4.2%	29.2%	35.4%	fixed

Table 9.4: Percentage of identifications in which each parameter presented uncertainty.

Values in Table 9.3 are related to the widths of the intervals in Table 9.4. If the interval is identified in many occasions as a real-valued parameter, the average width of that parameter will be very low. Parameters  $S_{IT}$  and  $k_{12}$  were the most uncertain parameters in both scenarios, which are interestingly both related to glucose transport

between plasma and interstitial fluid. However, the gastrointestinal system contributed to uncertainty for scenario 1 only in 6% of the trials despite the variability expected a priori according to tracer studies. This may be due, at least in part, to the consideration of two different  $t_{maxG}$  parameters dependent on the meal size, which were found statistically significantly different for both scenarios (see Table 9.3), and the few instances available of repeated meals during cross-validation (only two of the days used for each identification permutation were subject to pure gastrointestinal variability, i.e., the two days with identical meal input). This resulted in small widths for the common parameter  $\alpha$ . Richer (more repeated meals, tracer studies) data would most probably lead to larger variability in the gastrointestinal system. This is considered a limitation of the current data and further investigation is required. However, prediction ability of the method was demonstrated and no technical limitations exist should tracer data or longer time series be available.

As an illustration, identification results for two patients are shown in figures 9.4 and 9.5, for each identification-validation day permutation. Even though four simulated postprandial periods are displayed per identification, the simulations were not consecutive. Every day was simulated separately considering steady-state initial conditions in order to account for the pre-prandial stabilization period of the patients. Then the simulation profiles were concatenated for better illustration of the overall identification performance. Results are shown for scenario 1, although both scenarios reported very similar outcomes.

Figure 9.4 (patient 2) poses an example of an average patient's identification outcome. Results correspond to scenario 1 because of physiological soundness, although both scenarios reported very similar outcomes. For the first permutation (top left), validation was not successful ( $MARD_{tot} = 7.39\%$ ). Most of the postprandial data was overestimated, even though 5-hour glucose value was well captured. Though, in clinical terms, the identification implied no danger for the patient ( $gMARD_{tot} = 7.39\%$ ), since the clinical index was exactly the same than the prediction error. For the second permutation (top right), dynamics were modeled correctly, but postprandial peak was greatly underestimated ( $MARD_{tot} = 6.03\%$ ), and the effect of glucose infusion in the last hour to avoid hypoglycemia was overestimated. Since no dangerous zones were left unmodelled, clinical validation was also successful ( $gMARD_{tot} = 6.03\%$ ). Third permutation (bottom left) captured perfectly the dynamics of the postprandial period, even though no similar behaviors are observed in the identification days, only leaving out of prediction the last part of the 5-hour period ( $MARD_{tot} = 2.29\%$ ). However, clinical validation was not so good in this case, with  $gMARD_{tot} = 3.58\%$  due to the last part of the simulation not being

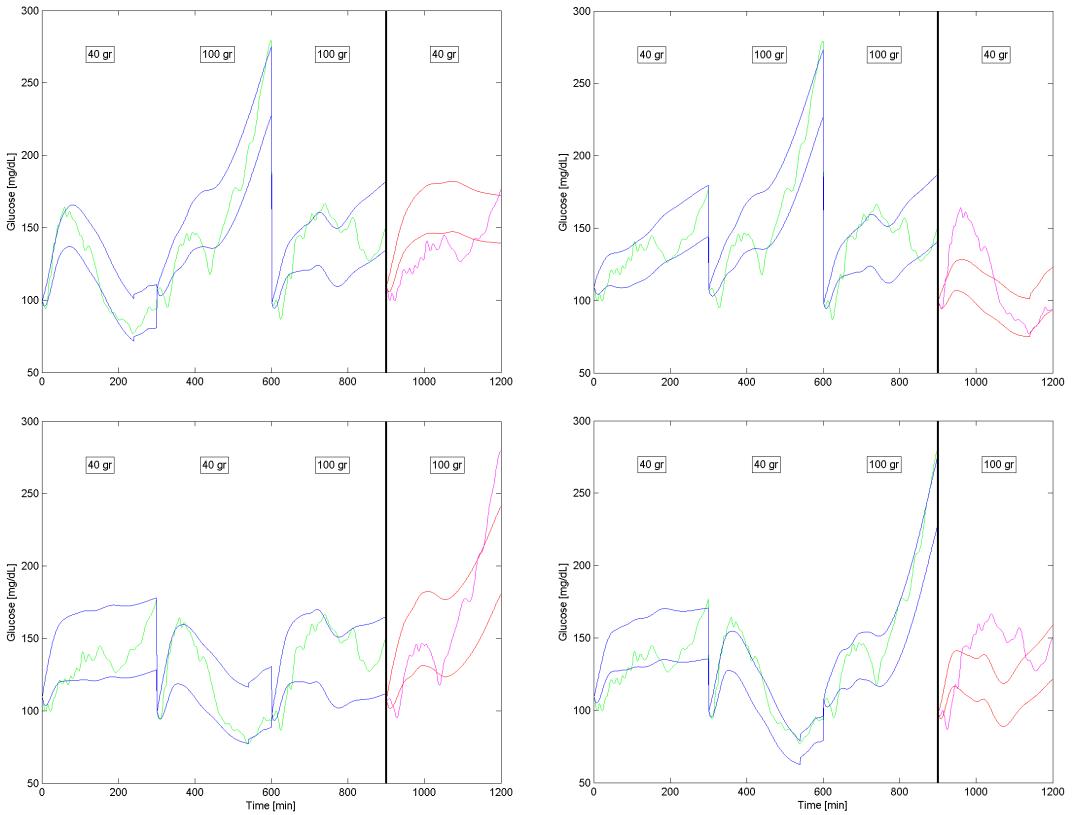


Figure 9.4: Patient 2 identification and validation results for the scenario 1. Day 1 validation is shown top left. Top Right graph shows validation for day 2. Bottom left validates day 3. Bottom right shows validation for day 4.

able to follow the rapid rise of glucose, missing thus the hyperglycemia risk. Fourth permutation (bottom right) showed some not predicted samples ( $MARD_{tot} = 9.18\%$ ), but no clinical risk was involved ( $MARD_{tot} = 9.2\%$ ).

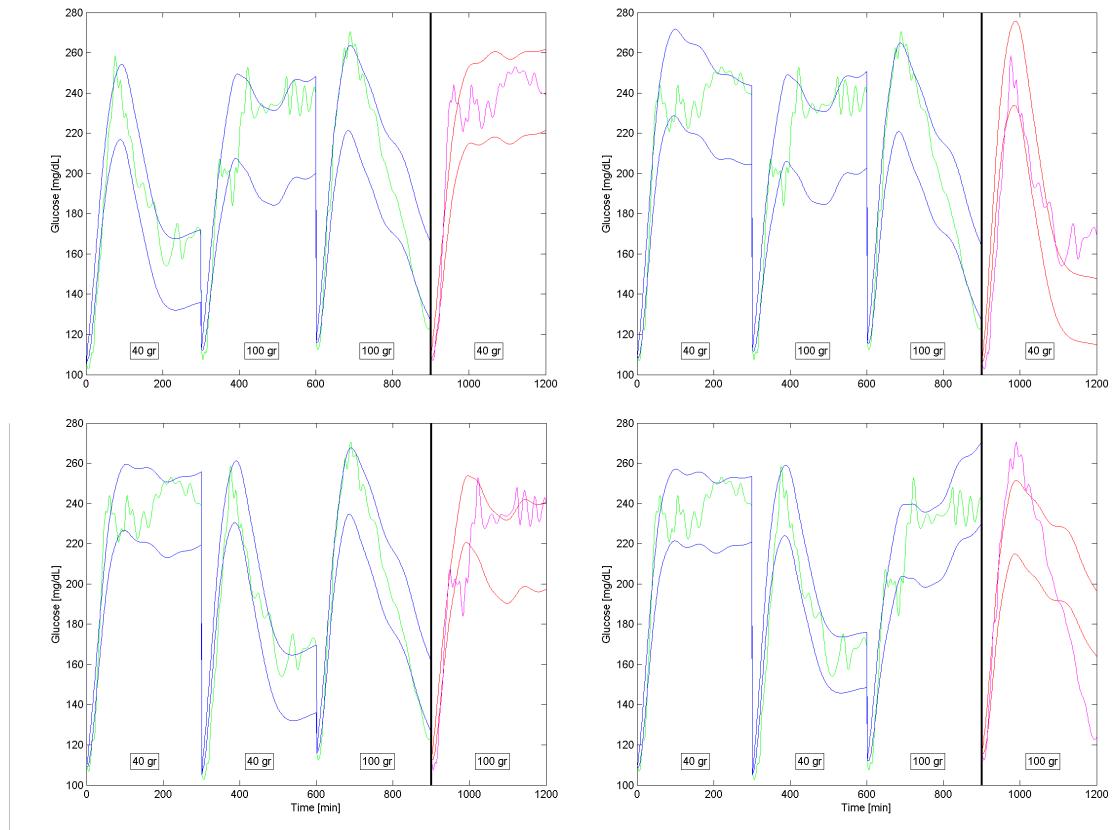


Figure 9.5: Patient 9 identification and validation results for the scenario 1. Day 1 validation is shown top left. Top Right graph shows validation for day 2. Bottom left validates day 3. Bottom right shows validation for day 4.

Figure 9.5 presents a case of a better identification, corresponding to patient 9. Day 1 prediction envelope (top left) covered almost all the blood glucose samples, and the overall error was very small ( $MARD_{tot} = 1.03\%$ ). Clinical error was not much higher ( $gMARD_{tot} = 1.8\%$ ) in this validation because most part of the postprandial period is spent in hyperglycemic region, and the model fails to predict a small part of it, although with very small error. The patient was well predicted when validating with day 2 (top right) except in the 5-hour horizon ( $MARD_{tot} = 4.93\%$ ). Validation was also really good for day 3 (bottom left), with a  $MARD_{tot}$  of  $1.88\%$ . Dynamics were well captured when validating

with day 4 (bottom right), with some lack of predictability from 3 hours onward in the postprandial period ( $MARD_{tot} = 6.88\%$ ). As happened with the first day, the clinical error is a little bit higher because the model fails to predict the hyperglycemic peak, but with very small error. As result, the patient was successfully identified for every permutation, capturing the dynamics of the patient independently of the days used for identification for an average envelope fitness function value of about 11.87 mg/dL (see Figure 9.2).

### 9.3 Best-case permutation

Figure 9.6 represents a patient with very good prediction for only one validation day, corresponding to patient 11. As with the other examples, results are shown for scenario 1, although both scenarios reported very similar outcomes.

Figure 9.6 shows the most illustrative case for the best case of a particular patient. We can see how only when days 1, 2 and 4 were used for identification (bottom left), dynamics of the validation day was well characterized. In all the other cases, the identification was unsuccessful. Envelope width for the best-case day was also wider than in the other cases (without width overestimation). This fact leads to think that variability of this particular patient showed its maximum for that particular permutation of monitoring days. Maximization of predictability is then straightforward if interval model identification works best when maximum variability is present in the data: the more days available the higher the variability, thus the better the prediction.

As it may be observed from patient 11, validation results are dependent on the permutation considered due to the different representativeness of the identification days with regard to the exhibited intra-patient variability. In the general case, out of the four permutations of the identification, usually one or two of them presented a very small error, while the rest did not capture correctly the behavior of the patient mainly due to lack of representativeness of the identification data with regard to the exhibited intra-patient variability. Indeed, it cannot be expected to predict a behavior in the validation day significantly different than the ones present in the identification data. The inclusion of more days for identification would reduce this problem. In this regard, a best-case permutation was defined as the permutation with minimum clinical error  $gMARD_{tot}$ , thus minimizing both the fitting error and the danger to the patient. When two of the permutations showed similar results, the solution with the minimum width was considered the best one. Table 9.5 shows the evaluation measures for the validation

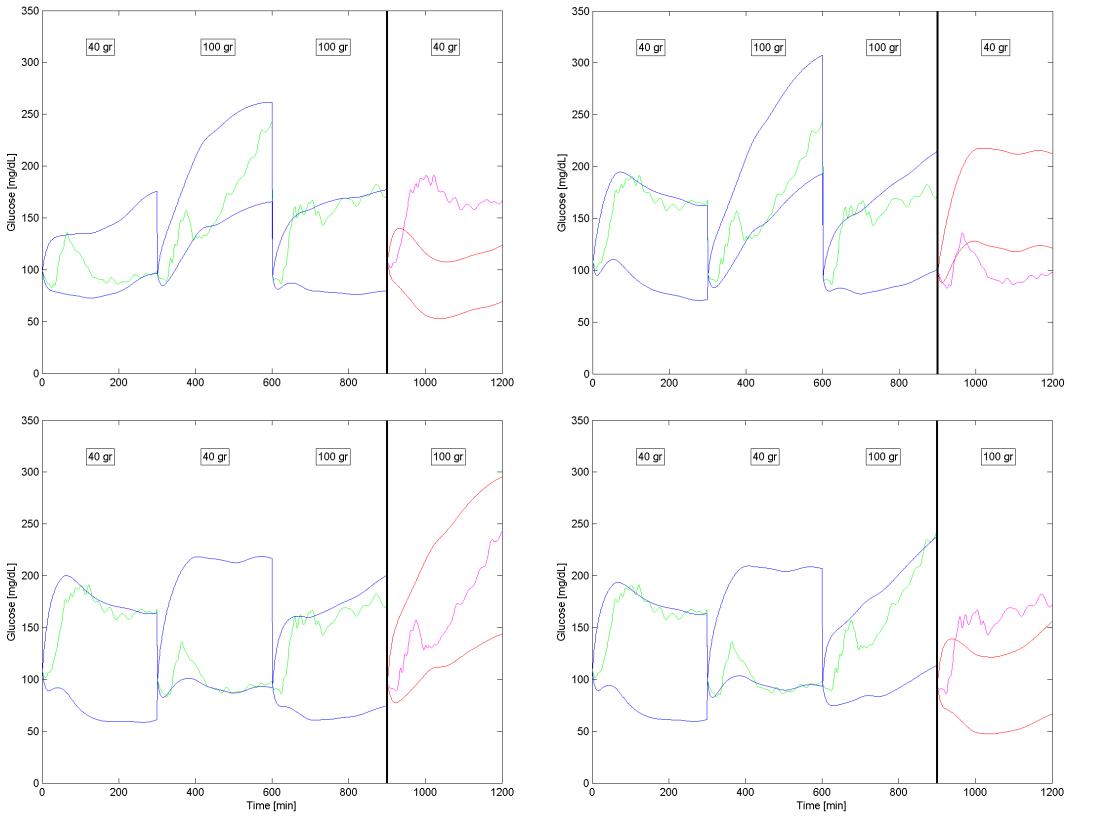


Figure 9.6: Patient 11 identification and validation results for the scenario 1. Day 1 validation is shown top left. Top Right graph shows validation for day 2. Bottom left validates day 3. Bottom right shows validation for day 4.

days considering the best-case permutation. Mean envelope fitness for the best cases in scenario 1 was 19.87 mg/dL (median 15.55 mg/dL), and for scenario 2 was 26.74 mg/dL (median 19.98 mg/dL).

