Lecture Notes in Bioengineering

Harald Kirchsteiger
John Bagterp Jørgensen
Eric Renard
Luigi del Re *Editors*

Prediction Methods for Blood Glucose Concentration

Design, Use and Evaluation

Springer

Lecture Notes in Bioengineering

More information about this series at http://www.springer.com/series/11564

Harald Kirchsteiger · John Bagterp Jørgensen
Eric Renard · Luigi del Re
Editors

Prediction Methods for Blood Glucose Concentration

Design, Use and Evaluation

Springer

Editors
Harald Kirchsteiger
Institute for Design and Control of Mechatronical Systems
Johannes Kepler University Linz
Linz
Austria

John Bagterp Jørgensen
Department of Applied Mathematics
Technical University of Denmark
Kongens Lyngby
Denmark

Eric Renard
Institut de Génomique Fonctionnelle de Montpellier
Montpellier
France

Luigi del Re
Institute for Design and Control of Mechatronical Systems
Johannes Kepler University Linz
Linz
Austria

ISSN 2195-271X ISSN 2195-2728 (electronic)
Lecture Notes in Bioengineering
ISBN 978-3-319-25911-6 ISBN 978-3-319-25913-0 (eBook)
DOI 10.1007/978-3-319-25913-0

Library of Congress Control Number: 2015953261

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

Standard diabetes insulin therapy for type 1 diabetes and late stages of type 2 is based on the expected development of blood glucose (BG) both as a consequence of the metabolic glucose consumption as well as of meals and exogenous insulin intake. Traditionally, this is not done explicitly, but the insulin amount is chosen using factors that account for this expectation.

The increasing availability of more accurate continuous blood glucose measurement (CGM) systems is attracting much interest to the possibilities of explicit prediction of future BG values. Against this background, in 2014 a two-day workshop on the design, use and evaluation of prediction methods for blood glucose concentration was held at the Johannes Kepler University Linz, Austria. One intention of the workshop was to bring together experts working in various fields on the same topic, in order to shed light from different angles on the underlying problem of modeling the glucose insulin dynamics of type 1 diabetes patients. Among the international participants were continuous glucose monitoring developers, diabetologists, mathematicians and control engineers, both, from academia and industry. In total 18 talks were given followed by panel discussions which allowed to receive direct feedback from the point of view of different disciplines.

This book is based on the contributions of that workshop and is intended to convey an overview of the different aspects involved in the prediction. The individual chapters are based on the presentations given by the authors at the workshop but were written afterward which allowed to include the findings and conclusions of the various discussions and of course updates.

The chapter "Alternative Frameworks for Personalized Insulin–Glucose Models" by Harald Kirchsteiger et al. asks the question whether more and more detailed physiological descriptions of the glucose metabolism with an ever-increasing degree of sophistication and number of modeled phenomena are really what is needed for pushing the boundaries in glucose prediction for control. As an alternative, the chapter introduces two data-based approaches that focus not on the prediction of exact future blood glucose values, but rather on the prediction of changes in the patients' blood glucose range.

The chapter "Accuracy of BG Meters and CGM Systems: Possible Influence Factors for the Glucose Prediction Based on Tissue Glucose Concentrations" by Guido Freckmann et al. discusses performance metrics used to characterize the accuracy of continuous glucose measurement devices. This topic is highly relevant for prediction models since many of them rely on the data given by the continuous sensors which are previously calibrated with blood glucose meter measurements which are also subject to measurement errors. Inaccurate measurements will directly affect the performance of the corresponding predictions.

The chapter "CGM—How Good Is Good Enough?" by Michael Schoemaker and Christopher G. Parkin also tackles the problem of continuous glucose monitor performance evaluation. Several performance metrics used in different published studies are compared and their individual characteristics analyzed. The chapter reveals why the comparison of a sensor evaluated in two different clinical studies is not always straightforward.

The chapter "Can We Use Measurements to Classify Patients Suffering from Type 1 Diabetes into Subcategories and Does It Make Sense?" by Florian Reiterer et al. makes use of continuous time prediction models to describe the interaction between ingested carbohydrates, subcutaneously injected insulin, and continuously measured glucose concentration. The identified model parameters of 12 subjects were analyzed and statistically significant correlations between the parameters and patient characteristics such as weight and age could be found.

The chapter "Prevention of Severe Hypoglycemia by Continuous EEG Monitoring" by Claus Borg Juhl et al. shows how to use EEG signals to predict upcoming hypoglycemic situations in real-time by employing artificial neural networks. The results of a 30-day long clinical study with the implanted device and the developed algorithm are presented.

The chapter "Meta-Learning Based Blood Glucose Predictor for Diabetic Smartphone App" by Valeriya Naumova et al. demonstrates how a highly sophisticated glucose prediction model can be ported from a development language running on a PC to a format such that it can be used conveniently by the patients. A unique feature of the algorithm is its independence of any user input other than historic CGM data which is automatically transmitted from a CGM device. No parameter estimation nor prediction model individualization is required.

The chapter "Predicting Glycemia in Type 1 Diabetes Mellitus with Subspace-Based Linear Multistep Predictors" by Marzia Cescon et al. uses data-based methods to develop individualized prediction models. The model can be considered as a combination of physiological models to precompute the rate of appearance of injected insulin and ingested carbohydrates in the bloodstream and of data-based models to combine this information and compute predictions up to 120 min in the future. The results show the performance on data from 14 type 1 diabetes patients in a clinical trial.

The chapter "Empirical Representation of Blood Glucose Variability in a Compartmental Model" by Stephen D. Patek et al. shows a modeling technique designed to extract the information on the net effect of meals on the blood glucose concentration. By assuming that all major unexplained glycemic excursions can be

attributed to oral glucose ingestion, a meal vector is estimated which significantly improves the mathematical model. Results are shown on three patients during a clinical trial and on virtual patients where it is shown how the method can be used for adjustments of the basal insulin rate.

The chapter "Physiology-Based Interval Models: A Framework for Glucose Prediction Under Intra-patient Variability" by Jorge Bondia and Josep Vehi tries to cope with the large intrasubject variability by using the concept of interval predictions. Instead of predicting a single blood glucose value in the future, a whole solution envelope is determined. With the presented theory it can be guaranteed that the real value is always inside of the envelope and moreover the envelope is not conservative. The method is evaluated on a physiological diabetes model.

The chapter "Modeling and Prediction Using Stochastic Differential Equations" by Rune Juhl et al. considers uncertainty in the dynamics between different patients as well as within a patient by making use of stochastic differential equations. It is shown how the mixed effects modeling methodology can be applied such that the underlying information of several datasets from different patients is extracted to form the model.

The chapter "Uncertainties and Modeling Errors of Type 1 Diabetes Models" by Levente Kovács and Péter Szalay analyzes the effect of prediction model uncertainties on the control system during a design procedure involving the steps model reduction by elimination of state variables, state estimation using extended Kalman Filters and Sigma Point filters and linear parameter-varying control synthesis.

The chapter "Recent Results on Glucose–Insulin Predictions by Means of a State Observer for Time-Delay Systems" by Pasquale Palumbo et al. introduces a prediction model which in real time predicts the insulin concentration in blood which in turn is used in a control system. The method is tested in simulation on a time-delay system representing the glucose–insulin system.

The chapter "Performance Assessment of Model-Based Artificial Pancreas Control Systems" by Jianyuan Feng et al. makes use of prediction models to compute treatment advices. The novelty of the proposed algorithm consists in explicitly considering (among others) the model prediction error and model error elimination speed. A retuning of the advisory system is done in case the prediction model does not perform well. Results on 30 virtual patients show the performance of the control system.

We would like to thank all people involved in the process of writing this book: All authors for their individual contributions, all reviewers of the book chapters, Daniela Hummer for the entire organization of the workshop, Boris Tasevski for helping with the typesetting, Florian Reiterer for his help editing the book, as well as Oliver Jackson and Karin de Bie for the good cooperation with Springer.

Linz
August 2015

Harald Kirchsteiger
John Bagterp Jørgensen
Eric Renard
Luigi del Re

Contents

Alternative Frameworks for Personalized Insulin–Glucose Models

Harald Kirchsteiger, Hajrudin Efendic, Florian Reiterer and Luigi del Re

Abstract The description of the insulin–glucose metabolism has attracted much attention in the past decades, and several models based on physiology have been proposed. While these models provide a precious insight in the involved processes, they are seldom able to replicate and much less to predict the blood glucose (BG) value arising as a reaction of the metabolism of a specific patient to a given amount of insulin or food at a given time. Data-based models have proven to work better for prediction, but predicted and measured values tend to diverge strongly with increasing prediction horizon. Different approaches, for instance the use of vital signs, have been proposed to reduce the uncertainty, albeit with limited success. The key assumption hidden behind these methods is the existence of a single "correct" model disturbed by some stochastic phenomena. In this chapter, instead, we suggest using a different paradigm and to interpret uncertainty as an unknown part of the process. As a consequence, we are interested in models which yield a similar prediction performance for all measured data of a single patient, even if they do not yield a precise representation of any of them. This chapter summarizes two possible approaches to this end: interval models, which provide a suitable range; and probabilistic models, which provide the probability that the BG lies in predetermined ranges. Both approaches can be used in the framework of automated personalized insulin delivery, e.g., artificial pancreas or adaptive bolus calculators.

1 Introduction

Glucose is the main energy source for the human body, but in excessive amounts it can be detrimental as well. In healthy persons, the correct glucose level is regulated by the insulin–glucose metabolism. In diabetes patients, this control system does not work properly any more, due, roughly speaking, either to lack of endogenous insulin

H. Kirchsteiger · H. Efendic · F. Reiterer (✉) · L. del Re
Institute for Design and Control of Mechatronical Systems,
Johannes Kepler University Linz, 4040 Linz, Austria
e-mail: florian.reiterer@jku.at
URL: http://desreg.jku.at/

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*,
Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_1

production or to a resistance to insulin action [23]. In such cases, the therapy of choice consists in providing the necessary insulin externally. However, the delay times associated with this procedure and the danger of removing too much glucose from the blood circulation—with possibly lethal consequences—make the determination of the correct insulin quantity quite difficult and delicate.

Against this background, there has been a sustained interest in describing the glucose–insulin metabolism by physiological models of increasing complexity [2, 3, 25, 44] . A model resulting from the cooperation between the Universities of Padova and Virginia [12] has been accepted in 2008 by the Food and Drug Administration (FDA) as a substitute for preclinical trials for certain insulin treatments. The model was recently extended with new features [11].