Scenario	1				2			
	Mean	SD	median	max/min	Mean	SD	median	max/min
Width [mg/dL]	104.6	51.1	94.3	201.1/42.5	103.8	43.6	101.3	179.2/43.5
Prediction [%]	87.3	17.6	97.3	100/52.3	87.1	18.4	96.8	100/45.3
MARDout [%]	4.38	5.78	1.51	17.81/0	3.32	4.41	0.87	13.17/0
MARDtot [%]	1.24	2.43	0.03	8.49/0	0.8	1.18	0.03	3.38/0
gMARDout [%]	5.08	6.27	1.51	17.81/0	4.24	5.79	0.87	15.94/0
gMARDtot [%]	1.41	2.51	0.03	8.49/0	1.06	1.72	0.03	5.63/0

Table 9.5: Results for both scenarios for the validation days of the best case permutation.

Table 9.5 shows the evaluation metrics for the best-case permutation for each patient. As expected, errors for the validation days were significantly smaller for the best cases with respect to results in Table 9.2 (scenario 1: 1.24 vs. 7.61 %  $p < 0.005$ ; scenario 2: 0.8 vs. 8.09 %,  $p < 0.005$ ). Widths were significantly larger for scenario 1 (104.6 vs. 78.8 mg/dL,  $p = 0.034$ ) but not significantly larger for scenario 2 (103.8 vs. 83.8 mg/dL,  $p = 0.075$ ), although we assume this lack of significance is due to the small sample size of 12 patients, and should the number of patients be larger, the difference in widths is expected to increase. This increment in the width was due to the increment of variability present in the identification data, improving representativeness of the patient's behavior, and thus validation results. No additional envelope width overestimation was found for scenario 1 (mean/median: 19.87/15.55 mg/dL versus 19.35/14.59 mg/dL for the whole data set). It was not so for scenario 2 (mean/median: 26.74/19.98 mg/dL versus 21.70/15.98 mg/dL) which compensated the non-physiological consideration of lack of variability in the gastrointestinal model with envelope width overestimation. Indeed, median Predictions value was over 95% (mean value over 85%) for both scenarios. This implied errors virtually equal to zero, and complete coverage of the validation data. No difference in the width between the scenarios was found (see Table 9.5). However, the observed envelope width overestimation for scenario 2 leads to the conclusion that scenario 1 better represents

intra-patient variability without incurring in model overfitting.

Finally, comparing Tables 9.2 and 9.5, clinical indexes were statistically significantly larger than their pure error counterparts, both for the whole dataset (scenario 1: 8.51 vs. 7.61 %,  $p < 0.005$ ; scenario 2: 8.95 vs. 8.09 %,  $p < 0.005$ ) and for the best-case permutations (scenario 1: 1.41 vs. 1.24,  $p < 0.005$ ; scenario 2: 1.06 vs. 0.8,  $p < 0.005$ ). However, differences between gMARDtot and MARDtot were reduced from approximately 1% for the whole dataset to approximately 0.2% (corresponding to underestimated mild hyperglycemia according to [20], Fig. 2) for the best-case permutation. Hypoglycemia risk was not at stake.

## 9.4 Discussion

Identification from experimental data supposes a great challenge. The case exposed in this chapter is the simplest situation of identification with real patient's data the author was aware of, and yet identification procedures turned out very complex. A complete identification study with implicit consideration of uncertainty in the system has been presented. Identification was performed using a hybrid cost index minimizing both the envelope width of the interval prediction model so that uncertainty is minimal, and fitting error to the data. Only the endogenous and gastrointestinal models where considered in this chapter, as a pilot identification study.

Identification of the 12 patients showed good prediction capabilities in average and especially when the maximum variability, or uncertainty for that particular patient, was represented in the identification days (best-case permutation). A larger number of monitoring days of the same patient would increase the probability of extreme variations of the metabolism of the patient, thus easing the identification and increasing the predictability of the data. A limitation of the study was precisely the nature of the data that prevented to draw conclusions on the source of the identified uncertainty. Contrary to tracer studies, the results here exposed did not show significant uncertainty in the gastrointestinal system compared to the insulin and endogenous subsystem, most probably due to data limitations.

The existence of best-case permutation and the possibility of predicting the behavior of the patient using this combination is the most important contribution of this chapter, and one of the main findings of the thesis. Although it is well known that patient variability in diabetes is very high, this work is the first successful identification process

performed on real data. The best-case permutation for each patient is still impossible to predict, and apparently appears at random in the dataset used for this chapter. Further work can be done on investigating the characteristics of the patient on the days of maximum variability, using for example pattern search on physiological variables.

The proposal of interval envelope fitness measures has been found of great use throughout all the pilot identification. Using a hybrid index for optimization limits the parameter possibilities that multiobjective optimization presents, and because of that, trivial solutions must be discarded using the fitness index. Envelope fitness seems to have much greater potential for the interval identification problem, and further investigations are required for its potential applications.

Only YSI data was used in this work, and it was assumed that plasma insulin values were available all throughout the postprandial period. The next logical step will be to evaluate the influence of the subcutaneous insulin route on the prediction capabilities. Further investigation has to be done on the influence of CGM to the identifications, instead of using YSI references, which are much more accurate.

# Chapter 10

## Experimental Interval Identification

Exploring the feasibility of full patient characterization is the main objective of this thesis, and the results of the *in silico* and pilot experimental identification were encouraging. However, many of the premises of the previous chapters are not realistic in the consideration of an ambulatory patient identification for the artificial pancreas. For practical purposes, a home-made patient identification must be performed over the information that a patient obtains in daily routine. This requires the use of CGM measures for the glucose estimation signal, and either pump or periodic injections, as well as meal estimation.

Identification methodology and the diabetic patient data for this chapter has already been presented in previous chapters. CMAES optimization algorithm was used for the minimization of the hybrid cost index explained in the pilot study in chapter 9. The hybrid index is not different from that explained previously, and the only difference remains in the vector of parameters to be optimized.

### 10.1 Optimization settings and parameters

In this chapter identification is performed realistically on a full physiologic model of a diabetic patient. The model used for the identification included, unlike the previous chapter, a system for the subcutaneous insulin absorption. This subsystem is known to suffer from enormous variabilities, and as such it is well suited for characterization

from interval models. The insulin absorption model used was the one included in the Cambridge simulator, in order to keep coherence with the rest of the models used. The parameter  $k_a$  defined in chapter 1.1.2 is referred here as  $t_{maxI}$  in order to keep the same notation for both the insulin and glucose absorption models.

In the case of insulin pump usage, which is the case of the dataset used, there is also the possibility of pump-induced errors, probably caused by total or partial occlusions. This type of errors is very similar to the estimation errors present in the amount of CHO of a meal. A very similar approach to that of the aggregated parameter  $\alpha$  in the gastrointestinal system was used here to gather all uncertainties in the insulin model input into one single parameter, defined  $\beta$ . This parameter was treated as an interval parameter and tested for identification capabilities in the following.

The other parameter influencing the insulin absorption is the insulin elimination rate from plasma  $k_e$ . This parameter was identified considering variability. Exactly in the same way as with the gastrointestinal model (in the  $t_{maxG}$ ), the flux constant  $t_{maxI}$  was identified on every patient, but due to the already mentioned interval model implementation limitations, no uncertainty was considered for it, so it was considered a constant value on each patient for the whole duration of the experiment. In practice, little difference exists when considering the uncertainty in the parameter  $\beta$  or in the  $t_{maxI}$  parameter when fitting experimental data, similarly to what happened when comparing  $t_{maxG}$  and  $\alpha$  in the gastrointestinal model. Very simple simulations of uncertainty are displayed in Figure 10.1 for both the  $\beta$  parameter and  $t_{maxI}$ .

Both parameters have virtually no influence in the band width of the glucose profile in the 45 minutes following the meal intake. Past that point  $t_{maxI}$  may present larger uncertainty related, but the glucose envelopes are very similar in the late postprandial period (two hours after ingestion). Even though the simulation shown in the gastrointestinal model showed better similarities between the parameters under comparison, the simulations for the model of subcutaneous insulin infusion are still considered to be very similar. Again, it is worth mentioning that these models will be used for data fitting, so the envelopes shown in Figure 10.1 are but a qualitative representation of the uncertainty related to each parameter throughout the postprandial period.

Finally, the complete set of parameters for identification is:

$$\theta = [S_{iT}, S_{iD}, S_{iE}, k_{12}, k_e, \alpha, \beta, t_{maxI}, t_{maxG}, V_g] \quad (10.1)$$

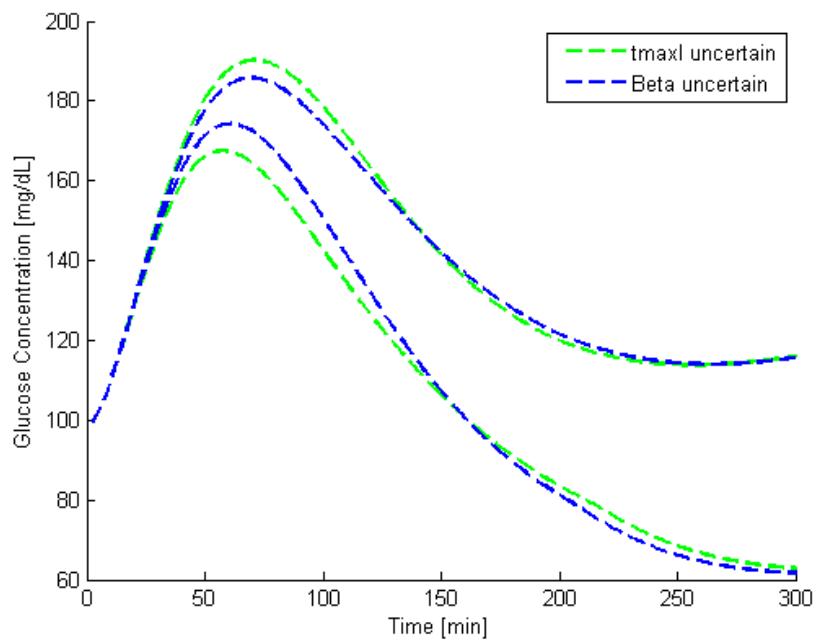


Figure 10.1: Simulation of a postprandial period considering a 15% uncertainty in the parameter  $t_{maxI}$  (in green) or in the aggregated parameter  $\beta$  (blue).

The parameter set proposed is equivalent to that used in the *in silico* identifications.

$\gamma$  range is one of the most important discussions in the completion of any identification with the proposed hybrid index. The selection of an appropriate  $\gamma$  is dependant on the type of the data being fitted, and of the error expected in the measurements in particular. The larger the error on the measurements is, the less importance is assigned to the fit part of the composite index. This approach is the opposite to the error-bound estimation explained in chapter 3.2.1, where the parameter set searched was of those parameters that are consistent with the data and its error. That approach lead to larger parameter sets as the uncertainty grows in the measurement. In the approach used here works exactly on the opposite way, ensuring a mimimum set of parameters that are consistent with the *physiology* of the patient and the uncertainty in the patient, not on the measurements. This physiologic set of parameters is smaller as less accurate estimations for glucose are available.

In the last chapter, a  $\gamma$  value of 100 was considered appropriate for the data being fitted. In the case at stake now, the data available is less compliant with the mathematical model because physiologic approximations are introduced in the subcutaneous insulin model. Therefore, less accurate predictions are expected from the model used and fitting error may be desired to be fit more loosely. In order to account for this, two different  $\gamma$  values are tested for every combination of parameters considered in the identifications:  $\gamma = 100$  and  $\gamma = 50$ . Having  $\gamma = 100$  allows for comparison with previous results, because the error ponderation is the same as in chapter 9.  $\gamma = 50$  reduces the influnce of the fitting error for accounting with the exposed model mismatches.

For the selection of identification parameters, several identifications were performed on reference glucose for different parameter vectors. A total of eight different permutations of the parameters were considered. The scenarios are detailed in table 10.1.