And, of course, there has been a large number of attempts to use the model information to determine the right quantities of external insulin needed, so to say to replace the missing control action, to "close the loop." Positive results have been reported for overnight operation [26], but also a reduction of variability for the daily use have been reported [33]. New and improved sensors have been introduced, faster insulin is entering the market, and all this nurtures the hope that more progress will come. Unfortunately, in spite of over 40 years of research [1], the results of this "artificial" or "virtual pancreas" are still not there where they should be and simple safety rules—e.g., avoiding insulin infusion during or near to hypoglycemia—seem to be able to offer the largest part of the benefits of closed-loop control in a much simpler way as well.

Against this background, it is natural to lean back for a moment and wonder whether we are asking the right questions. In this chapter, we are addressing this very issue by presenting two promising alternatives to modeling, namely interval models and probabilistic modeling techniques.

2 Alternatives for Modeling

The key assumption for modeling is the reproducibility of results. To some extent, of course, this is always true, for instance insulin does reduce the blood glucose (BG) concentration and carbohydrates (Carb) intake increases it. Many other effects are known but hard to quantify—e.g., some hormones increase it as well, muscular work reduces it—but there are many control loops in the human body which can lead to a complex response, e.g., in the case of prolonged physical work [21]. Other effects, like the circadian variation of sensitivities, are known as well [46].

Physiological models describe all these phenomena relying on deep understanding of physiology and frequently on very specific measurements, e.g., tracers [24, 42]. It is usually impossible to determine all parameters of such models with simple "external" measurements, e.g., taking only Carb and insulin administration and BG values. Indeed, it has been shown that the "external" behavior—the relation between insulin and Carb intake and BG—of these complex models can be approximated very well with very simple ones [17, 45] (see Fig. 1). This is the one

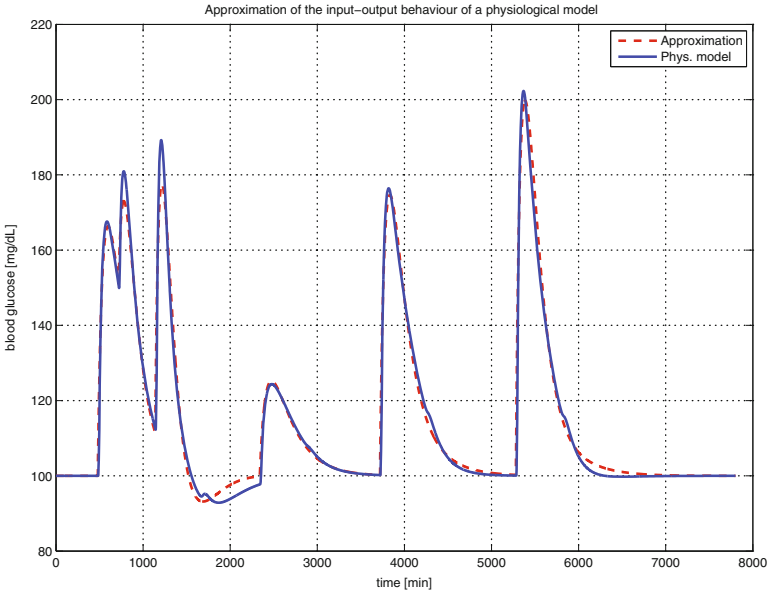

Fig. 1 Approximation of the input–output relationship of model [12] as shown in [45]

reason why comparisons between the values predicted by physiological models and measurements are very seldom, exceptions being for instance [8, 47].

Indeed, in general, physiological models are not able to provide a personalized description.

As the comparison makes clear, the measurements have a high degree of complexity not reflected in the physiological model, the main reason being the many unmodeled effects, e.g., related to the emotional state, which may affect very strongly the BG values, and which cannot be captured by the model because the critical quantities, in this case the concentration of some hormones, are not known.

There are several ways to cope with this problem. On one side, the attempt can be made to find additional measurements to extend the model, e.g., vital signs—acceleration, heart frequency, body temperature, and so on. Similar techniques have proved very useful in the industrial framework, e.g., to detect changes in machines [13], but have never really succeeded in the case of diabetes treatment.

Another approach, related to another chapter of this book (see "Empirical Representation of Blood Glucose Variability in a Compartmental Model" by S. Patek et al.) consists essentially in estimating a "corrective" Carb input to explain the difference between measured and computed values. While this method cannot be used in real time, it allows to study the effect of some changes in therapy, e.g., different amounts of insulin.

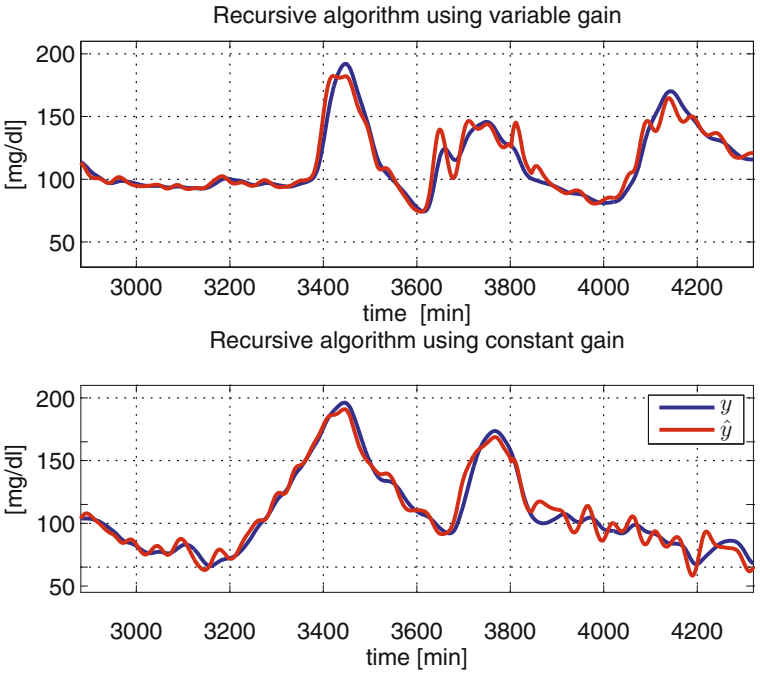

Fig. 2 45 min ahead prediction of glucose for two different patients from [6]

If we are interested in obtaining models which are sufficiently simple to allow their use and parameter estimation in real time, it might be better to look for other approaches. A very efficient adaptive model was developed by [5] which relies on a simple hypothesis, a so-called ARX model, and determines the parameters continuously, concentrating on the most critical BG ranges. Figure 2 shows the performance of such a model as predictor.

In this chapter, however, we suggest two different approaches. The key idea is not to get rid of uncertainty, assuming one particular value to be true, but to design models valid for the whole region, implicitly assuming that a full range of values are possible and in some sense true. One possible approach to this end are interval models, i.e., models which compute an output range and not a single value. The other alternative is using a probabilistic approach. Indeed, the exact BG value is not really important in itself, the clinician is more interested in keeping it inside the usual ("euglycemic") range and preventing to reach a dangerous one, e.g., hypoglycemia. Markov jump model can help in describing the physiology the way it really is—i.e., to some extent random.

Both models can be used for automated insulin delivery as well. In the case of interval models, the problem can be stated in terms of a min/max problem such that the carbohydrate amount optimizes the cost functions for all cases of uncertainty. In

the probabilistic case, it becomes the minimization of the probability that, under the action of insulin and meals, the BG leaves the good range to reach a dangerous one.

The further sections of this chapter are organized as follows: Before coming to the topic of interval models itself, a review about empirical continuous time transfer function models is given in Sect. 3 and the differences between possible model structures are discussed. The models presented in that section are a special form of control-oriented data-based models for describing the input–output relationship of the glucose metabolism that have proven quite powerful in the recent past (see e.g., [37, 38]). One of those presented model structures is then further used for the interval modeling introduced in Sect. 4. After a quick overview about the topic of interval modeling in general, some details about the methods used here for deriving interval models from data are given in Sect. 4.1. In following Sect. 4.1, results for the interval modeling are shown, both for simulated and for real patient data. The subsequent Sect. 5 then describes a probabilistic framework that can be used for predicting changes from one BG range to another. It starts in Sect. 5.1 with an overview about Gaussian mixture models that have been used in this context. Section 5.2 gives some details about the used model structure and the methodology of predicting transitions in the BG range, whereas actual prediction results for real patient data are presented in the following Sect. 5.3. The chapter finished with some final conclusions and discussion given in Sect. 6.

3 Model Structures

A model consists of a mathematical structure and of parameters. Thus the first step in modeling consists of fixing the model structure, and thereafter the parameters have to be tuned to get the best correspondence between measured and computed values. Of course, every model represents a simplification of the real system, and not every model structure is able to capture the behavior of the system under observation in a sufficiently general way. This is especially true in the case of a simplified model we are interested in.

Table 1 summarizes previously proposed model structures to describe the blood glucose dynamics where $BG(s)$, $Carb(s)$, and $I(s)$ correspond to the blood glucose concentration, ingested meal carbohydrates, and subcutaneously injected insulin bolus, respectively, all transformed into the Laplace domain. The table also lists the number of parameters which need to be estimated from data. All models thus use the same amount of information. A typical dataset which could be used for parameter estimation is shown in Fig. 3 (data from the DIAdvisor project [14]).