The described scenarios were repeated for two different  $\gamma$  values, resulting in a total 16 identification trials. The three parameters that define each scenario are identified as interval parameters. Parameters  $t_{maxI}$  and  $t_{maxG}$  are identified as real-valued parameters, and two different  $t_{maxG}$  values are identified in the same was as in the previous chapter, considering that the  $t_{maxG}$  parameter for the 100 grams meals is forced to be larger than the 40 grams parameter, reflecting a slower absorption of a much larger meal.

Identification boundaries for the parameters were the same for all the scenarios, for sake of comparison between scenarios. Repeatability of the results was found to be lower when introducing the new set of parameters correspondant to the subcutaneous

Scenario	$\alpha$	$\beta$	$k_e$
A	✓	✓	✓
B	✓	✓	✗
C	✓	✗	✓
D	✓	✗	✗
E	✗	✓	✓
F	✗	✓	✗
G	✗	✗	✓
H	✗	✗	✗

Table 10.1: List of parameters to be identified in each scenario.

insulin model. In order to improve the convergence of the identification solution, the CMAES optimization algorithm permits the tuning of several parameters. It was decided to increase the size of the initial population of individuals for better coverage of the parameter space, which is now significantly larger. Nevertheless, 2 identifications of each scenario with the exact same optimization parameters were performed for repeatability testing, and solution with the best index was chosen.

Finally, for the desired scenario chosen from the identification of a full patient model using YSI measurements, CGM identification is to be performed. Identification settings are exactly the same as in the YSI reference frame, but the data used is not trusted to be as accurate. Indeed, it has been discussed before in this thesis (chapter 6) how the CGM estimations can be as far off as 20% from the glucose reference signal. This lack of trust on the data translates in a smaller  $\gamma$  in the optimization index. In order to search for the most satisfactory scenario, several identifications are performed for different gamma values, decreasing from the value of 100 stated in the first identifications with the YSI data. Six different  $\gamma$  values were considered in intervals of 15: 100, 85, 70, 55, 40 and 25. The combination of all the results for different  $\gamma$  values results in a pseudo-PF that provides a selection of identification results suitable for the data available.

For evaluation of the results several metrics that have been explained introduced are used:

- Width [mg/dL]: Is the maximum width of the envelope resultant from the simulation of the postprandial periods. Is one of the parts of the composite index in the optimization.
- Prediction [%]: Percentage of samples predicted within the envelope.

- MARD [%]: Mean relative deviation of the envelope form fitting or validation data.
- gMARD [%]: Mean relative deviation weighted depending on the danger to the patient, following the index proposed by del Favero [20].
- env\_fit [mg/dL]: Envelope fitness measure defined in chapter 9. Describes the maximum separation of the envelope to the fitting data, and quantifies the overestimation of the band width.

Optimal scenario selection is done by analyzing the described metrics and also the optimization results in the parameter space. Although in the previous chapter it was concluded that the characterization of the uncertainty is very difficult for the models used now in literature, identifications perfomed with the full model, and especially considering the new introduced parameters, are understood to be faulty if parameters are fixed to the optimization boundaries in most of the permutations.

## 10.2 Identification from reference glucose

The first results for the identification of the full model are displayed in table 10.2. Only the results from the identification sub-set are displayed in the table. Computation times vary on the number of parameters being fit. The larger number of parameters was 18 for the A scenario and the computation time for all the patients was over 70 hours on a workstation Intel®Xeon®CPU 2.67 GHz with 4 GB of RAM memory running under Windows 7. For the fastest identification, which consisted on 12 parameters, the computation time was approximately 50 hours on the same computation setting.

The results metrics show very similar fit for the use of the full model to the fit of the pilot identification (table 9.2). This fact suggests that fitting capabilities are mainly dependant on the quality of the data being fit, although it is assumed that the quality of the predictions is lower. Comparison between different  $\gamma$  values is as expected: smaller widths are identified for  $\gamma = 50$  than in the larger  $\gamma = 100$  scenario, but at the cost of larger fitting errors, reflected in the MARD metric. Comparing envelope fitness to the scenarios of the pilot study reveals that no further overestimation is introduced by adding the subcutaneous insulin route to the model, because the env\_fit values are very similar to those of the scenario 1 in the pilot study.

Scenario		A	B	C	D	E	F	G	H
$\gamma = 100$									
Width	mean	80.4	85.9	86.9	98.8	87.6	92.2	92.9	105.2
	median	71.6	78.2	77	101.8	82.7	88	87.1	109.5
Prediction	mean	76.6	75.8	77.3	74.8	77.2	77.6	78	77
	median	78.2	76.8	78.8	77.6	77.6	79.1	79.2	77.7
MARD	mean	0.9	0.97	0.88	1.12	0.85	0.87	0.85	1.01
	median	0.76	0.95	0.76	1.08	0.77	0.83	0.77	1
gMARD	mean	0.93	1	0.91	1.14	0.88	0.9	0.88	1.05
	median	0.79	0.96	0.79	1.08	0.8	0.86	0.8	1.02
env_fit	mean	16.2	18.7	17	19.9	17.5	18.7	18.2	20.7
	median	14.4	16.5	15	16.7	15.1	17.1	17.8	18
$\gamma = 50$									
Width	mean	70.9	77.2	77.9	89.1	79	81.7	84.3	95.9
	median	63.4	69.6	70.1	90.3	71.3	75.8	80.4	98.5
Prediction	mean	67.1	67.7	68.8	66	69.7	69.4	70.4	69
	median	70.4	70.4	72	70.9	70.8	72.1	72.2	71.3
MARD	mean	1.6	1.54	1.51	1.79	1.49	1.52	1.47	1.64
	median	1.29	1.45	1.26	1.67	1.28	1.45	1.34	1.59
gMARD	mean	1.67	1.61	1.59	1.87	1.57	1.6	1.55	1.73
	median	1.35	1.5	1.3	1.74	1.34	1.53	1.38	1.65
env_fit	mean	14.6	16.1	15.7	17.3	15.6	16.4	16.3	18.2
	median	11.4	13.5	12.5	14.8	12.7	14.8	15.4	15.9

Table 10.2: Results for all the scenarios for the 3 identification days considering all the possible permutations of days from the dataset.

The prediction capabilities of the models identified and the envelope widths related to the validation days are listed in table 10.3. Even though widths of identification and validation days have high correlations, the reproduction of a pseudo pareto front with the data from the validation days is the best visual representation of the results for scenario selection.

Scenario		A	B	C	D	E	F	G	H
$\gamma = 100$									
Width	mean	77.2	78.7	82.4	90.1	81.5	83.5	87.6	95.3
	median	67.3	65.6	72.7	80.3	72.2	69.7	81	84
Prediction	mean	51.3	53	53.6	48.1	52.9	53	55.6	56.2
	median	56.5	56.3	56.8	49.5	57.3	57.8	57.8	61.5
MARD	mean	9.96	9.41	9.63	10.93	9.83	10.81	9.88	9.27
	median	4.82	5.17	4.34	5.51	5.67	4.03	4.06	3.73
gMARD	mean	11.60	10.98	11.24	12.65	11.39	12.4	11.47	10.87
	median	5.84	5.19	4.59	5.68	5.67	4.15	4.07	3.74
$\gamma = 50$									
Width	mean	67.1	71.1	74.1	80.3	74	73.4	79.7	87.4
	median	55.8	59.8	69.8	68.8	62.2	63.1	76.5	78
Prediction	mean	45.7	45.5	48.8	42.2	49.8	46.3	51.4	51.1
	median	46.3	46.7	52.7	40.2	55.2	51	56.2	56.2
MARD	mean	11.94	10.82	10.73	12.22	11.3	12.48	10.75	10.15
	median	5.89	7.26	6.44	7.20	6.54	6.2	5.41	5.04
gMARD	mean	13.7	12.55	12.43	13.92	12.92	14.22	12.42	11.85
	median	7.01	7.84	6.86	7.42	6.54	6.37	5.51	5.14

Table 10.3: Results for all the scenarios for the validation day considering all the possible permutations from the dataset.

Samples predicted by the scenarios considered are very similar to those presented in the pilot study, but MARD values are larger, meaning a poorer prediction of the results. This prediction is not a consequence of overfitting the data since it was already stated that the envelope fitness values are similar. This fall of prediction capabilities is then assumed to be a consequence of poor physiology modeling of the subcutaneous route, and the less information present in the data for identification, especially in the insulin input data, which now only comes from the piecewise function of insulin infusion through the pump.

The results of the cross-validation study are more clearly displayed in figure 10.2.

Only mean values are plotted since comparison is better applied to average values where outliers are more easily filtered.

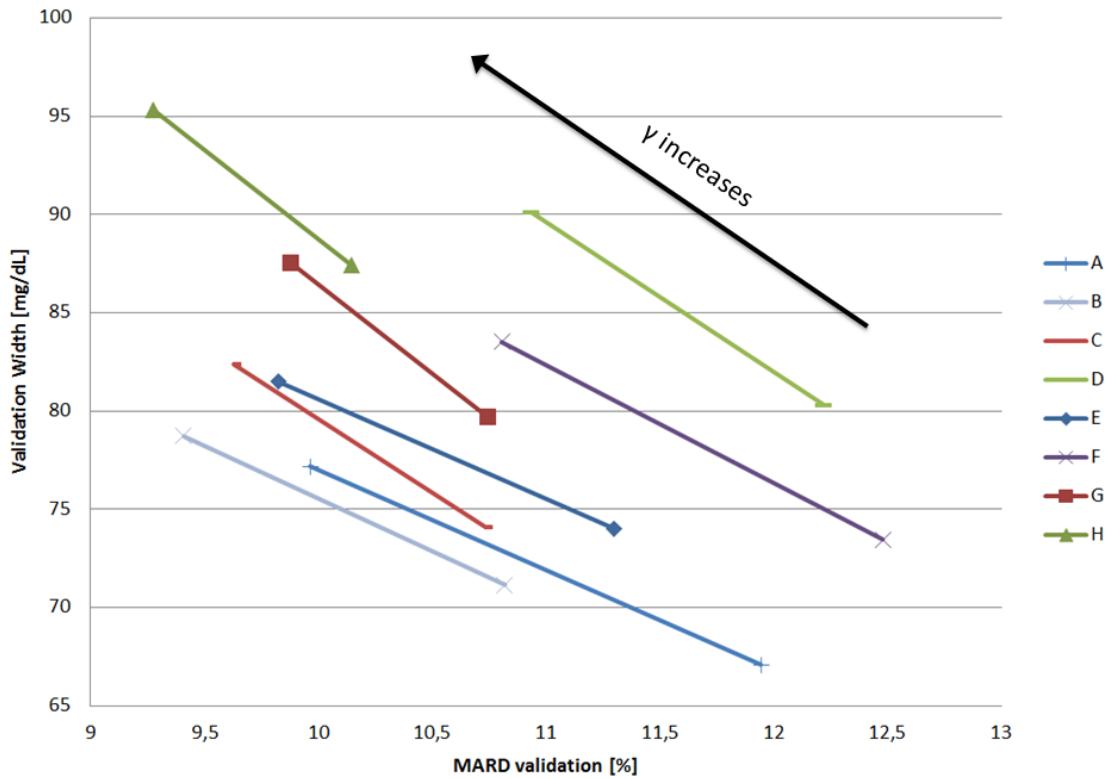


Figure 10.2: Representation in the validation Pareto Space of the results for the identification of all the scenarios considered. Two identifications for different  $\gamma$  value are displayed, and the approximate direction of the  $\gamma$  increment is plotted.

It is clear that some combinations of parameters yield better predictions than others.  $\gamma$  gradient clearly displaces the validation results towards the top left direction, even though the results displayed do not correspond to the direct application of the weighting index, but rather the validation of those. Out of each pair of points in the graph, the top left point correspond to the  $\gamma = 100$  experiment of each identification scenario, and the  $\gamma = 50$  is located in the bottom right of the pair. It is worth noting that two points hardly define a pareto front and as such, no assumptions of the optimality of one scenario or the other must be taken by looking at the pair of points as a whole. Each scenario, taking

into account the parameter vector identified and  $\gamma$  must be regarded separately for the selection of a better case in the full model identification paradigm.

Scenario B seems at a first glance to be closer to the ideal solution than the rest of the scenarios (being the ideal solution the best of both objectives for all the individuals). However, this scenario presents a much larger env\_fit value than the other solutions close to the ideal (mean env\_fit B<sub>100</sub>=18.73>A<sub>100</sub>=16.24 p<0.005 and B<sub>100</sub>=18.73>C<sub>100</sub>=16.95 p=0.041). Other solutions close to the ideal point are the scenarios A and C. All of this cases include the identification of the parameter  $\alpha$ , thus reinforcing the hypothesis drawn in the pilot study on the importance of this parameter.

The ruled out scenario B considers the newly introduced parameter  $\beta$  for the identification of the subcutaneous insulin subsystem. This parameter is identified in almost every patient and permutation to the optimization boundary, hinting a problem with identifiability even though the formal identifiability analysis proved it sensible for identification. This problem with parameter  $\beta$  is not only a matter of scenario B; scenario A, which also identifies parameter  $\beta$ , identifies this aggregated parameter to the optimization boundary 87% of the experiments. Simulations of the identified sets of parameters proved that plasma insulin levels were being identified too low for the data available.

In order to finally discard scenario A from identification feasibility, a comparison to the competing scenarios is to be performed on the prediction capabilities. Envelope width for the validation days is smaller of the scenario A<sub>100</sub> (77,1 mg/dL) than in scenario C<sub>100</sub> (82,4 mg/dL) with a significance p<0.005. However, no difference in the envelope fitness is appreciated (mean env\_fit A<sub>100</sub>=16.24>C<sub>100</sub>=16.95 p=0.078) meaning that no overestimation is assigned into the width difference. Also, MARD values are very similar, with no significance (p=0.226) on the error of the scenario C<sub>100</sub> being better than scenario A<sub>100</sub>.