It is immediately visible that the model inputs are impulse-shaped quantities which are zero most of the time. That is because subcutaneous insulin injections are discrete events and meal ingestions, regardless of the quantity and time it takes to actually consume the food, are commonly treated as discrete events also. Therefore, for model analysis, the *impulse* responses are of great interest whereas *step* responses, commonly used in control engineering, do not provide sufficient information. A step

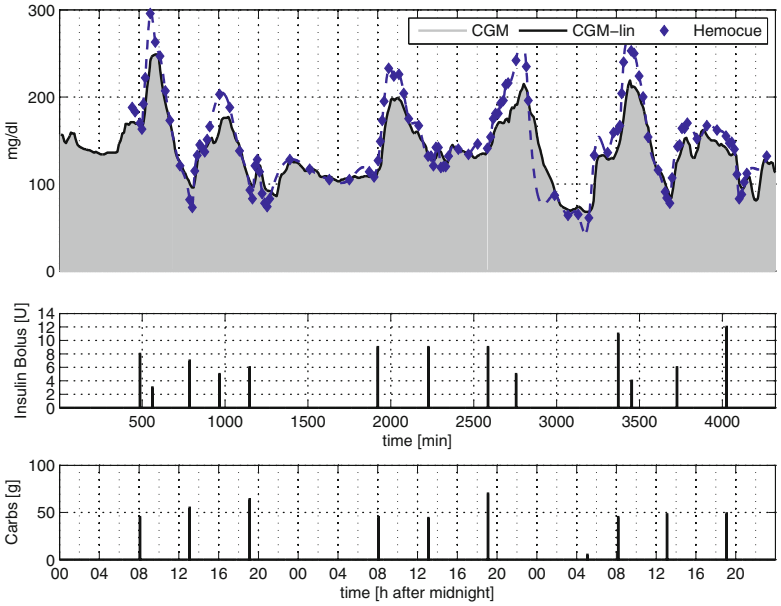

Fig. 3 Illustration of the relevant measurement data for a specific patient, CHU0102 (data from the DIAdvisor project [14])

Table 1 Selected model structures previously published

No.	Model structure	Parameters	Reference
1	$BG(s) = \frac{K_1}{(1+sT_1)^2 s} Carb(s) + \frac{K_2}{(1+sT_2)^2 s} I(s)$	4	[29]
2	$BG(s) = \frac{K_1 \exp(-\tau_1 s)}{(1+sT_1)s} Carb(s) + \frac{K_2 \exp(-\tau_2 s)}{(1+sT_2)s} I(s)$	6	[36]
3	$BG(s) = \frac{K_1}{(1+sT_1)s} Carb(s) + \frac{K_2}{(1+sT_2)s} I(s)$	4	[9]
4	$BG(s) = \frac{K_1}{(1+sT_1)^2} Carb(s) + \frac{K_2}{(1+sT_2)^2} I(s)$	4	[4][a]
5	$BG(s) = \frac{K_1 \exp(-\tau_1 s)}{(1+sT_1)(1+sT_2)} Carb(s) + \frac{K_2 \exp(-\tau_2 s)}{(1+sT_3)s} I(s)$	7	[10]

[a]Extended with a dynamic model for carbohydrates

response would actually mean that food or insulin is added to the metabolism in a continuous way over an extended period of time. While this might be true for the case of insulin, it is definitely an unrealistic case for food. In all given references, continuous insulin delivery is not considered for modeling.

From a physiological point of view, the impulse response gives precise information on the effect of one gram of carbohydrate and one unit of insulin, respectively, on the blood glucose concentration.

The impulse responses for both inputs of model structure 1 in Table 1 are shown in Fig. 4 for various selections of the parameters T_1 and T_2. It is immediately clear that the parameters K_1 and K_2 correspond to the steady-state change in BG. Furthermore, the parameters T_1 and T_2 are time constants which determine the time it takes until

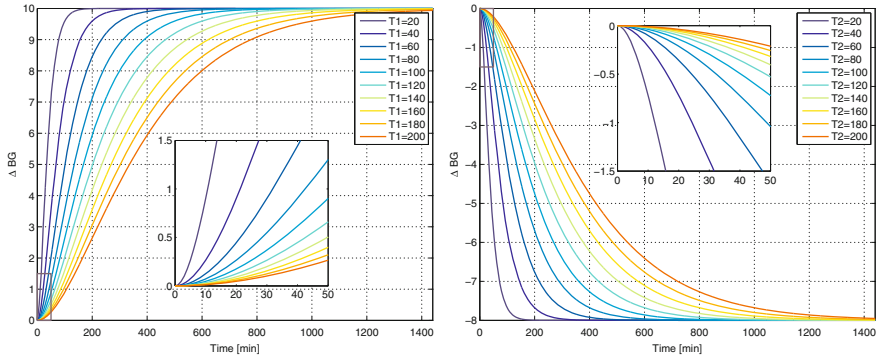

Fig. 4 Impulse response of model structure 1 from Table 1 using $K_1 = 10$, $K_2 = -8$

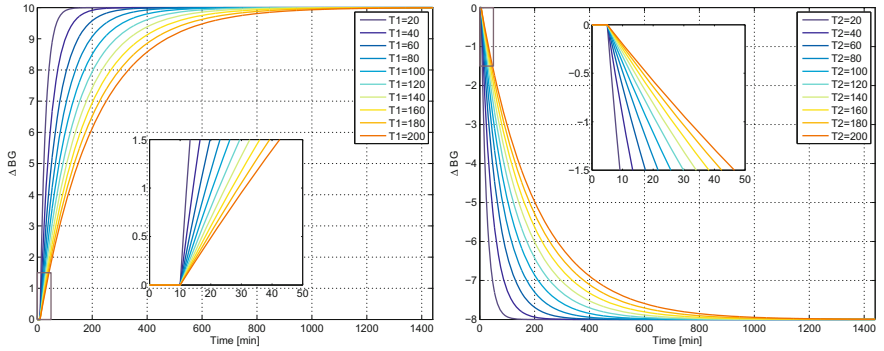

Fig. 5 Impulse response of model structure 2 from Table 1 using $K_1 = 10$, $\tau_1 = 10$, $K_2 = -8$, $\tau_2 = 5$

this steady state is reached. From this perspective, the physiological interpretation of the model parameters is straightforward.

The impulse responses for both inputs of model structure 2 in Table 1 are shown in Fig. 5 for various selections of the parameters T_1 and T_2. Compared to the impulse responses in Fig. 4, model structure 2 involves a time delay determined by the parameters τ_1 and τ_2. Furthermore, after this delay, the impulse response shows a discontinuity (see the magnified plots in Fig. 5) which hardly appears in real subjects. However, the overall responses of model structures 1 and 2 are similar.

The impulse responses for model structure 3 in Table 1 are not shown explicitly since they are very similar to those of model structure 2 shown in Fig. 5, except for the time delay which is zero in this case ($\tau_1 = 0$, $\tau_2 = 0$).

Model structure 4 of Table 1 gives a significantly different impulse response than the models discussed above, see Fig. 6. Since there is no integrating behavior, the impulse responses return to the steady state of zero. Physiologically, the assumption is that even in the absence of insulin, glucose will be removed from the circulation after meal ingestion. This is fundamentally different than the assumptions in the

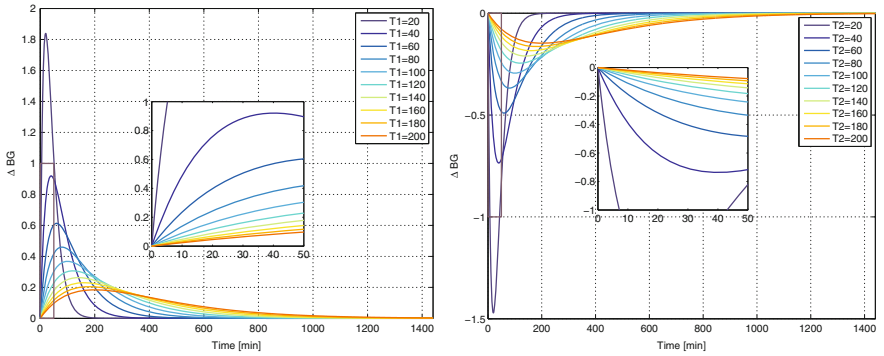

Fig. 6 Impulse response of model structure 4 from Table 1 using $K_1 = 100$, $K_2 = -80$

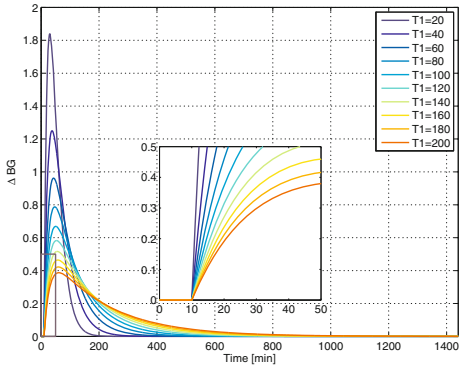

Fig. 7 Impulse response of model structure 5 from Table 1 using $K_1 = 100$, $T_2 = 20$, $\tau_1 = 10$

previous structures, where glucose will not be removed until insulin is supplied. As it can be seen in the experimental results in the corresponding papers, both model assumptions can lead to a good approximation of real data. Note also that model parameters are not directly visible from the impulse response.

Finally, the impulse response related to the carbohydrate input from model structure 5 in Table 1 is shown in Fig. 7. The response related to the insulin input is the same as for model structure 2 and can be seen in Fig. 5. The main difference in this structure is the assumption of two different time constants T_1, T_2, both for the carbohydrate dynamics. Therefore, there is more flexibility in the response compared to the case of using only one value for both constants ($T_1 = T_2$). However, it is questionable whether distinctive values for T_1 and T_2 can really be identified from clinical data [29].

4 Interval Models

There are essentially two ways to characterize interval models: one viewpoint is to design the model such that all possible realizations of the uncertainties will be considered and the true system output will be within the computed model output all the time, minimizing at the same time the conservativeness. The other viewpoint is to allow a certain amount or points to lie outside of the computed interval and thereby obtaining a much smaller interval [35]. The first approach is treated in another chapter of this book (see "Physiology-Based Interval Models: A Framework for Glucose Prediction Under Intra-Patient Variability" by J. Bondia and J. Vehi) while we focus on the latter here.

We will choose the model structure 1 from Table 1

$$BG(s) = \frac{K_1}{(1+sT_1)^2 s} Carb(s) + \frac{K_2}{(1+sT_2)^2 s} I(s) \tag{1}$$

because it contains only four parameters to be estimated and does not lead to a discontinuity in the impulse response. The generalized description of this model (assuming only one input for simplicity of notation) is a continuous time process model of the form

$$G(s) = \frac{B(s)}{A(s)} = \frac{b_0 + b_1 s + \cdots + b_m s^m}{a_0 + a_1 s + \cdots + s^n}. \tag{2}$$

Modeling a specific patient means to assign values to the variables a_i, b_i in (2). This is an identification problem [34, 43] and can be tackled in two ways. By transforming (2) into an equivalent discrete-time formulation, all the available tools of discrete-time system identification can be applied. However, such a transformation might introduce additional parameters and the initial (physiological) interpretation of the continuous time parameters gets lost [20]. On the opposite, there are continuous time identification methods which directly estimate the parameters in (2) without any need for transformation [19]. The benefits of such a direct estimation in the context of the human glucose insulin system have been treated in [10, 30]. In this contribution, we will thus also focus on a direct continuous time estimation method.