Scenario C<sub>50</sub> can be also taken into consideration as the optimal solution. Every comparison of the metrics of validation for the identifications results statistical significative when comparing same parameter vectors but different  $\gamma$  values. Therefore scenario C<sub>50</sub> can be compared to scenario C<sub>100</sub> taking into account that envelope fitness and widths of the simulations of the validation days is lower, but the MARD and predicted samples are worse. A  $\gamma$  value of 100 is preferred for coherency sake with the pilot study, where the weighting factor was the same, and considering that the quality of the fitting data is the same, an equal value of  $\gamma$  is preferred for comparison purposes.

Considering finally scenario C100, prediction capabilities are very good, with an average MARD of 9.63% (median 4.34%). These MARD values are consistent with the previous findings and represent a degree of error equal or lower than those of CGM machines in the current market. gMARD metric for scenario C100 displays an average of 11.24% (median 4.59%) posing relatively small danger to the patient.

Using scenario C100, patients that were considered good identifications on the pilot study do not behave as well in the case of introducing the subcutaneous insulin subsystem. That is the case of patient 9, which was reported as an example of good identification independently of the permutations in the pilot study, but when plasma insulin is removed of the identification it does not work as well. Patient 9 identification can be seen in figure 10.3.

By looking at the permutations of patient 9, can be clearly seen that plasma insulin dynamics in the pilot study introduced trends in the identification of glucose that helped the model to predict the patient easier. The subcutaneous model of insulin absorption can simulate very simple dynamics (even with uncertainty), but physiology of the subcutaneous route is more complex. Loss of predictability with respect to the pilot study is attributed to this lack of dynamic modelling.

Nevertheless, patient predictability is considered to be very good in most of the patients. In figure 10.4, simulation of all the permutations available for patient 1 is displayed.

Patient 1 is considered to be an average case of identification for this set of identifications. Two of the permutations show very good prediction capabilities, validating day 1 and day 2. The other two permutations fail at capturing the full variability of the patient and over (day 3) or underestimate (day 4) the postprandial response of the patient almost completely. The identification performed using the second permutation is considered the best case for this patient.

Best cases metrics of all the scenarios are presented in table 10.4. Best cases are calculated by selecting the best gMARD permutation for each patient.

As it happened with the pilot study, there is always a minimum of 1 permutation per patient that predicts with accuracy the validation's day postprandial response. Best cases statistics are calculated from the 12 best days of the 12 patients under study, and being such a smaller population, are much more sensible to outliers. Despite this, it can be observed again that widths are always larger for the best case of each scenario than for

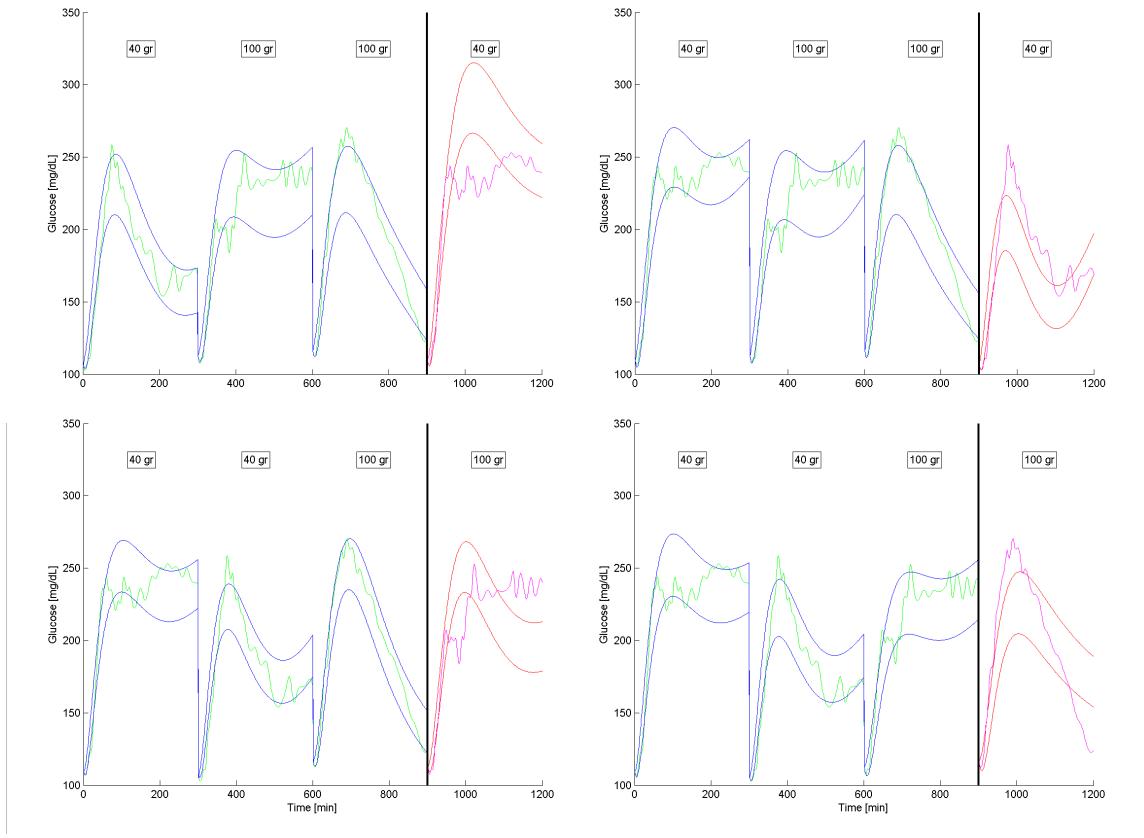


Figure 10.3: Patient 9 identification and validation results for the scenario C100. Day 1 validation is shown top left. Top Right graph shows validation for day 2. Bottom left validates day 3. Bottom right shows validation for day 4.

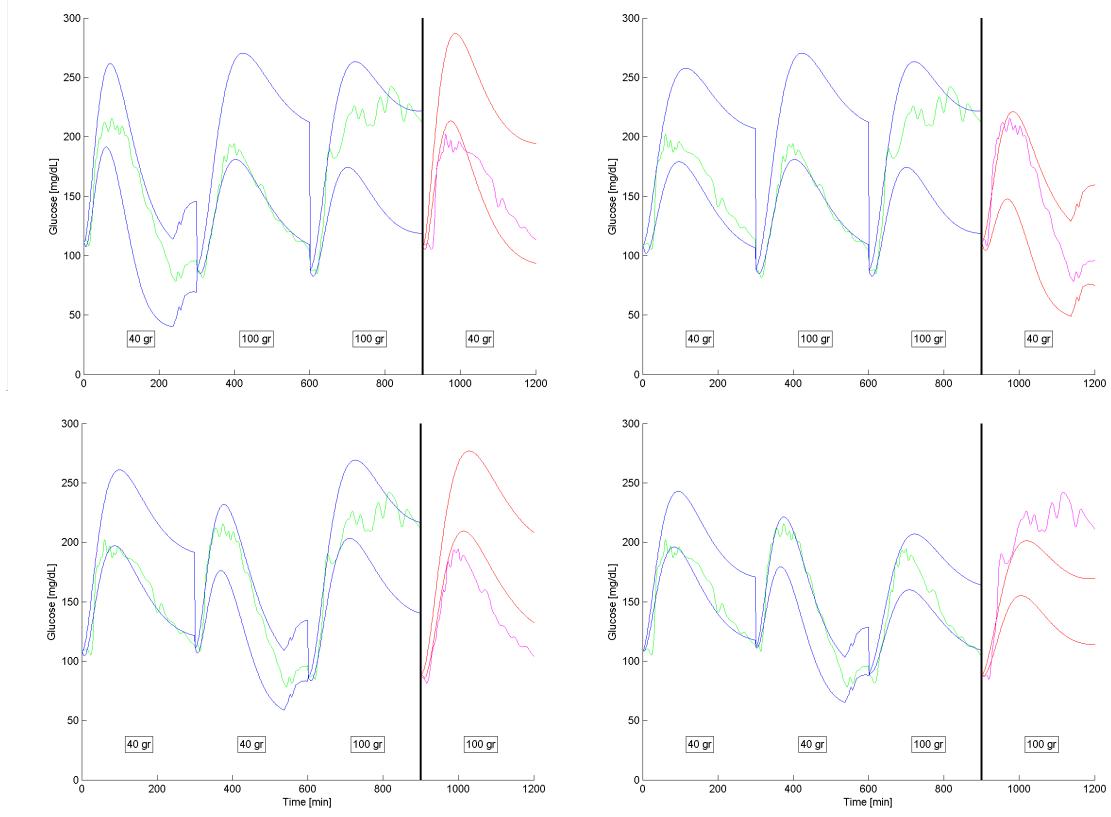


Figure 10.4: Patient 1 identification and validation results for the scenario C100. Day 1 validation is shown top left. Top Right graph shows validation for day 2. Bottom left validates day 3. Bottom right shows validation for day 4.

Scenario		A	B	C	D	E	F	G	H
$\gamma = 100$									
env_fit	mean	17.8	23.8	21	21.3	19.7	22	21	22.9
	median	18.4	17.6	19.7	17.3	18.7	18.4	20.9	18.9
Width	mean	93.6	107.8	114.5	112.4	100.2	102	104.6	113
	median	93.2	95.6	99.1	113	97.1	103.7	98.9	120.9
Prediction	mean	77.7	83.6	84.6	82.6	82.8	86.8	87.8	85.9
	median	86.2	88.2	94.5	91.3	89.5	92.3	92.8	88.8
MARD	mean	1.26	0.96	1.25	1.25	0.83	0.83	0.66	0.73
	median	0.65	0.46	0.11	0.53	0.46	0.29	0.12	0.28
gMARD	mean	1.43	0.99	1.27	1.28	0.89	0.84	0.73	0.8
	median	0.72	0.5	0.11	0.58	0.48	0.29	0.12	0.29
$\gamma = 50$									
env_fit	mean	16.7	19.4	18.4	19.7	19.6	19.3	19.4	20.9
	median	16.3	15.2	16.3	15.5	18.5	15.7	19.4	17.8
Width	mean	83.4	89.8	99.3	100.9	94.7	94.8	96.6	105.3
	median	85.2	85.6	90.5	98.4	93	96.7	94.3	111.6
Prediction	mean	72.2	79.4	79.6	74.1	81.9	82.1	84.9	82.8
	median	83.2	86.8	85.5	85.7	86.7	87.2	92	86.7
MARD	mean	1.81	1.35	1.6	1.96	0.99	1.05	0.91	0.99
	median	0.85	0.68	0.75	0.78	0.59	0.55	0.21	0.57
gMARD	mean	1.91	1.42	1.61	2.02	1.08	1.18	0.98	1.10
	median	1.05	0.83	0.81	0.79	0.67	0.55	0.22	0.63

Table 10.4: Results for all the scenarios for best case permutation of each patient.

the full cross-validation study. gMARD is obviously smaller since it is the decision metric to select each permutation to be the best case. MARD is also very small, at around 1% error for every scenario.

A representative case of the best case hypothesis is presented in figure 10.5.

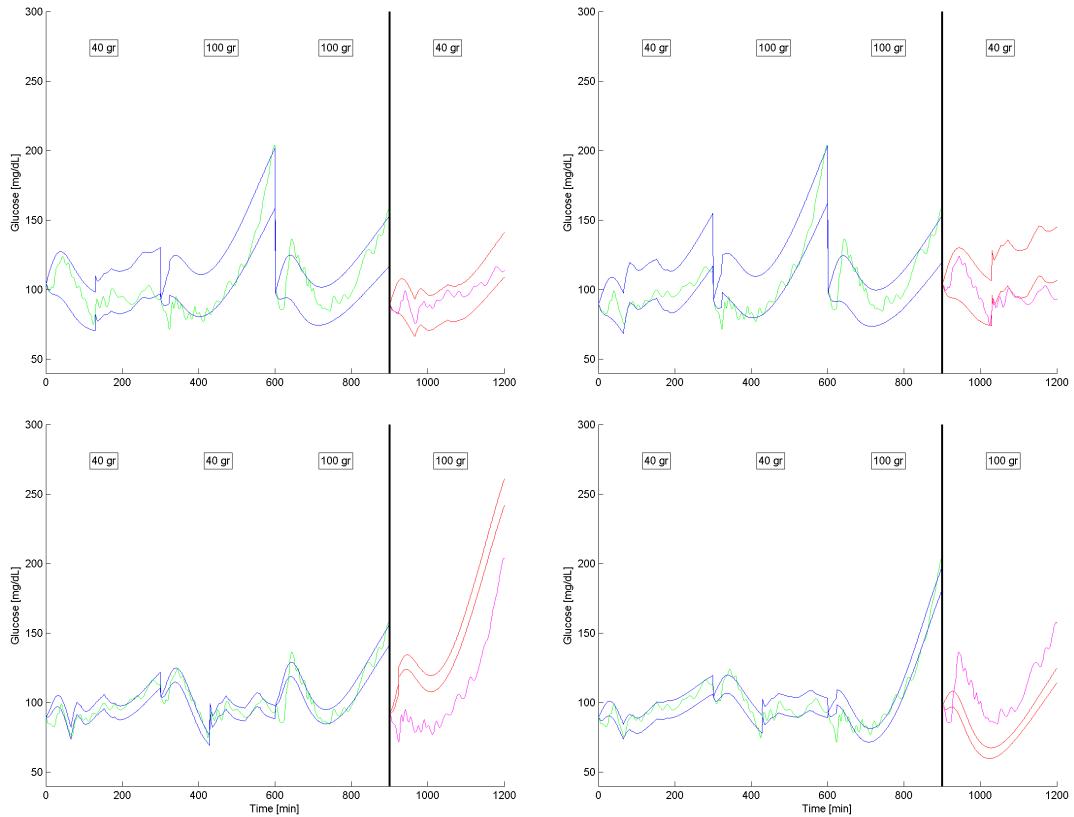


Figure 10.5: Patient 6 identification and validation results for the scenario C100. Day 1 validation is shown top left. Top Right graph shows validation for day 2. Bottom left validates day 3. Bottom right shows validation for day 4.

In patient 6 case only 1 day is well validated while the rest of days show no prediction capabilities. Note that patient 6 presented tendency to hypoglycemia at the moment of the monitoring, and IV glucose infusion was delivered in days 1 and 2. This intravenous glucose infusion prevented the patient to go into hypoglycemia but introduced a new perturbation to be modeled and identified. While the choice of this hypoglycemia prevention method was to use IV glucose infusion in order to minimize the influence of

further glucose absorption dynamics (e.g intestinal absorption), the endogenous model still may use this information in the parameter tuning routine. This is the case on the validation of day 2, where glucose infusion was delivered throughout the whole postprandial period. If this infusion is left out of the identification days, as in the second permutation, the validation is much more difficult than in case of including this information, as happens in permutation 1, even though envelope validation widths may be larger (validation width permutation 1 = 32 mg/dL vs 39.2 mg/dL in permutation 2). Also, days 3 and four represent clearly the extremes of variability of the patient, given that when both are included in the identification set, the width is much larger.

The hypothesis of best cases introduced in the pilot study seems to be valid in this case as well. MARD in the best cases is virtually zero and there is no difference between scenario. Concerning scenario C100, which has been chosen as the optimal parameter set, widths associated to the best case are larger (mean 114.5 mg/dL) than the widths for the whole dataset (mean 82.4 mg/dL) and this difference is significant with a factor  $p=0.0272$ . This does not translate however into a poor fitting if the identification data, since envelope fitness is not significantly different in the best case set and in the whole dataset (mean best case env\_fit=21 > cross-validation env\_fit=17;  $p=0.074$ ). gMARD difference for the best cases in comparison to the MARD is virtually non-existent, posing these identification no danger to the patient's safety.

The number of patients available is a very important issue and should be focus of the discussion in this type of analysis. It is clear that 12 patients is hardly enough to draw conclusions in a matter that affect very large populations, as is the case of diabetes modelling. It is very difficult however to find dataset open for research groups to access and experiment with. The reason for this conservative behavior from large research groups is of course that gathering real patient data from type 1 diabetic patients is extremely difficult and costly. Considering different sample sizes, some of the trends exposed in this thesis may become significant where now are not, or vice-versa.

However, interval identification poses a major tool for characterization of patient's postprandial behavior, even when great difficulties are tested against it such as the uncertainties of daily life in a diabetic patient, or the lack of reliable models for subcutaneous insulin infusion. This methodology is yet to be tested to one of the greatest difficulties in the process of diabetes modeling: the continuous glucose monitor signals, which will be accounted for in the following lines.

## 10.3 Identification from CGM

CGM error has been proven a big disturbance in the measurements of glucose, and ultimately one of the main problems in full patient identification in diabetes. Previous effort were performed on this thesis in analysis and modelling of the error signal produced by this type of devices (see chapter 6), proving very useful for simulation purposes and for virtual identification studies. The same CGM signals used for the deduction of the models for the Dexcom® SEVEN® PLUS and the Medtronic® Paradigm® Veo™ were included in the patients dataset for the final identification procedure.

By default the signal from the SEVEN® PLUS monitor was used for the identification of the whole dataset, given that there were more postprandial periods provided by this monitor than with the Paradigm® Veo™, and also regarding at the results of the error analysis performed before, were the monitor from Dexcom provided significantly better glucose estimations for this particular dataset.

Data fitting methodology is exactly the same as in the works with YSI reference glucose measurements. From all the scenarios considered in the previous chapter, only the settings of scenario C are used for CGM data fitting, given that the results with YSI hinted a superior capacity of prediction of data samples than in the other configurations. The final set of parameters identified, separating the interval parameters as their upper and lower bounds, is as follows:

$$\theta = [\underline{S_{iT}}, \overline{S_{iT}}, \underline{S_{iD}}, \overline{S_{iD}}, \underline{S_{iE}}, \overline{S_{iE}}, \underline{k_{12}}, \overline{k_{12}}, \underline{k_e}, \overline{k_e}, \underline{\alpha}, \overline{\alpha}, t_{maxI}, t_{maxG40}, t_{maxG100}, V_g] \quad (10.2)$$

The total number of parameters being identified is 16, including 6 interval parameters.

The weighting factor  $\gamma$  of the identification index is considered to be an estimation of the researcher of the confidence that is entrusted in the data used for the fitting. CGM is known to introduce significant errors in the glucose profiles of diabetic patients, especially in the postprandial period, and this is logically translated in a mistrust attitude to this type of data.  $\gamma$  is then supposed to be lower than the values used for the identification of YSI data, but the values to be used is not known. In here, 6 different values of  $\gamma$  are considered in order to create an approximately continuous front in the pareto space from where to extract a desired scenario which related to a defined  $\gamma$  value. The six values of  $\gamma$  are, decreasing from the value of the YSI identification: 100, 85, 70,

55, 40, 25.

CGM influence on the identification experiment is enormous. Out of the 12 patients available in the dataset, 2 were discarded due to CGM malfunctions. An example of CGM malfunction that nullifies the patient's identification is displayed in figure 10.6, corresponding to patient 7.

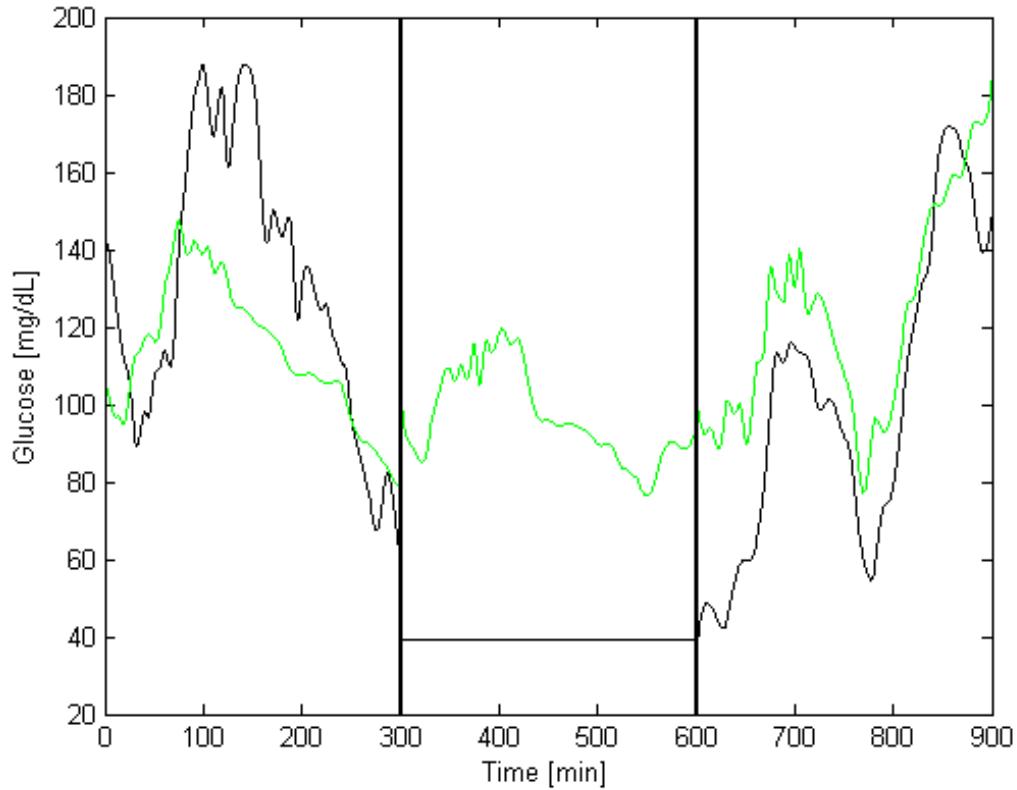


Figure 10.6: One of the permutations of patient 7 where two of the CGM monitoring signals are faulty and start. YSI signal is plotted in green and CGM estimation is plotted in black.

Even though the CGM devices were calibrated before the experiment and new sensors were inserted the day before to the experiment, several cases of monitors starting the experiment in saturation levels are reported (lower saturation level = 40 mg/dL). Examples of this faulty behavior are patients 7, 9 and 11. All of these patients wore 2 different sets of CGM as explained before. Unfortunately, for patients 7 and 11 both

monitors failed in the same days, yielding the data for those patients useless for the cross-validation study. Possible reasons for this concurrence of faults in different sensors could be misunderstanding of the use of the sensors by the patients, where sudden movements or external factors can displace the inserted subcutaneous sensor leading to loose of the CGM signal. It is worth remembering that all the patients were successfully trained in the use of CGM and were long-term users of the CGM technology.

Patient 9 was saturated in one postprandial period for the Dexcom® SEVEN® PLUS monitor, but Medtronic® Paradigm® Veo™ performed correctly in that day. Patient 9 CGM data was modified to include the use of the Paradigm® Veo™ in the identification for the postprandial period where the SEVEN® PLUS failed. The final resulting set of patients is then reduced to 10.

The evaluation of the identifications is performed following the same metrics that were applied in the previous chapter, with different application focus. CGM signal is the data to be fit in the optimization process, and all the identification metrics and the identification composite index are evaluated in the CGM data of each patient. Therefore, performance of the methodology in the identification data for each permutation of the cross-validation study is applied to the corresponding CGM data. However, for prediction of the glucose behavior, it is the “real” blood glucose which wants to be predicted, not the CGM estimation. YSI data is thus the correct data to be used for the evaluation of the performance metrics in the validation days. It is worth noting that envelope fitness measurements only makes sense if evaluated to the same data used to fit, therefore the env\_fit values are always reported on CGM data.

Computation cost on the CGM optimization was very similar to that of the YSI optimizations. Computation times for the algorithm used in this work are highly dependent on the number of parameters being tuned, and given that the number of parameters is fixed to 16, no difference between all the optimizations reported here is appreciated. Average optimization time was approximately 60 hours on a workstation Intel® Xeon® CPU 2.67 GHz with 4 GB of RAM memory running under Windows 7.

Results of the identification days for all the possible permutations are shown in table 10.5.

Fitting statistics and envelope widths are worst than those obtained from YSI identification, as expected. When looking at the  $\gamma = 100$  case, for sake of comparison, average width for the identification days goes from 86.9 mg/dL in the YSI identification to 104.8 mg/dL. env\_fit is also worsened going from 17 mg/dL up to 22.5 mg/dL. This results

$\gamma$		100	85	70	55	40	25
Width	mean	104.8	102.4	99.5	94.1	88.2	76.4
	median	96.5	94.2	90.6	85.5	80.1	69.5
Prediction	mean	77.4	75.6	73.4	69.4	65.1	56.6
	median	79.9	78	75.8	72.7	68.6	60.9
MARD	mean	1.24	1.37	1.59	2	2.53	3.69
	median	1.11	1.24	1.43	1.74	2.31	3.34
gMARD	mean	1.67	1.85	2.16	2.73	3.44	4.93
	median	1.49	1.67	2.07	2.39	3.21	4.65
env_fit	mean	22.5	22	21.5	20.2	19	17.7
	median	21	20.2	19.2	18	16.1	15.9

Table 10.5: Results for all the  $\gamma$  values for the three identification days considering all possible permutations of days from the dataset.

in an overestimation of the width of the interval model, responding to the much noisier data source of the CGM. MARD is also larger (note that in identification data MARD is evaluated against CGM), going from 0.88% of the YSI identification up to 1.24%, even though prediction samples are very similar (YSI 77.3% vs CGM 77.4%).

Validation data is displayed in table 10.6.

$\gamma$		100	85	70	55	40	25
Width	mean	101.2	99.6	95.9	90.1	84.8	73
	median	91	84.5	85.6	79.6	73.9	62.4
Prediction	mean	61.1	60.7	58.4	54.8	54.5	48.3
	median	69.3	67.8	62.5	57.7	61.8	49.7
MARD	mean	9.6	9.68	10.04	11.16	11.8	14.21
	median	4.32	4.68	5.14	6.1	6	7.07
gMARD	mean	11.63	11.75	12.17	13.38	14.14	16.74
	median	4.35	4.71	5.18	6.53	6.34	8.16

Table 10.6: Results for all the  $\gamma$  values for the validation day considering all possible permutations of days from the dataset.

Calculation of significance between the data of the CGM identifications and YSI identifications is not straightforward. Given that CGM identifications are only performed on 10 patients, paired permutation significance tests are not possible, unless the faulty CGM patients are removed from the YSI dataset. It is also possible to perform unpaired

significance tests, but it is more reasonable to use the information available from the patient to perform the test. To do so, patients 7 and 11 are discarded, and the permutations tests performed as usual.

Validation widths for the  $\gamma = 100$  scenario of the CGM identification are larger than those of the identification with blood glucose reference ( $101.2 \text{ mg/dL} > 82.4 \text{ mg/dL}$ ;  $p < 0.005$ ), but MARD values are very similar (YSI MARD = 9.63% vs. CGM MARD = 9.6%). Only  $\gamma$  value of 100 is tested for this comparison, because it is the same as in the YSI experiment. Results for different  $\gamma$  values are displayed in figure 10.7 for easier visualization.