4.1 Continuous Time System Identification

Beginning first without taking into account the uncertainty, estimates of the model parameters can be found by minimizing a quadratic criterion of the form

$$J_1(\theta) = \frac{1}{N} \sum_{k=1}^{N} \varepsilon^2(k, \theta) \tag{3}$$

where ε denotes the difference between measurement and model output,

$$\varepsilon(t, \theta) = y(t) - \hat{y}(t, \theta) \tag{4}$$

N is the total number of available measurements and the vector θ contains all parameters. Various optimization techniques exist for actually minimizing the criterion (3) (see e.g., [43]).

Now, considering uncertainty in the system to be modeled we assume two separate datasets of the same system with different underlying dynamics. The cost function which is minimized now is the sum of two similar terms as above, again taking into account the deviation between model output and measurement. Additionally, we will introduce a third term which penalize the standard deviation of the estimated model parameters for the first dataset compared to the second dataset (5). For example, if this third term is not present, the model parameters for both datasets will be estimated independently and might deviate to a great extent from each other. By introducing and weighting (using the weighting matrix Π which has the tuning parameters in the main diagonal) the third term, a compromise between a good model fit of the individual datasets and compact parameter sets can be obtained.

$$J_2(\theta_1, \theta_2) = \frac{1}{N}\sum_{k=1}^{N} \varepsilon_1^2(k, \theta_1) + \frac{1}{N}\sum_{k=1}^{N} \varepsilon_2^2(k, \theta_2) + \frac{1}{N} \left\| \bar{\sigma}(\theta_1,\ \theta_2) \right\|_{\Pi}^2 \tag{5}$$

The function $\bar{\sigma}$ denotes a vector standard deviation operator which determines component-wise the standard deviation of the model parameters of the parameter vector θ.

Minimization can be done with a Gauss–Newton algorithm [43]. Thanks to the definition of the cost and the model structure, gradient vector and Hessian matrix can be determined analytically [30]. Nevertheless, the optimization is nonlinear and iterative and thus care has to be taken to avoid improper starting values for the parameters. A generalization of the cost function to N_{exp} datasets and the corresponding gradient vectors and the Hessian was derived in [30].

As a result of the continuous time system identification, there is one parameter vector per experiment. The next step is to extract parameter intervals which then define the model output interval. Considering again the model structure 1 from Table 1, there are four model parameters which means the parameter vector has the form

$$\theta = [K_1,\ K_2,\ T_1,\ T_2]^{\top} \tag{6}$$

Denoting with superscripts the corresponding dataset number, at the end of the identification N_{exp} estimates $K_1^1, K_1^2, \ldots K_1^{N_{exp}}$ are available and their maximum value is K_1^{max} and minimum value K_1^{min}. The computation of the interval model output is then

$$\hat{Y}^{\max}(s) = \frac{K_1^{\max}}{(1+T_1^{\min}s)^2 s}U^{(1)}(s) + \frac{K_2^{\max}}{(1+T_2^{\max}s)^2 s}U^{(2)}(s) \tag{7a}$$

$$\hat{Y}^{\min}(s) = \frac{K_1^{\min}}{(1+T_1^{\max}s)^2 s}U^{(1)}(s) + \frac{K_2^{\min}}{(1+T_2^{\min}s)^2 s}U^{(2)}(s) \tag{7b}$$

Note that K_2 describes the effect of insulin and is thus negative. In the following subsection, we will present results when applying the interval model estimation to short data segments representing the breakfast period. The starting point of the data is breakfast time (around 8:00 in the morning) and data end points just before lunch were taken (around 12:30).

4.2 Interval Model Results

Figure 8 shows a typical result obtained with the methodology described above for three independent measurement sets of a single patient. The black stars are the measured glucose levels using a Yellow Spring Instrument *YSI 2300 STAT Plus™ Glucose Analyzer* (YSI) device. When modeling each dataset independently (individual models), the green curves apply which show of course the best performance. Depending on the choice of the tuning matrix Π, the identified parameters then depend stronger or less strong on each other and an output interval according to (7) can be computed, see the red dashed lines in Fig. 8. Finally, also the mean interval model response is shown, which is computed when using as parameter the average values of all the estimated parameters of the interval model.

Remark 1 Again note the dependency of the results on the tuning matrix Π. When choosing $\Pi = 0$ a zero matrix, the corresponding term in the cost function (5) will disappear and the result is an independent estimation of the parameters for each experiment. This also results in the largest possible output interval. In the other extreme, choosing very large values in Π results in very small standard deviations of the model parameters, making them equivalent. This also results in an output interval which degenerates to a single line and is the same result as for a standard multiexperiment identification setup [34].

For an in-depth analysis of the method, it was applied to 10 simulated and 10 clinical datasets. The simulated data was obtained from a time-varying metabolic simulation model [28]. In the following, we will use the term individual model for a model estimated on a single dataset (experiment), the term interval model for a model according to (7), and the term mean interval model when a simulation with the mean value of the interval model parameters is done.

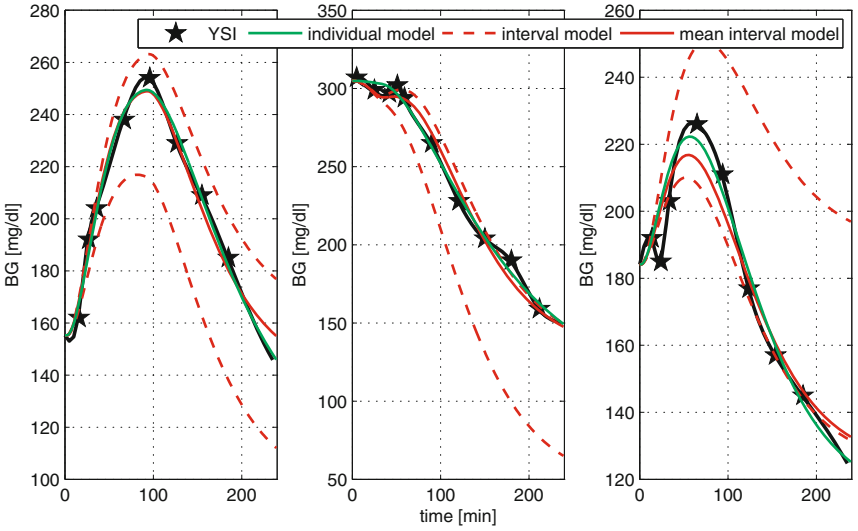

Fig. 8 Output response of the interval model compared to measurements for three datasets

4.2.1 Performance Metrics

Results will be evaluated based on a fit value (8) and on the error in detecting the peak BG. This is visualized in Fig. 9: ΔBG is the difference between the measured and predicted peak BG and ΔT is the difference in time. Both quantities can be positive or negative, where positive means the measured BG peak is higher and later in time than the simulated. The motivation of using those metrics is their importance in diabetes treatment.

$$\text{fit} = 100\left(1 - \frac{\|\hat{y} - y\|_2^2}{\|y - \bar{y}\|_2^2}\right) \tag{8}$$

4.2.2 Results Using the Simulated Data

First, we will present results from a single simulated patient before presenting statistics of all 10 considered patients. The parameter values and performance metrics reported in Table 2 are from simulated patient number 2. For the interval model, Π (5) was tuned in such a way, that the fit value of the interval model is not significantly less than 90 % of the fit of the individual model, and that the ratio between mean value and standard deviation $(1/CV)$ of each of the four parameters is at least five. In this way, it is ensured that the interval predictions are reasonably tight.

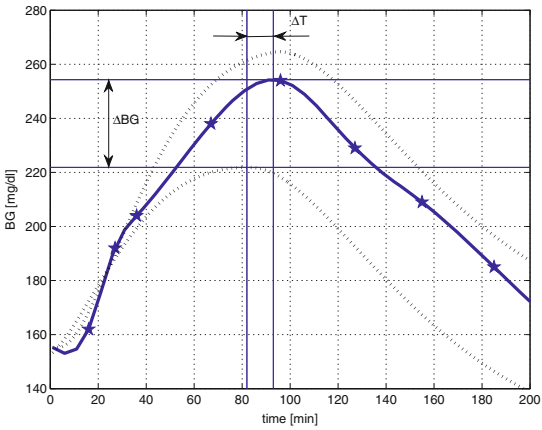

Fig. 9 Performance metrics $(\Delta BG, \Delta T)$ for evaluation of the results. *Blue stars* indicate BG measurements, interpolated with second-order polynomial. *Black dashed* are the min/max responses of the interval model (color figure online)

For the presented results in Table 2 and Fig. 10, the elements in the main diagonal of Π are [15000, 125, 20, 0.3] result in $1/CV = [5.02, -5.90, 5.94, 6.62]$. The parameter estimation was done on the first 3 days (up to $t = 72$ h). For the individual models, the validation is based on mean parameters of the first 3 days. For the interval models, the performance metrics were calculated based on the exact estimations for days 1–3 and validation was done with the mean interval model.

From Table 2, we see that there is a rather small ratio between mean value and standard deviation of the estimated parameters (especially for K_2)—which would mean that the insulin effect is very different—when the experiments are identified independently of each other. This ratio is greatly increased when applying the proposed interval model estimation, without decreasing the performance (on the three training days). Considering validation, the interval model shows better performance compared to the individual model. Note also that in this particular case, the individually estimated model for day 3 is not useful, because $K_2 < 0$, i.e., an insulin injection would result in a glucose increase.

The results for all 10 patients are summarized in Table 3 where the mean values of the performance metrics over all 5 days are shown. The interval model shows the highest fit values in every case, and the average error (for all 10 patients) of correct estimation of the BG peak is only 3.79 mg/dl compared to 10.80 mg/dl of the individual models. Note that this rather high error is to a large extent caused by the specific glucose responses to meals by the simulator for some patients, where the response does not have a single maximal value. In those cases there are two almost equivalent high peaks, but separated in time. Such a shape cannot be reproduced with the chosen model structure (1), and the approximation typically lies in the middle of the two peaks. It is of interest to note that from the 30 individual models, there were 8 that showed to be incorrect from a physiological point of view ($K_1 < 0$ and/or $K_2 > 0$). Considering the interval model, only for patient #6 no physiologically correct models could be found, as this patient's glucose dynamics have a high time

Table 2 Estimated parameters and performance metrics for simulated patient # 2

Individual model							
	Meal #1	Meal #2	Meal #3	Meal #4	Meal #5	Mean	Mean/Std.
Fit	66.54	66.85	87.66	12.73	42.02	55.16	1.92
ΔBG	2.96	0.56	1.35	−10.72	−11.06	5.33	1.03
ΔT	−7.00	−133.00	99.00	−129.00	−132.00	100.00	1.86
K_1	1.28(±0.08)	0.99(±0.05)	0.77(±0.04)	1.01	1.01	1.01	3.97
K_2	−13.83(±0.59)	−5.56(±0.41)	7.37(±0.48)	−4.01	−4.01	−4.01	−0.38
T_1	24.99(±1.33)	20.28(±0.91)	14.69(±0.60)	19.98	19.98	19.98	3.88
T_2	99.77(±12.02)	92.55(±16.48)	61.43(±4.25)	84.58	84.58	84.58	4.15
Interval model							
Fit	72.02	69.68	65.40	33.76	51.93	58.56	3.68
ΔBG	−0.13	1.51	−2.23	−7.20	−6.14	3.44	1.12
ΔT	−9.00	−121.00	−99.00	−114.00	−123.00	93.20	1.94
K_1	1.00	0.96	1.36	1.10	1.10	1.10	5.02
K_2	−16.19	−13.63	−11.53	−13.78	−13.78	−13.78	−5.90
T_1	19.00	21.16	26.26	22.14	22.14	22.14	5.94
T_2	211.86	253.75	287.56	251.05	251.05	251.05	6.62

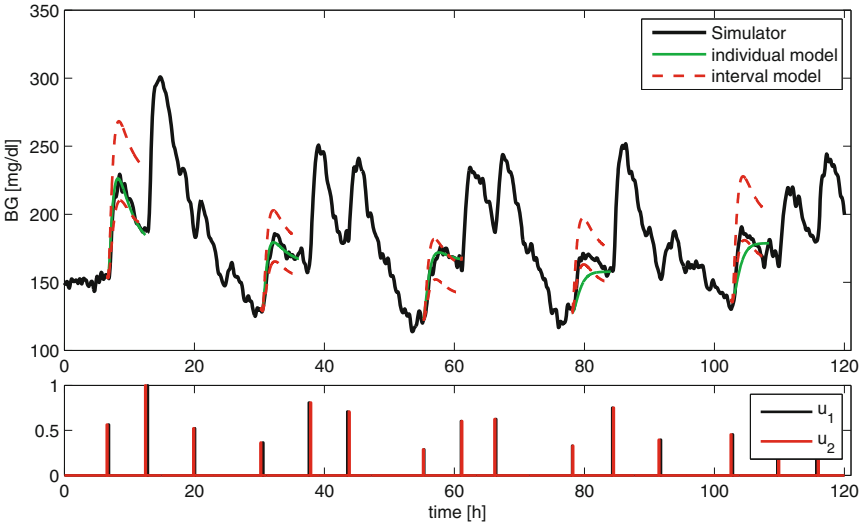

Fig. 10 Simulated patient #2: simulated glucose profile (*black solid*) individual model outputs for the breakfasts (*green solid*) and worst-case interval model outputs for breakfasts (*red dashed*). The first 3 days (up to $t = 72$ h) were used for estimation of the models, the remaining 2 days show validation results. *Bottom panel* shows the scaled impulse carbohydrate (*black solid*) and insulin (*red solid*) inputs

Table 3 Statistics for all 10 virtual patients, mean values of the performance metrics

	Individual model			Interval model		
Pat	Fit	ΔBG	ΔT	Fit	ΔBG	ΔT
P1	46.08	5.08	108.00	61.82	5.68	23.40
P2	55.16	5.33	100.00	58.56	3.44	93.20
P3	86.68	2.72	2.40	87.05	2.73	2.20
P4	32.15	16.36	76.00	74.01	4.64	49.60
P5	−8.26	33.73	18.80	89.69	2.67	18.20
P6	63.35	6.63	71.40	67.32	6.74	78.80
P7	65.20	6.76	8.00	71.41	3.68	6.00
P8	65.32	3.83	69.00	67.57	3.22	57.80
P9	70.78	3.08	6.40	71.57	2.48	5.40
P10	15.04	24.48	35.40	82.74	2.62	4.20
Avg.	49.15	10.80	49.54	73.18	3.79	33.88

constant, and BG is only rising in the observational period. To correct this, a longer simulation time would be necessary.

4.2.3 Results Using the Clinical Data

The data used here was collected in [14]. Similar to the previous section, let us first analyze the estimated parameters for a specific patient in Table 4. The ratio between mean value and standard deviation was again greatly increased while the mean performance decrease is only moderate.

A graphical representation of the results obtained for this patient (#0107) is given in Fig. 11. In contrast to the previously simulated cases, there is a considerable difference between CGM and YSI measurements, available only at (infrequent) time intervals. The sampling rate is however small enough such that interpolation with second-order polynomials does not result in excessive overshoots. The figure does not only show that individual and interval model are well suited to represent dynamics of T1DM patients, but also that the estimated parameter intervals are small enough such that min/max simulations (7) result in reasonably small BG prediction intervals. A detailed view of the three breakfast periods is given in Fig. 8 for patient #0115.

Table 5 complements Table 3 and demonstrates the usefulness of the proposed methodology for real measurements on a group of 28 patients (only 10 are listed for

Table 4 Estimated parameters and performance metrics for simulated patient #0107

	Meal #1	Meal #2	Meal #3	Mean	Mean/Std.
Individual model					
Fit	91.59	93.51	91.39	92.16	78.85
ΔBG	−5.75	4.04	−2.19	3.99	2.24
ΔT	5.00	4.00	−1.00	3.33	1.60
K_1	5.18(±0.16)	4.34(±0.04)	5.68(±0.12)	5.07	7.50
K_2	−65.64(±1.86)	−49.01(±0.51)	−76.76(±0.76)	−63.80	−4.57
T_1	20.15(±0.38)	18.14(±0.20)	20.67(±0.29)	19.65	14.70
T_2	67.71(±3.63)	72.41(±1.68)	56.00(±1.74)	65.37	7.73
Interval model					
Fit	91.27	89.52	86.11	88.97	33.89
ΔBG	−6.87	7.76	−10.77	8.47	4.14
ΔT	5.00	6.00	−8.00	6.33	4.15
K_1	4.94	4.69	4.55	4.72	24.03
K_2	−61.59	−55.36	−66.39	−61.11	−11.05
T_1	19.44	20.12	17.15	18.90	12.16
T_2	71.32	73.74	67.86	70.97	24.01

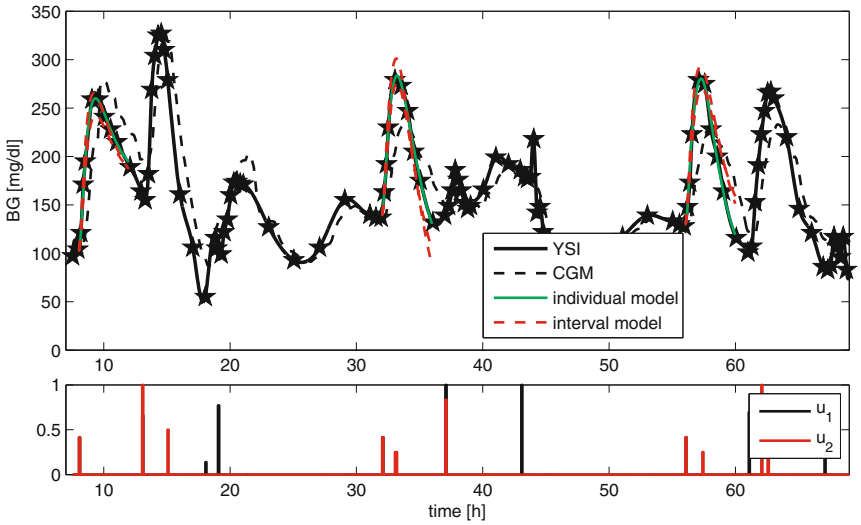

Fig. 11 Patient #0107: YSI measurements (*stars*), interpolated with second-order polynomial, CGM readings (*black dashed*), individual model outputs for the breakfasts (*green solid*), and worst-case interval model outputs for breakfasts (*red dashed*). All 3 days were used for estimation of the models. *Bottom panel* shows the scaled impulse carbohydrate (*black solid*) and insulin (*red solid*) inputs

Table 5 Statistics for clinical data, mean values of the performance metrics

	Individual model			Interval model		
Pat	Fit	ΔBG	ΔT	Fit	ΔBG	ΔT
P0102	84.84	5.70	12.33	83.07	7.83	11.00
P0103	71.35	1.13	18.00	56.99	5.34	20.00
P0104	82.10	4.55	11.33	75.51	11.92	16.00
P0106	81.81	7.93	6.00	70.56	12.93	5.67
P0107	92.16	3.99	3.33	88.97	8.47	6.33
P0108	66.21	9.49	5.67	55.82	15.89	6.67
P0109	90.67	3.82	6.67	84.52	11.57	10.67
P0110	58.81	8.36	21.67	49.63	12.79	45.67
P0111	71.14	10.10	19.00	64.53	10.86	13.00
P0112	68.27	7.58	33.67	58.43	15.17	45.33
Avg.*	77.62	5.78	17.76	68.21	10.91	20.64

*average values calculated based on 26 patients.

the sake of brevity). Out of this group, only for 2 patients no suitable, i.e., sufficiently small parameter intervals could be found because the variability was too high.

Remark 2 The estimation of the model parameters was done based on YSI measurements, which are probably not available in practice. It should be noted that the

proposed model identification procedure is independent of the source of the measurements, the YSI data can be replaced with data from glucose meters which are usually calibrated to show plasma equivalent values. As pointed out in the literature, e.g., in [7], the difference between those two measures can be up to 11 %, with plasma being higher than whole blood values, which directly effects the models. Similarly, a model estimated with YSI data could be used afterwards with a BG meter which introduces an uncertainty in the offset, directly related to the accuracy of the glucose meter.