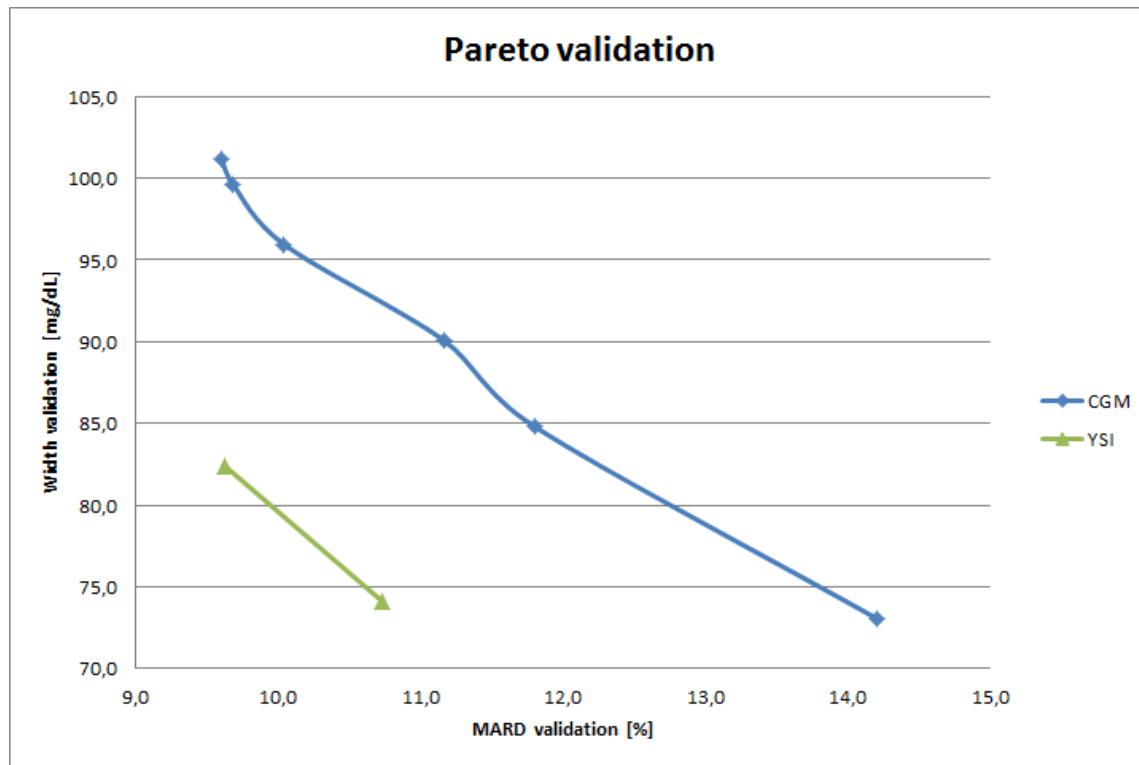


Figure 10.7: Representation in the validation Pareto Space of the results for the identification of all the  $\gamma$  values considered.

The range of  $\gamma$  considered seems appropriate for the problem at stake, since the range covers from the same MARD level of the YSI identification (MARD=9.6% at  $\gamma=100$ ) up to very similar envelope widths from both identifications at  $\gamma = 25$  (YSI width 74.1 mg/dL vs CGM width 73 mg/dL). Looking at the whole set of identifications,  $\gamma = 70$

is closer in the pareto space to the much better solution that YSI identification poses than any other identification performed by CGM. Also, width decreases significantly from  $\gamma = 100$  to  $\gamma = 70$  for very similar MARD results, which makes this solution to be preferred. For lower  $\gamma$  values, improvements on the envelope width come at the cost of dramatically increasing the average MARD for the whole dataset. All simulations onward are presented for the  $\gamma = 70$  scenario.

An example of identification with CGM is shown in figure 10.8.

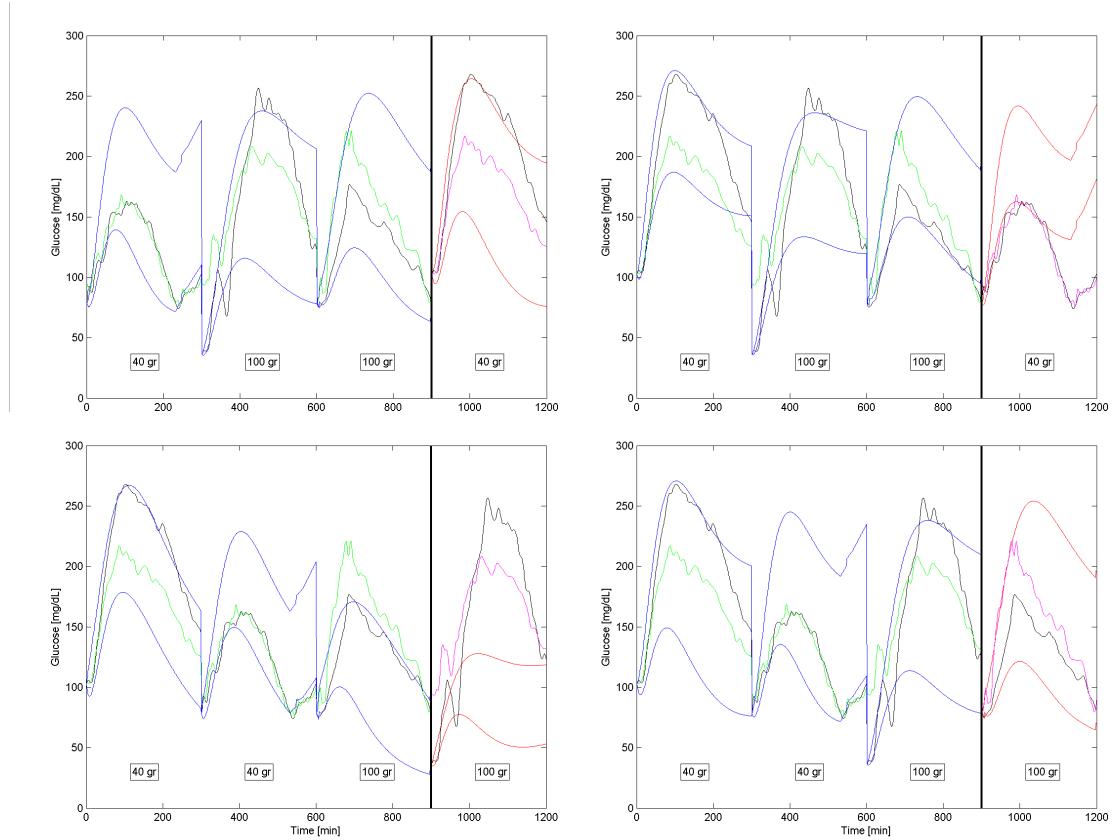


Figure 10.8: Patient 5 identification and validation results for the scenario C100. Day 1 validation is shown top left. Top Right graph shows validation for day 2. Bottom left validates day 3. Bottom right shows validation for day 4.

Patient number 5 shows very clearly the inconvenient of CGM identification. Even though the CGM was not considered faulty for the postprandial periods of this patient, several days present clear over or underestimations of the blood glucose signals. This

behavior is very common on any CGM device and is part of the assumed error of the sensor's estimations. It does, however, affect greatly on the identifications performed. Disregarding the possibilities of inherent variability of the patient, it can be observed how overestimation of the signal causes the validation of the model to also overestimated the blood glucose reference signal. For example, in the validation of day 2 for patient 5, in two of the days used for identification the CGM rises approximately 50 mg/dL over the blood glucose reference, causing the band with to grow much larger and misplaced in the patient's average postprandial behavior. This problem is diminished in situations like the validation of day 1, where overestimation and underestimation occur in separate days of the identification set, causing the envelope width to be larger than required, but centered in the average postprandial behavior of the patient, thus including the validation data within the envelope. The larger envelopes identified in average for the CGM identification are explained by the performance of the identification of patient 5, which is a representative sample of an identification for this chapter.

Patient 9, for whom the Dexcom® SEVEN® PLUS monitor presented flat estimations, also is a very good example of identification with CGM. The identification performed for patient 9 is plotted in figure 10.9.

This patient supposes a good identification for all possible identifications paradigms: it was the best example of identification for the pilot study, and although the introduction of the subcutaneous insulin model in the identification scenario worsened the outcome of the optimizations, it was yet a good example of identification with YSI for a full model of a patient. CGM performs reasonably well for this patient leading to a very similar performance in the identification to that obtained from YSI identification.

Permutations 1 and 2 (validation days 1 and 2) cover completely the validation data on the model's envelope. Permutation 3 does not present a good validation even though the data fitting is satisfactory. The last permutation represents a failed validation but also a poor fitting of the glucose signal in day 3 mostly caused by a bad CGM estimation.

The final example is patient 6, whose identification is displayed in figure 10.10.

Patient 6 can be categorized as a patient with very low variability. Envelope widths are much lower than with the rest of the patients. Also, glucose excursions for this patient are smaller than the rest of the dataset. This permits a clearer observation of the best case scenario of the patient. CGM readings are satisfactory for all the postprandial periods of this patient, resulting in good identification for the CGM identification too. The first permutation is the best example of satisfactory identification, where all the

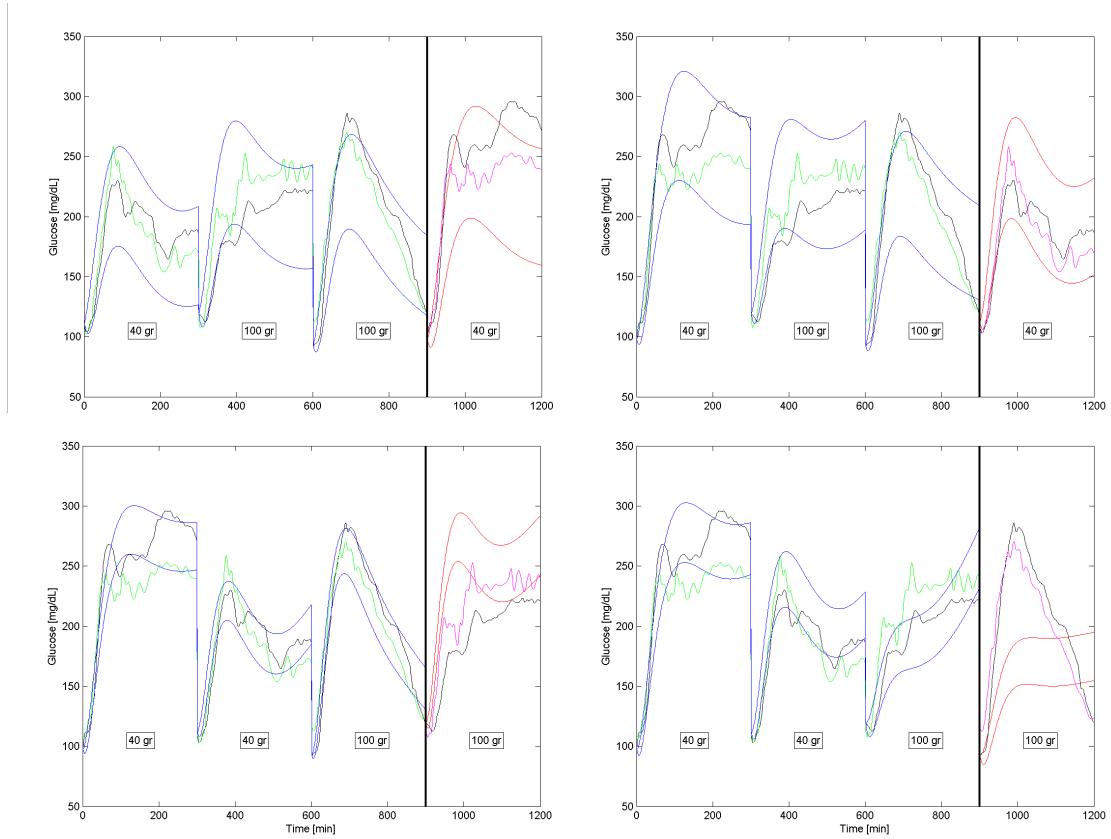


Figure 10.9: Patient 9 identification and validation results for the scenario C100. Day 1 validation is shown top left. Top Right graph shows validation for day 2. Bottom left validates day 3. Bottom right shows validation for day 4.