Remark 3 The proposed model structure approximates BG dynamics based on carbohydrate and insulin inputs. There are several other factors which have an impact, e.g., physical activity, stress, and variation in nutrients in different meals. Activity did not play an important role in the analyzed data, because patients were hospitalized and stress could not be assessed with the available measurements. Complete information of nutrients (carbohydrates, proteins, and lipids) was available, but not utilized. Extended versions of the model structure making use of those data were tested (results not shown), but did not result in improvements. Changing meal compositions cause variations in the individual BG responses and are therefore indirectly captured by the interval model, resulting in enlarged intervals.

5 A Probabilistic Approach

Classical methods use various mathematical techniques to predict a single value in the future (e.g., the expected value in 30 min.). Contrary to that, in the preceding section, interval models were proposed which estimate a predicted range of the future BG. Probabilistic models, the topic of this section, go one step further and are solely concerned with the probability that the future BG will be inside a given region defined on the basis of physiological criteria. This corresponds more to the interest of the clinician, who aims to keep BG in a good range and not to some specific value. This allows also redefining the control task in terms of minimizing a given transition probability.

To estimate the mentioned probabilities for certain BG regions, we propose Gaussian models, in particular Gaussian mixture models (GMM), which are commonly used for modeling of biometric and biological systems with continuous measurements (see for example [15, 16] for an in-depth discussion). According to [39], a GMM is defined as a parametric probability density function represented as a weighted sum of Gaussian component densities.

5.1 Gaussian and Generalized Gaussian Mixture Models

Let us assume that a given (biological) system has n measured inputs. If we assume that each input is continuously measured with a sampling frequency f, a data vector recorded over given time t will have a total of $m = f \cdot t$ samples. Combined together, the individual measurement vectors of the system variables will form a measurement matrix x of dimensions $n \times m$. A Gaussian Mixture Model of the system is given as

$$p(x|\lambda) = \sum_{i=1}^{n} w_i \cdot q\,(x|\mu_i, \Sigma_i) \tag{9}$$

where $w_i, i = 1, \ldots n$ are the mixture weights and the component Gaussian densities are determined as

$$q\,(x|\mu_i, \Sigma_i) = \frac{1}{(2 \cdot \pi)^{m/2}\,|\Sigma_i|^{1/2}} \cdot e^{-\frac{1}{2} \cdot (x-\mu_i)' \cdot \Sigma_i^{-1} \cdot (x-\mu_i)} \tag{10}$$

with mean vector μ_i and covariance matrix $\Sigma_i, i = 1, \ldots n$. The mixture weights satisfy the constraint that

$$\sum_{i=1}^{n} w_i = 1 \tag{11}$$

The complete GMM is represented by the mean vectors, covariance matrices, and mixture weight from all variable densities:

$$\lambda = \{w_i, \mu_i, \Sigma_i\}, \quad i = 1, \ldots, n \tag{12}$$

There are several well-established algorithms for estimation of the parameters of GMM from measurement data. Two of them, namely the iterative expectation-maximization (EM) algorithm and the maximum a-posteriori (MAP) from a well-trained prior model are most frequently used. More details about these two algorithms can be found, for example, in [39] or [40].

However, the statistical analysis of the BG data used in this study showed that, in the general case, they cannot be represented by the Gaussian distribution. In principle, this problem can be solved by the approximation of a non-Gaussian distribution. The near-normal distributions, i.e., the distributions which moderately deviate from the Gaussian distribution, can be approximated using disjunction into Gaussian components. The details about this well-known approach can be found, for example, in [27]. Now, modeling of the data with non-Gaussian distributions would be possible using its Gaussian components. This method is known as the independent component analysis (ICA). The fundamental restriction in ICA is that the independent components must be non-Gaussian for ICA to be possible [27]. Under the assumption that the data is distributed according to multivariate Gaussian distribution, i.e., where the

components are also distributed according to Gaussian distribution, the ICA method is equivalent to the principal component analysis (PCA). The generalized mixture model using ICA is discussed in [32]. This model is an extension of the GMM where the model components have non-Gaussian distributions. The deviation of Gaussian distribution is modeled using the generalized Gaussian density. It is assumed that function $p(x|\mu)$ is a differentiable function μ, so that the log-likelihood is given as:

$$L = \sum_{i=1}^{n} \log p(x|\mu_i) \tag{13}$$

Thus, the class probabilities for the data matrix x are given as

$$q\left(\lambda_k|x_n, \Sigma\right) = \frac{p\left(x_n|\Sigma_k, \lambda_k\right) \cdot p(\lambda_k)}{\sum_k p\left(x_n|\Sigma_k, \lambda_k\right) \cdot p(\lambda_k)} \tag{14}$$

The detailed discussion about the GMM using ICA, especially about generation of model parameters, can be found, e.g., in [31] or [32]. The cited references also give examples of applications of these models.

5.2 Modeling Method and Model Structure

The main goal of this study was to generate probabilistic mixture models of the blood glucose levels and to use the generated models for short-term predictions, up to 30 min in advance. Those predictions are being calculated on the basis of the current and previous values of BG.

As already mentioned, the models were not designed to predict the exact future BG concentrations in patients. Instead, they predict in which of five predefined regions (which could be easily changed and adapted) the BG would most probably lie after the prediction interval:

<70 [mg/dl] (hypoglycemia)
70–120 [mg/dl] (normglycemia)
120–180 [mg/dl] (elevated normglycemia)
180–300 [mg/dl] (hyperglycemia)
>300 [mg/dl] (acute hyperglycemia)

Data from various patients are, in general case, not normally distributed. Under such circumstances, the blood glucose levels of such patients must be modeled using generalized GMM, as previously discussed in Sect. 5.1.

Several sources have already reported about periodic properties of the BG concentration in the human body. A reader can find in-depth analyses, e.g., in [41] or [22]. Starting from the assumption that the BG exhibits periodic properties, the fast Fourier transformation (FFT) analysis has been conducted in order to determine the

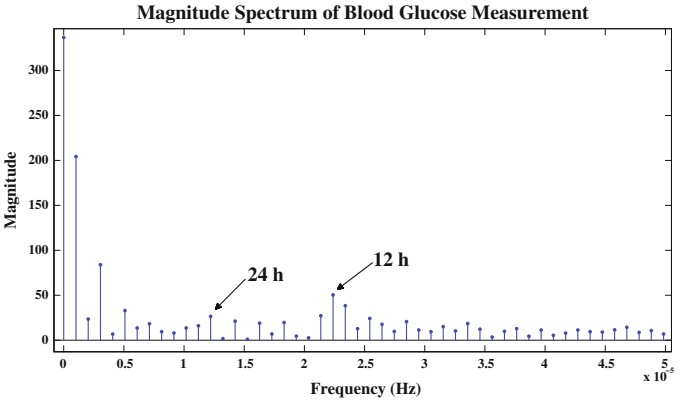

Fig. 12 Magnitude spectrum of the BG concentration measurements for Patient #6

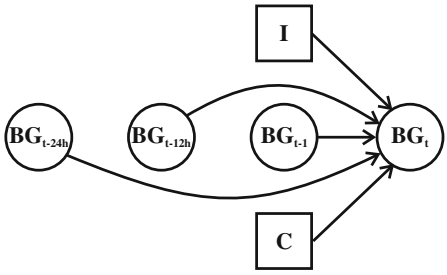

Fig. 13 Model structure used in generation of GMM of the BG concentration for patients selected for the study

dominant periodic components in the measured data. Figure 12 shows the magnitude spectrum of one representative patient (Patient #6 from clinical study [18]). The results are obtained by analyzing a 7-days blood glucose time series. Figure 12 shows the frequencies of the recurring patterns and the estimation of the importance of each cycle in the glucose time series. As it can be seen in the figure, a 12-h pattern repetition appears to be the most significant. Besides the 12-h pattern, the 24-h pattern has also been identified as important due to the circadian nature of glucose control [6]. The confirmation of these results can also be found in the extensive study of frequency characterization of blood glucose dynamics [22].

Thus, based on the previous considerations, a model structure shown in Fig. 13 has been selected. It includes the previous levels of blood glucose as inputs, specifically the ones measured 24 h before (BG_{t-24h}), 12 h before (BG_{t-12h}) and the last available measurement (BG_{t-1}). The model structure also includes the intake of insulin (I) and carbohydrates (C). It is important to note that, differently from the continuous 24-h blood glucose level measurements, the intake of insulin and carbohydrates were discrete events typically occurring 3–6 times a day during the clinical stay of the patients. It is also important to mention that the carbohydrate values used are not measured but estimated on the basis of the actual content of patients meals.

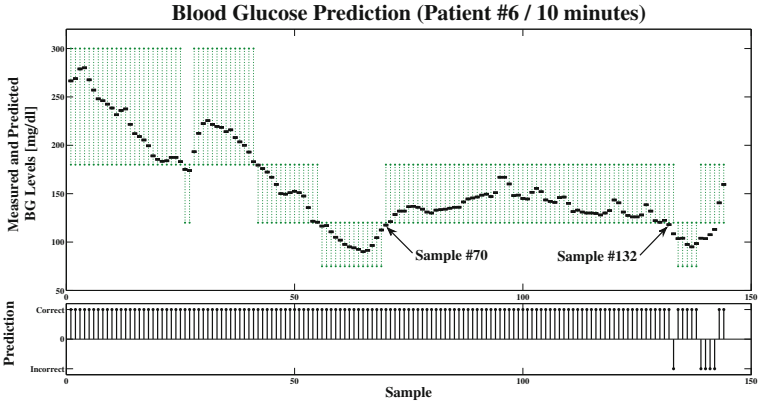

Fig. 14 10-min prediction of the BG concentration levels (Patient #6) (color figure online)

5.3 Modeling Results

The data used here were collected in the clinical trial [18] conducted at the Institute for Diabetes-Technology GmbH in Ulm, Germany between October and December 2011. In this study, a total of 12 patients with type-1 diabetes spent 7 days hospitalized and were equipped with six CGM sensors in parallel. For the purpose of the present study, the data from one of the sensors (a DexCom™Seven® Plus sensor by DexCom, San Diego, CA) was used which provides continuous information on subcutaneous glucose for 7 days with a sampling time of 5 min. Additionally, we used recorded information about meal intakes and bolus insulin injections.

The measurement data from clinical patients used in this study were divided into two parts. The larger part of the data which corresponded to the first 6 clinical days, was used to train the GMM. The smaller part of the data corresponded to the final (seventh) clinical day and was used for prediction and testing of the trained models. The models were used first to predict the blood glucose levels during the seventh clinical day and then to compare the prediction with the actual measurements in order to assess the correctness of the model.

Figure 14 shows the results of prediction of the BG levels (concentrations) for a clinical patient (designated for the purpose of this study as Patient #6), in particular the predictions of the final clinical day (24 h). The measured blood glucose levels from the previous 6 clinical days are used to generate the prediction model. The measurements and predictions in Fig. 14 are shown in 10 min equidistance (or time slices). To prevent possible misunderstanding, it is important to underline that, although the results are shown in 10 min time slices, the figures show future trends 10 min from the observation point at any particular time.

Figure 14 is divided into two parts. The upper part shows the actual (measured) blood glucose levels (showed using black dots) and the predicted blood glucose intervals (showed as vertical green lines limited at their ends by dots). The lower portion

of figures indicates if the prediction above was correct or incorrect. A particular prediction was considered as a correct one, in a strict sense, if the actual measurement fell within its limits.

However, some of the predictions which are classified as incorrect can be found at or very close to the limits of the prediction segments and they are actually predicting the general trend of the blood glucose signal well. Examples of that can be seen in Fig. 14 (e.g., samples #70 or #132). Therefore, some of those slightly incorrect predictions can be considered to be borderline cases. This was one of the reasons to define an evaluation algorithm which could cope with different kinds of borderline behavior of the BG models. The evaluation algorithm, therefore, includes the following rules:

- Segment overlapping: each two neighboring segments are overlapped for 5 [mg/dl]; therefore, at some prediction instances the modeling algorithm can extend the width of the prediction segments.
- Persistence principle: in order to eliminate the prediction outliers, all single prediction segments which are in opposition to their neighboring prediction segments (from left and right) are changed to correspond to their neighboring predictions; this rule is not applied in cases when there is a persistent transition from one to another prediction segment.

The results shown in Fig. 14 are obtained by the application of so defined evaluation algorithm. The model has 96.53 % correct prediction rate, having only 3.47 % incorrect predictions during the final 24-h period.
The evaluation algorithm was able to deliver usable predictions of blood glucose trends up to 30 min in the future. The future predictions are generated stepwise, in 10-min steps, where the second (20-min prediction) and the third step (30-min prediction) are using the predictions from their preceding steps in calculations. The 20-min predictions are, therefore, calculated on the basis of the preceding 10-min. predictions, and the 30-min predictions are calculated on the basis of the preceding 20-min predictions. As a consequence, the evaluation algorithm was able to make corrections of the results with 10 min delay.

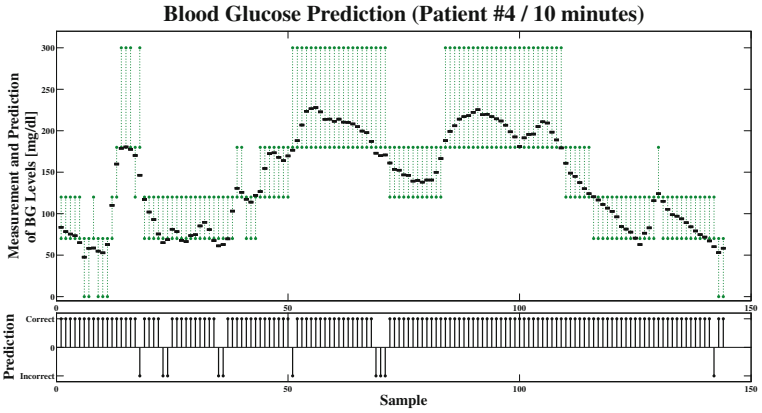

Fig. 15 10-min prediction of the BG concentration (Patient #4)

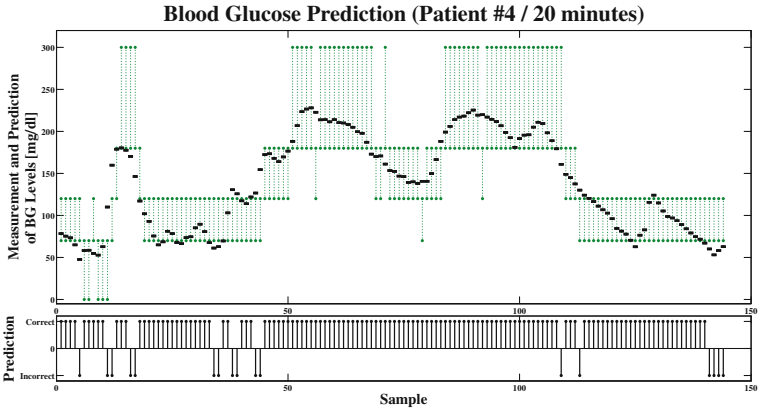

Fig. 16 20-min prediction of the BG concentration (Patient #4)

Figures 15, 16 and 17 show blood glucose prediction results for a clinical patient (Patient #4). The figures show the 10-, 20-, and 30-min predictions of the final (seventh) clinical day of Patient #4. Similarly as in the case of Patient #6, the measured blood glucose levels from the previous six clinical days are used to generate the prediction model. The model performance for the tests shown in Figs. 15, 16 and 17 are summarized in Table 6.

For the purpose of comparison, the blood glucose levels of the same patient (Patient #4) are modeled using linear second-order autoregressive models (AR) for the same prediction horizons (10-, 20-, and 30-min). The linear AR models used the same data as the GMM models. However, due to the fundamental difference between these two model types (the linear AR models predict single values, while the proposed prediction method using GMM models predicts intervals with certain probabilities), in order to make the comparison possible, the single value predictions of the linear

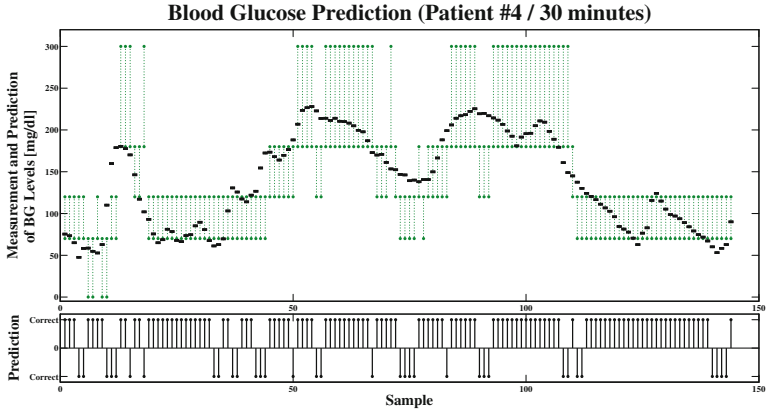

Fig. 17 30-min prediction of the BG concentration (Patient #4)

Table 6 Correct and incorrect prediction rates for 10-, 20-, and 30-min predictions of the BG concentrations for Patient #4 (GMM and linear AR model)

Patient #4	GMM model		Linear AR model	
	Correct %	Incorr. %	Correct %	Incorr. %
10-min prediction	**93.06**	**6.94**	96.56	3.45
20-min prediction	**88.19**	**11.81**	84.83	15.17
30-min prediction	**76.39**	**23.61**	67.59	32.41

AR models are used to predict the blood glucose intervals. Figure 18 shows both the single value predictions as well as the corresponding interval predictions of blood glucose level; the model performance for the tests shown in Fig. 18 is also summarized in Table 6.

In general, based on the results summarized in Table 6, it can be concluded that the GMM models deliver mostly correct predictions. The correct prediction rate for the 20- and 30-min predictions is expectedly lower than for the 10-min prediction because, as explained earlier, the 20- and 30-min predictions use the outputs from their preceding predictions which are already uncertain to the some extent.

The comparison of the GMM models with the linear AR models showed that the linear AR models exhibit slightly better 10-min predictions, while the GMM models exhibit superior results for 20- and especially 30-min predictions. However, the performance of the linear AR models in predicting intervals should be taken with caution. The linear AR models predict single values of blood glucose concentration. Thus, a probability could not be assigned to the BG intervals detected using those single values predictions of the linear AR models, making them not entirely comparable to the GMM models. Instead, if a fit value is used to assess the performance of the linear AR models, they showed much worse results: 84.76 %, 68.82 %, and 54.01 % fit for 10-, 20-, and 30-min predictions, respectively. Thus it can be concluded that

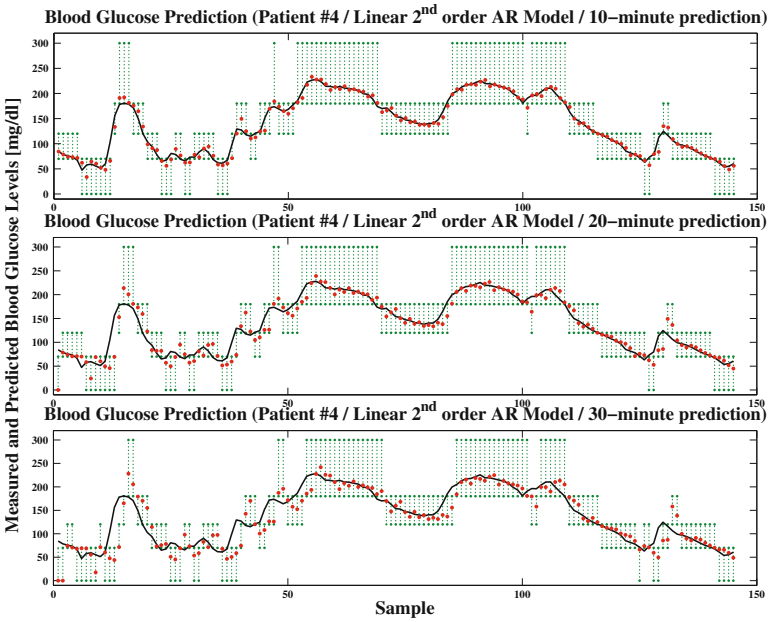

Fig. 18 10-, 20- and 30-min prediction of the BG concentration using linear second-order AR models (Patient #4)

the proposed method using GMM models showed superior performance especially as the prediction window increased. It is important to stress that in no case a wrong prediction of the proposed method using GMM models could have led to a dangerous patient decision (as in the case of a hyperglycemia predicted when the patient was in fact in hypoglycemia).

6 Conclusion and Outlook

The increasing availability of new sensors, smart insulin pumps, and more in general of computing power opens new chances for personalized diabetes care, in particular for closed-loop control, as in the artificial pancreas. First results in this direction are promising.