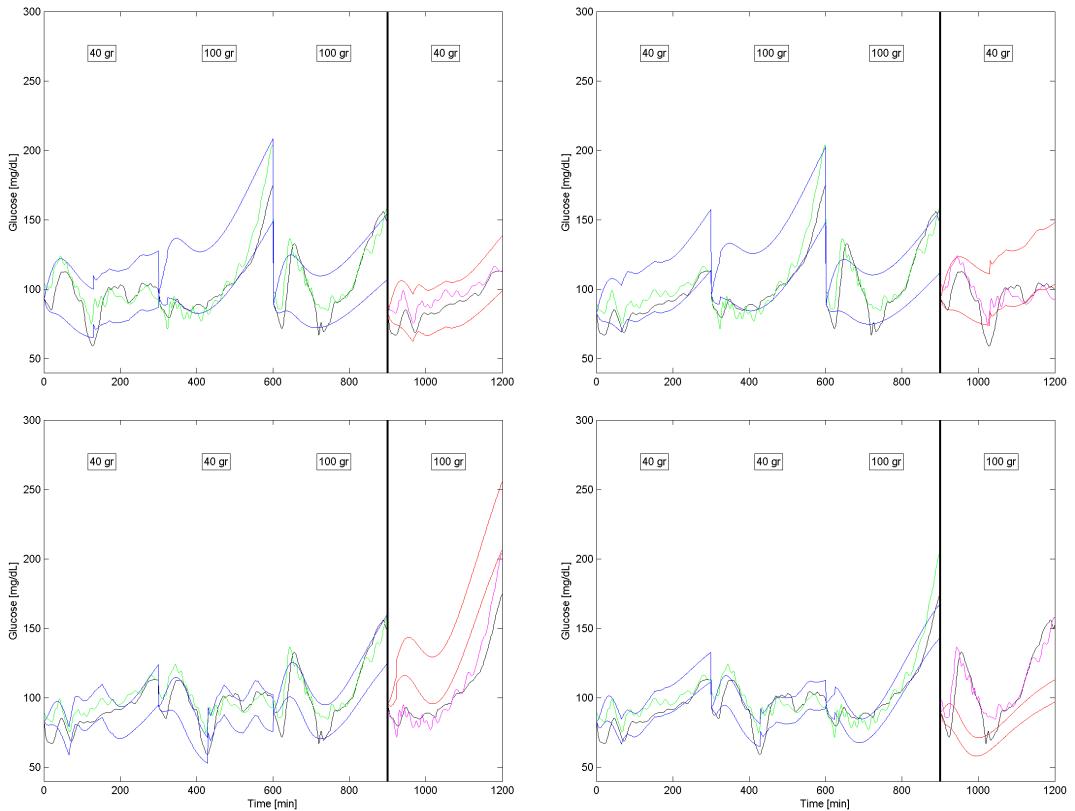


Figure 10.10: Patient 6 identification and validation results for the scenario C100. Day 1 validation is shown top left. Top Right graph shows validation for day 2. Bottom left validates day 3. Bottom right shows validation for day 4.

prediction band envolves the YSI measures, even though the initial point of the model simulation is taken from the CGM signal (it is the only known data to the model). Permutation 2 is a very similar case of very good prediction, but in this case it is improved by the larger envelope widths caused by CGM identification, because this particular permutation of this patient did not perform as well with YSI identification due to the small width (therefore small variability) of the patient's identification days. The two last permutations' validation days are a clear example of both patient's extremes of variability within the 4-day experiment, falling both of them out of the prediction bands

Patient's 6 example hints that the best case permutation hypothesis that was proved valid in the previous chapters is still possible for the CGM identifications. Best case validation metrics and envelope fitness measure are listed in table 10.7.

$\gamma$		100	85	70	55	40	25
env_fit	mean	27.3	26.8	25.8	25.4	24.9	22.5
	median	22.1	21.8	20	19.4	20	16
Width	mean	130.4	127	122.5	112	112.3	91
	median	115.4	110.8	113.1	102.9	95.1	67.6
Prediction	mean	90.9	90.6	88.3	82.1	89.9	80.1
	median	97.8	97.3	96.7	97.2	97	96.2
MARD	mean	0.58	0.59	0.74	1.82	0.81	2.06
	median	0.14	0.15	0.19	0.19	0.2	0.27
gMARD	mean	0.59	0.6	0.75	1.83	1.02	3.15
	median	0.14	0.15	0.19	0.19	0.2	0.27

Table 10.7: Results for all the  $\gamma$  values for the best case permutation in each patient.

MARD measures are very small, comparable in magnitude with those of the previous chapters. gMARD difference form the MARD counterpart is no relevant and supposes no danger to the patients. Average predicted samples is approximately 90% for all  $\gamma$  values considered, therefore consideration of full coverage of the validation days for the best case identification is valid.

When comparing the metrics of the best cases with the whole dataset, even though widths appear to be larger for the best cases ( $122.5 \text{ mg/dL} > 95.95 \text{ mg/dL}$ ) this difference is actually no significative ( $p=0.109$ ) because of the small number of best case samples available. It was discussed before in this chapter how the small number of patients is one of the main problems for the validation of the methodology at stake in this thesis. This problem gets aggravated by the removal of the two patients whose CGM was faulty,

effectively reducing the samples size 16%. Significative differences in the metrics are much more difficult to find in smaller samples of the diabetic patient's population. Nevertheless, envelope fitness of the best cases at stake was not different from that of the whole dataset ( $21.75 \text{ mg/dL} > 21.49 \text{ mg/dL}$ ;  $p=0.457$ ), validating the fitting process.

## 10.4 Discussion

Full model patient identification considering uncertainty has been performed in this chapter, from both YSI and CGM data sources, displaying the decreasing prediction capabilities of the models involved as the complexity of the system is increased and the data source loses quality (increases the error associated).

For the identification including the subcutaneous insulin route, 8 different scenarios were tested for identification feasibility composing different parameter vector to be used in the identification. Even though all the parameters in the model were identifiable following the theoretical identifiability tests performed, parameters were left out of the identification sequentially in order to test the practical identifiability of the parameters, along with the prediction capabilities of the model with each combination of parameters. Only parameter  $\beta$ , corresponding to the aggregated error parameter of the subcutaneous insulin route, was discarded due to grave identifiability problems.

The inclusion of CGM supposed a great set-back for the prediction power of the models identified. The width associated for the selected optimum scenario was in average 95.9 mg/dL (median 85.6 mg/dL), far away of the average 82.4 mg/dL (median 72.7 mg/dL) necessary for the YSI study that included the subcutaneous route of insulin. When these figures are compared to the 78.8 mg/dL (median 66.4 mg/dL) width obtained in the pilot study, it is possible to view the error introduced by the CGM devices (difference of 13.5 mg/dL in comparison to the model mismatch introduced by the subcutaneous insulin route of 3.6 mg/dL).

Prediction error increased with the inclusion of the subcutaneous route and the CGM measures. Starting with a MARD of 7.61% (median 5.56%) in the pilot study, this error metric increased up to 9.63% (median 4.34%) with the subcutaneous insulin route inclusion in the model, and even larger error was reported for with the inclusion of continuous monitoring technology, up to 10.04% (median 5.14%). When comparing the increment of envelope width and MARD, it can be appreciated that most of the uncertainty from the CGM inclusion is translated into envelope width by the prediction

algorithm, while the inclusion of the input model for subcutaneous insulin contributes to the methodology by increasing the MARD. This is a consequence of the choice of a  $\gamma = 100$  for both experiments involving YSI data, weighting the envelope width similarly in the optimizations for both experiments.  $\gamma$  value was smaller for the CGM experiment, giving more importance to error minimization and therefore increasing the envelope width.

Finally, the number of patients is the great limitation to the findings here exposed. CGM errors and model mismatches suppose a great challenge to the identification techniques available, introducing relevant noises in the final performance metrics used for evaluation of the identification methodologies, which will only be palliated if large databases of patients' information are available to the scientific community.

# **Chapter 11**

## **Conclusions**



# Bibliography

- [1] Crisostomo Alberto Barajas Solano. Identificación bayesiana de los modelos de absorción de glucosa para comidas mixtas. Master's thesis, Departamento de Eléctrica, Electronica y Automática. Universidad de Girona, 2009.
- [2] R. Basu, B. Di Camillo, G. Toffolo, A. Basu, P. Shah, A. Vella, R. Rizza, and C. Cobelli. Use of a novel triple-tracer approach to assess postprandial glucose metabolism. *Am J Physiol-Endoc M*, 284(1):E55, 2003.
- [3] M.S. Bazaraa, H.D. Sherali, and CM Shetty. *Nonlinear programming: theory and algorithms*. John Wiley and Sons, 2006.
- [4] RN Bergman, LS 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. *Journal of Clinical Investigation*, 68(6):1456, 1981.
- [5] G.E.P. Box and G.M. Jenkins. *Time series analysis: forecasting and control*. Prentice Hall PTR, 1994.
- [6] M. Breton and B. Kovatchev. Analysis, modeling, and simulation of the accuracy of continuous glucose sensors. *J Diab Sci Technol*, 2(5):853–862, 2008.
- [7] B.A. Buckingham, C. Kollman, R. Beck, A. Kalajian, R. Fiallo-Scharer, M. Tansey, L.A. Fox, D.M. Wilson, S.A. Weinzimer, K.J. Ruedy, and W.V. Tamborlane. Evaluation of factors affecting cgms calibration. *Diabet Technol Ther*, 8:318–325, 2006.
- [8] R Calm, M García-Jaramillo, J Bondia, MA Sainz, and J Vehí. Comparison of interval and monte carlo simulation for the prediction of postprandial glucose under uncertainty in type 1 diabetes mellitus. *Comput Meth Prog Bio*, 104(3):325–332, 2010.

- [9] R Calm, M Garcia-Jaramillo, J Vehí, J Bondia, C Tarin, and W Garcia-Gabin. Prediction of glucose excursions under uncertain parameters and food intake in intensive insulin therapy for type 1 diabetes mellitus. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 1770–1773. IEEE, 2007.
- [10] Fraser Cameron, Günter Niemeyer, and B Wayne Bequette. Extended multiple model prediction with application to blood glucose regulation. *Journal of Process Control*, 22(8):1422–1432, 2012.
- [11] Marzia Cescon, Fredrik Ståhl, Mona Landin-Olsson, and Rolf Johansson. Subspace-based model identification of diabetic blood glucose dynamics. *Insulin*, 2(4):6, 2009.
- [12] W.L. Clarke, S. Anderson, M. Breton, S. Patek, L. Kashmer, B. Kovatchev, et al. Closed-loop artificial pancreas using subcutaneous glucose sensing and insulin delivery and a model predictive control algorithm: the virginia experience. *J Diab Sci Technol*, 3(5):1031–1038, 2009.
- [13] W.L. Clarke, D. Cox, L.A. Gonder-Frederick, W. Carter, and S.L. Pohl. Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care*, 10(5):622–628, 1987.
- [14] C. Cobelli and K. Thomaseseth. The minimal model of glucose disappearance: optimal input studies. *Math Biosci*, 83(2):127–155, 1987.
- [15] T.F. Coleman and Y. Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6:418–445, 1996.
- [16] Harald Cramer. *Mathematical models of statistics*. Number 9. Princeton University Press, 1946.
- [17] C Dalla Man, M. Camilleri, and C. Cobelli. A system model of oral glucose absorption: validation on gold standard data. *IEEE T Bio-Med Eng*, 53(12 Part 1):2472–2478, 2006.
- [18] C Dalla Man, RA Rizza, and C. Cobelli. Meal simulation model of the glucose-insulin system. *IEEE T Bio-Med Eng*, 54(10):1740–1749, 2007.
- [19] D de Pereda, S Romero-Vivo, B Ricarte, and J Bondia. On the prediction of glucose concentration under intra-patient variability in type 1 diabetes: a monotone systems approach. *Comput Meth Prog Bio*, 108 (3):993–1001, 2012.

- [20] S. Del Favero, A. Facchinetto, and C. Cobelli. A glucose-specific metric to assess predictors and identify models. *Biomedical Engineering, IEEE Transactions on*, 59(99):1281–1290, 2012.
- [21] J.A. Egea, M. Rodríguez-Fernández, J.R. Banga, and R. Martí. Scatter search for chemical and bio-process optimization. *Journal of Global Optimization*, 37(3):481–503, 2007.
- [22] P.G. Fabietti, V. Canonico, M.O. Federici, M.M. Benedetti, and E. Sarti. Control oriented model of insulin and glucose dynamics in type 1 diabetics. *Medical and Biological Engineering and Computing*, 44(1):69–78, 2006.
- [23] A. Facchinetto, G. Sparacino, and C. Cobelli. Reconstruction of glucose in plasma from interstitial fluid continuous glucose monitoring data: role of sensor calibration. *J Diab Sci Technol*, 1(5):617–623, 2007.
- [24] A. Facchinetto, G. Sparacino, and C. Cobelli. Modeling the error of continuous glucose monitoring sensor data: critical aspects discussed through simulation studies. *J Diab Sci Technol*, 4(1):4–14, 2010.
- [25] DA Finan, FJ Doyle III, CC Palerm, WC Bevier, HC Zisser, L Jovanović, and DE Seborg. Experimental evaluation of a recursive model identification technique for type 1 diabetes. *J Diab Sci Technol*, 3(5):1192–1202, 2009.
- [26] G. Franceschini and S. Macchietto. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63(19):4846–4872, 2008.
- [27] Eleni I Georga, Vasilios C Protopappas, and Dimitrios I Fotiadis. Glucose prediction in type 1 and type 2 diabetic patients using data driven techniques. *Knowledge-Oriented Applications in Data Mining*, 2011.
- [28] N. Hansen. The cma evolution strategy: a comparing review. *Towards a new evolutionary computation*, 192:75–102, 2006.
- [29] N. Hansen and S. Kern. Evaluating the cma evolution strategy on multimodal test functions. In *Parallel Problem Solving from Nature-PPSN VIII*, pages 282–291. Springer, 2004.
- [30] K.L. Helms and K.W. Kelley. Insulin glulisine: An evaluation of its pharmacodynamic properties and clinical application. *The Annals of Pharmacotherapy*, 43(4):658, 2009.

- [31] J Herrero, M. Martínez, J. Sanchis, and X. Blasco. Well-distributed pareto front by using the  $\epsilon$ -moga evolutionary algorithm. *Computational and Ambient Intelligence*, 4507:292–299, 2007.
- [32] P Herrero, CC Palerm, E Dassau, H Zisser, L Jovanović, C Dalla Man, C Cobelli, J Vehí, and FJ Doyle III. A glucose absorption model library of mixed meals for in silico evaluation of artificial  $\beta$  cell control algorithms. In *Journal of Diabetes Science and Technology*, March 2008.
- [33] Pau Herrero, Jorge Bondia, Cesar C Palerm, Josep Vehí, Pantelis Georgiou, Nick Oliver, and Christofer Toumazou. A simple robust method for estimating the glucose rate of appearance from mixed meals. *Journal of diabetes science and technology*, 6(1):153, 2012.
- [34] R. Hovorka, J.M. Allen, D. Elleri, L.J. Chassin, J. Harris, D. Xing, C. Kollman, T. Hovorka, A.M.F. Larsen, M. Nodale, et al. Manual closed-loop insulin delivery in children and adolescents with type 1 diabetes: a phase 2 randomised crossover trial. *Lancet*, 375(9716):743–751, 2010.
- [35] R. Hovorka, V. Canonico, L.J. Chassin, U. Hauerter, M. Massi-Benedetti, M. Orsini Federici, T.R. Pieber, H.C. Schaller, L. Schaupp, T. Vering, et al. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiological Measurement*, 25:905–920, 2004.
- [36] R. Hovorka, J. Kremen, J. Blaha, M. Matias, K. Anderlova, L. Bosanska, T. Roubicek, M.E. Wilinska, L.J. Chassin, S. Svacina, et al. Blood glucose control by a model predictive control algorithm with variable sampling rate versus a routine glucose management protocol in cardiac surgery patients: a randomized controlled trial. *Journal of Clinical Endocrinology & Metabolism*, 92(8):2960, 2007.
- [37] R. Hovorka, F. Shojaee-Moradie, P.V. Carroll, L.J. Chassin, I.J. Gowrie, N.C. Jackson, R.S. Tudor, A.M. Umpleby, and R.H. Jones. Partitioning glucose distribution/transport, disposal, and endogenous production during ivgtt. *Am J Physiol-Endoc M*, 282(5):E992, 2002.
- [38] Roman Hovorka, Kavita Kumareswaran, Julie Harris, Janet M Allen, Daniela Elleri, Dongyuan Xing, Craig Kollman, Marianna Nodale, Helen R Murphy, David B Dunger, et al. Overnight closed loop insulin delivery (artificial pancreas) in adults with type 1 diabetes: crossover randomised controlled studies. *BMJ: British Medical Journal*, 342, 2011.

- [39] JN Hunt and DF Stubbs. The volume and energy content of meals as determinants of gastric emptying. *The Journal of Physiology*, 245(1):209, 1975.
- [40] Maira Alejandra García Jaramillo. *Prediction of postprandial blood glucose under intra-patient variability and uncertainty and its use in the design of insulin dosing strategies for type 1 diabetic patients*. PhD thesis, Universitat de Girona, 2011.
- [41] Luc Jaulin, Michel Kieffer, Olivier Didrit, and Eric Walter. *Applied interval analysis: with examples in parameter and state estimation, robust control and robotics*, volume 1. Springer Verlag, 2001.
- [42] A. Kamath, A. Mahalingam, and J. Brauker. Analysis of time lags and other sources of error of the dexcom seven continuous glucose monitor. *Diabet Technol Ther*, 11(11):689–695, 2009.
- [43] D.B. Keenan, R. Cartaya, and J.J. Mastrototaro. Accuracy of a new real-time continuous glucose monitoring algorithm. *J Diabetes Sci Technol*, 4(1):111, 2010.
- [44] H. Kirchsteiger, G.C. Estrada, S. Pölzer, E. Renard, and L. del Re. Estimating interval process models for type 1 diabetes for robust control design. In *Proceedings of the 18th IFAC World Congress*, volume 18, pages 11761–11766, 2011.
- [45] William C Klingensmith, Katherine L Rhea, Elizabeth A Wainwright, and Orlin Woodie Hopper. The gastric emptying study with oatmeal: reference range and reproducibility as a function of age and sex. *Journal of Nuclear Medicine Technology*, 38(4):186–190, 2010.
- [46] T. Kobayashi, S. Sawano, T. Itoh, K. Kosaka, H. Hirayama, and Y. Kasuya. The pharmacokinetics of insulin after continuous subcutaneous infusion or bolus subcutaneous injection in diabetic patients. *Diabetes*, 32(4):331, 1983.
- [47] Boris Kovatchev, Claudio Cobelli, Eric Renard, Stacey Anderson, Marc Breton, Stephen Patek, William Clarke, Daniela Bruttomesso, Alberto Maran, Costa Silvana, et al. Multinational study of subcutaneous model-predictive closed-loop control in type 1 diabetes mellitus: summary of the results. *Journal of diabetes science and technology*, 4(6):1374, 2010.
- [48] B.P. Kovatchev, M. Breton, C. Dalla Man, and C. Cobelli. In silico preclinical trials: A proof of concept in closed-loop control of type 1 diabetes. *J Diab Sci Technol*, 3(1):44–55, 2009.

- [49] B.P. Kovatchev, L.A. Gonder-Frederick, D.J. Cox, and W.L. Clarke. Evaluating the accuracy of continuous glucose-monitoring sensors continuous glucose-error grid analysis illustrated by therasense freestyle navigator data. *Diabet Care*, 27(8):1922–1928, 2004.
- [50] EW Kraegen and DJ Chisholm. Insulin responses to varying profiles of subcutaneous insulin infusion: kinetic modelling studies. *Diabetologia*, 26(3):208–213, 1984.
- [51] Alejandro José Laguna Sanz. Model identification from ambulatory data for post-prandial glucose control in type 1 diabetes. Master's thesis, Departament d'Enginyeria de Sistemes i Automàtica, 2010.
- [52] ED Lehmann and T. Deutsch. A physiological model of glucose-insulin interaction in type 1 diabetes mellitus. *Journal of Biomedical Engineering*, 14(3):235–242, 1992.
- [53] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Upper Saddle River, NJ,, 1999.
- [54] L. Magni, D.M. Raimondo, L. Bossi, C. Dalla Man, G. De Nicolao, B. Kovatchev, and C. Cobelli. Model predictive control of type 1 diabetes: an in silico trial. *J Diab Sci Technol*, 1(6):804–812, 2007.
- [55] R.S. Mazze, E. Strock, S. Borgman, D. Wesley, P. Stout, and J. Racchini. Evaluating the accuracy, reliability, and clinical applicability of continuous glucose monitoring (cgm): Is cgm ready for real time? *Diabet Technol Ther*, 11(1):11–18, 2009.
- [56] Mario Milanese and Antonio Vicino. Estimation theory for nonlinear models and set membership uncertainty. *Automatica*, 27(2):403–408, 1991.
- [57] JB Mitchell, DL Costill, JA Houmard, WJ Fink, RA Robergs, and JA Davis. Gastric emptying: influence of prolonged exercise and carbohydrate concentration. *Medicine & Science in Sports & Exercise*, 21(3):269, 1989.
- [58] RE Moore. *Methods and applications of interval analysis*, volume 2. SIAM, 1979.
- [59] E. Mosekilde, K.S. Jensen, C. Binder, S. Pramming, and B. Thorsteinsson. Modeling absorption kinetics of subcutaneous injected soluble insulin. *Journal of Pharmacokinetics and Pharmacodynamics*, 17(1):67–87, 1989.
- [60] B. Nowotny, PJ Nowotny, K. Strassburger, and M. Roden. Precision and accuracy of blood glucose measurements using three different instruments. *Diabetic Medicine*, 29(2):260–265, 2012.

- [61] G. Nucci and C. Cobelli. Models of subcutaneous insulin kinetics. a critical review. *Comput Meth Prog Bio*, 62(3):249–257, 2000.
- [62] C. Palerm, M. Rodríguez-Fernández, WC. Bevier, H Zisser, JR. Banga, L Jovanović, and FJ. Doyle 3r. Robust parameter estimation in a model for glucose kinetics in type 1 diabetes subjects. In *Proceedings of the 28th IEEE EMBS Annual International Conference New York City, USA, Aug 30-Sept 3, 2006*, 2006.
- [63] S. Panunzi, P. Palumbo, and A. De Gaetano. A discrete single delay model for the intra-venous glucose tolerance test. *Theoretical Biology and Medical Modelling*, 4(1):35, 2007.
- [64] M.W. Percival, W.C. Bevier, Y. Wang, E. Dassau, H.C. Zisser, L. Jovanović, and FJ Doyle 3rd. Modeling the effects of subcutaneous insulin administration and carbohydrate consumption on blood glucose. *J Diab Sci Technol*, 4(5):1214–1228, 2010.
- [65] M.W. Percival, E. Dassau, H. Zisser, L. Jovanovic, and F.J. Doyle III. Practical approach to design and implementation of a control algorithm in an artificial pancreatic beta cell. *Industrial and Engineering Chemistry Research*, 48:6059–6067, 2009.
- [66] MJD Powell. A fast algorithm for nonlinearly constrained optimization calculations. In *Numerical analysis: proceedings of the Biennial Conference held at Dundee, June 28-July 1, 1977*, page 144, 1978.
- [67] WR Puckett and EN Lightfoot. A model for multiple subcutaneous insulin injections developed from individual diabetic patient data. *Am J Physiol-Endoc M*, 269(6):E1115, 1995.
- [68] MJ Quon, C. Cochran, SI Taylor, and RC Eastman. Non-insulin-mediated glucose disappearance in subjects with iddm. discordance between experimental results and minimal model analysis. *Diabetes*, 43(7):890, 1994.
- [69] C Radhakrishna Rao. *Linear statistical inference and its applications*, volume 22. Wiley. com, 2009.
- [70] W. Regitnig, Z. Trajanoski, HJ Leis, M. Ellmerer, A. Wutte, G. Sendlhofer, L. Schaupp, GA Brunner, P. Wach, and TR Pieber. Plasma and interstitial glucose dynamics after intravenous glucose injection: evaluation of the single-compartment glucose distribution assumption in the minimal models. *Diabetes*, 48(5):1070, 1999.

- [71] M. Rodriguez-Fernandez, J.A. Egea, and J.R. Banga. Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. *BMC bioinformatics*, 7(1):483, 2006.
- [72] Maria Rodriguez-Fernandez, Pedro Mendes, and Julio R Banga. A hybrid approach for efficient and robust parameter estimation in biochemical pathways. *Biosystems*, 83(2):248–265, 2006.
- [73] Derrick K Rollins, Nidhi Bhandari, Jim Kleinedler, Kaylee Kotz, Amber Strohbehn, Lindsay Boland, Megan Murphy, Dave Andre, Nisarg Vyas, Greg Welk, et al. Free-living inferential modeling of blood glucose level using only noninvasive inputs. *Journal of Process Control*, 20(1):95–107, 2010.
- [74] P Rossetti, FJ Ampudia-Blasco, AJ Laguna, A Revert, J Vehí, JF Ascaso, and J Bondia. Evaluation of a novel continuous glucose monitoring-based method for mealtime insulin dosing -the ibolus- in subjects with type 1 diabetes using continuous subcutaneous insulin infusion therapy: A randomized controlled trial. *Diabetes Technol Ther*, 14 (11):10.1089/dia.2012.0145, 2012.
- [75] A. Roy and R.S. Parker. Dynamic modeling of exercise effects on plasma glucose and insulin levels. *J Diab Sci Technol*, 1(3):338–347, 2007.
- [76] G. Scheiner and B.A. Boyer. Characteristics of basal insulin requirements by age and gender in type-1 diabetes patients using insulin pump therapy. *Diabetes Research and Clinical Practice*, 69(1):14–21, 2005.
- [77] F. Ståhl and R. Johansson. Diabetes mellitus modeling and short-term prediction based on blood glucose measurements. *Math Biosci*, 217(2):101–117, 2009.
- [78] C Tarin, E Teufel, J Pico, J Bondia, and HJ Pfleiderer. Comprehensive pharmacokinetic model of insulin glargine and other insulin formulations. *IEEE T Bio-Med Eng*, 52(12):1994–2005, 2005.
- [79] E. Torlone, C. Fanelli, AM Rambotti, G. Cassi, F. Modarelli, A. Vincenzo, L. Epifano, M. Ciofetta, S. Pampanelli, P. Brunetti, et al. Pharmacokinetics, pharmacodynamics and glucose counterregulation following subcutaneous injection of the monomeric insulin analogue [lys (b28), pro (b29)] in t1dm. *Diabetologia*, 37(7):713–720, 1994.
- [80] KJ Versyck, JE Claes, and JF Van Impe. Optimal experimental design for practical identification of unstructured growth models. *Mathematics and computers in simulation*, 46(5-6):621–629, 1998.

- [81] E. Walter and L. Pronzato. *Identification of parametric models*. Springer Heidelberg, 1997.
- [82] ME Wilinska, LJ Chassin, CL Acerini, JM Allen, DB Dunger, and R Hovorka. Simulation environment to evaluate closed-loop insulin delivery systems in type 1 diabetes. *J Diab Sci Technol*, 4(1):132–144, 2010.
- [83] ME Wilinska, LJ Chassin, HC Schaller, L. Schaupp, TR Pieber, and R. Hovorka. Insulin kinetics in type-1 diabetes: continuous and bolus delivery of rapid acting insulin. *IEEE T Bio-Med Eng*, 52(1):3–12, 2005.
- [84] M.E. Wilinska and R. Hovorka. Simulation models for< i> in silico</i> testing of closed-loop glucose controllers in type 1 diabetes. *Drug Discovery Today: Disease Models*, 5(4):289–298, 2009.
- [85] H.A. Wolpert. The nuts and bolts of achieving end points with real-time continuous glucose monitoring. *Diabet Care*, 31:146–149, 2008.
- [86] H.C. Zisser, T.S. Bailey, S. Schwartz, R.E. Ratner, and J. Wise. Accuracy of the seven® continuous glucose monitoring system: Comparison with frequently sampled venous glucose measurements. *J Diab Sci Technol*, 3(5):1146–1154, 2009.