Nevertheless, variability, in particular intrapatient variability, remains a challenge. It also means that some methods may work well for some patient in most cases, but not for all patients and all cases. Of course, having a perfect model which can predict the future BG value with a high precision would be the best. Clinical experience, however, shows that some conditions, like anger, can affect enormously the BG development without being predictable.

This chapter intended to show that it might be wiser to accept this limitation from the very beginning and to look for "range" models, either in the form of moving or of prefixed ranges. While they will almost never provide an exact prediction of the next BG values, their error can be limited in such a way to be clinically easier to cope with than classical models.

Of course, the validity of the proposed approach needs testing in a clinical study and there is still much work to be done to tailor the methods to the specific problem.

Acknowledgments This chapter is derived in part from an article published in the International Journal of Control, published online 25 Mar 2014, copyright by Informa UK Limited trading as Taylor & Francis Group, available online: http://www.tandfonline.com/doi/full/10.1080/00207179.2014.897004, as well as in part from an article published in the Proceedings of the 2014 22nd Mediterranean Conference of Control and Automation (MED), copyright by The Institute of Electrical and Electronics Engineers, Incorporated (IEEE), available online: http://dx.doi.org/10.1109/MED.2014.6961587 .

References

1. Bequette, B.W.: Challenges and recent progress in the development of a closed-loop artificial pancreas. Annu. Rev. Control **36**(2), 255–266 (2012). doi:10.1016/j.arcontrol.2012.09.007
2. Bergman, R.N.: Minimal model: perspective from 2005. Horm. Res. **64**(suppl 3), 8–15 (2005)
3. Bergman, R.N., Ider, Y.Z., Bowden, C.R., Cobelli, C.: Quantitative estimation of insulin sensitivity. Am. J. Physiol. **236**(6), E667–677 (1979)
4. Boiroux, D., Schmidt, S., Frøssing, L., Nørgaard, K., Madsbad, S., Skyggebjerg, O., Duun-Henriksen, A.K., Poulsen, N.K., Madsen, H., Jørgensen, J.B.: Control of blood glucose for people with type 1 diabetes: an in-vivo study. In: Proceedings of the 17th Nordic Process Control Workshop, pp. 133–140 (2012)
5. Castillo Estrada, G., Del Re, L., Renard, E.: Nonlinear gain in online prediction of blood glucose profile in type 1 diabetic patients. In: Proceedings of the 49th IEEE Conference on Decision and Control (CDC), pp. 1668–1673 (2010)
6. Castillo Estrada, G.E.: Identification of glucose dynamics in type 1 diabetes patients using multiple daily injection therapy. Ph.D. thesis, Johannes Kepler University Linz, Austria (2011)
7. Cengiz, E., Tamborlane, W.V.: A tale of two compartments: interstitial versus blood glucose monitoring. Diabetes Tech. Ther. **11**(S1), S11–S16 (2009)
8. Dalla Man, C., Guerra, S., Sparacino, G., Renard, E., Cobelli, C.: A reduced type 1 diabetes model for model predictive control. In: Book of abstracts, 3rd International Conference on Advanced Technologies and Treatments for Diabetes (ATTD), Basel, Switzerland (2010)
9. Cescon, M., Stemmann, M., Johansson, R.: Impulsive predictive control of T1DM glycemia: an in-silico study. In: 2012 ASME Dynamic Systems and Control Conference. Fort Lauderdale, FL, USA (2012)
10. Cescon, M., Johansson, R., Renard, E., Maran, A.: Identification of individualised empirical models of carbohydrate and insulin effects on T1DM blood glucose dynamics. Int. J. Control **87**(7), 1438–1453 (2014). doi:10.1080/00207179.2014.883171
11. Dalla Man, C., Micheletto, F., Lv, D., Breton, M., Kovatchev, B., Cobelli, C.: The UVA/PADOVA type 1 diabetes simulator: new features. J. Diabetes Sci. Technol. **8**(1), 26–34 (2014)
12. Dalla Man, C., Rizza, R.A., Cobelli, C.: Meal simulation model of the glucose-insulin system. IEEE Trans. Biomed. Eng. **54**(10), 1740–1749 (2007). doi:10.1109/TBME.2007.893506
13. del Re, L., Grünbacher, E., Pichler, K.: Verfahren und Vorrichtung zum Diagnostizieren des Zustandes eines Maschinenbauteils (2008)

14. DIAdvisor: European commission IST programme FP7 project DIAdvisor, n° 216592. http://www.diadvisor.org
15. Duda, R., Hart, P.: Pattern Classification and Scene Analysis. Wiley, New York (1973)
16. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, 2 edn. Wiley (2000). ISBN 978-0-471-05669-0
17. Fang, Q., Yu, L., Li, P.: A new insulin-glucose metabolic model of type 1diabetes mellitus: an in silico study. Comput. Method. Program. Biomed. **120**, 16–26 (2015)
18. Freckmann, G., Pleus, S., Link, M., Zschornack, E., Klötzer, H.M., Haug, C.: Performance evaluation of three continuous glucose monitoring systems: comparison of six sensors per subject in parallel. J. Diabetes Sci. Technol. **7**(4), 842–853 (2013)
19. Garnier, H., Mensler, M., Richard, A.: Continuous-time model identification from sampled data: implementation issues and performance evaluation. Int. J. Control **76**, 1337–1357 (2003)
20. Garnier, H., Young, P.: What does continuous-time model identification have to offer? In: Proceedings of the 16th IFAC Symposium on System Identification, vol. 16 (2012)
21. Goodyear, L.J., Kahn, B.B.: Exercise, glucose transport, and insulin sensitivity. Annu. Rev. Med. **49**, 235–261 (1998)
22. Gough, D.A., Kreutz-Delgado, K., Bremer, T.M.: Frequency characterization of blood glucose dynamics. Ann. Biomed. Eng. **31**(3), 91–97 (2003)
23. Guyton, A.C.: Textbook of Medical Physiology, 11th edn. Elsevier Inc., Philadelphia (2006)
24. Hinshaw, L., Dalla Man, C., Nandy, D.K., Saad, A., Bharucha, A.E., Levine, J.A., Rizza, R.A., Basu, R., Carter, R.E., Cobelli, C., Kudva, Y.C., Basu, A.: Diurnal pattern of insulin action in type 1 diabetes, implications for a closed-loop system. Diabetes **62**(7), 2223–2229 (2013)
25. Hovorka, R., Canonico, V., Chassin, L.J., Haueter, U., Massi-Benedetti, M., Federici, M.O., Pieber, T.R., Schaller, H.C., Schaupp, L., Vering, T., Wilinska, M.E.: Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. Physiol. Meas. **25**(4), 905–920 (2004)
26. Hovorka, R., Elleri, D., Thabit, H., Allen, J.M., Leelarathna, L., El-Khairi, R., Kumareswaran, K., Caldwell, K., Calhoun, P., Kollman, C., Murphy, H.R., Acerini, C.L., Wilinska, M.E., Nodale, M., Dunger, D.B.: Overnight closed-loop insulin delivery in young people with type 1 diabetes: a free-living, randomized clinical trial. Diabetes Care **37**(5), 1204–1211 (2014). doi:10.2337/dc13-2644
27. Hyvärinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. Wiley, New Jersey (2001)
28. Kirchsteiger, H., del Re, L.: Robust tube-based predictive control of blood glucose concentration in type 1 diabetes. In: IEEE 52nd Annual Conference on Decision and Control (CDC), 2013, pp. 2084–2089 (2013). doi:10.1109/CDC.2013.6760189
29. Kirchsteiger, H., Castillo Estrada, G., Pölzer, S., Renard, E., Del Re, L.: Estimating interval process models for type 1 diabetes for robust control design. In: Proceedings of the 18th IFAC World Congress, pp. 11,761–11,766. Milano, Italy (2011)
30. Kirchsteiger, H., Johansson, R., Renard, E., Re, L.D.: Continuous-time interval model identification of blood glucose dynamics for type 1 diabetes. Int. J. Control **87**(7), 1454–1466 (2014). doi:10.1080/00207179.2014.897004
31. Lee, T.W., Lewicki, M.S.: The generalized gaussian mixture model using ica. In: Proceedings of International Workshop ICA, pp. 239–244 (2000)
32. Lee, T.W., Lewicki, M., Sejnowski, T.: Unsupervised classification with non-gaussian mixture models using ICA. Adv. Neural Inf. Process. Syst. **11**, 508–514 (1999)
33. Leelarathna, L., Dellweg, S., Mader, J.K., Allen, J.M., Benesch, C., Doll, W., Ellmerer, M., Hartnell, S., Heinemann, L., Kojzar, H., Michalewski, L., Nodale, M., Thabit, H., Wilinska Malgorzata, E., Pieber, T.R., Arnolds, S., Evans, M.L., Hovorka, R.: On behalf of the AP@home Consortium: day and night home closed-loop insulin delivery in adults with type 1 diabetes: three-center randomized crossover study. Diabetes **37**, 1931–1937 (2014)
34. Ljung, L.: System Identification, 2nd edn. PTR Prentice Hall (1999)
35. Milanese, M., Norton, J., Piet-Lahanier, H., Walter, E. (eds.): Bounding Approaches to System Identification. Kluwer Academic/Plenum Publishing Corporation, London (1996)

36. Percival, M.W., Bevier, W.C., Wang, Y., Dassau, E., Zisser, H.C., Jovanovic, L., Doyle, F.J.: Modeling the effects of subcutaneous insulin administration and carbohydrate consumption on blood glucose. J. Diabetes Sci. Technol. **4**(5), 1214–1228 (2010)
37. Reiterer, F., Kirchsteiger, H., Assalone, A., Freckmann, G., del Re, L.: Performance assessment of estimation methods for CIR/ISF in bolus calculators. In: Proceedings of the 9th IFAC Symposium on Biological and Medical Systems, pp. 231–236 (2015)
38. Reiterer, F., Kirchsteiger, H., Freckmann, G., del Re, L.: Identification of diurnal patterns in insulin action from measured CGM data for patients with T1DM. In: Proceedings of the 2015 European Control Conference, pp. 1–6 (2015)
39. Reynolds, D.: Encyclopedia of Biometric Recognition, chap. Gaussian mixture models, p. 1217. MIT Lincoln Laboratory, Cambridge, MA (2008)
40. Reynolds, D., Rose, R.: Robust text-independent speaker identification using gaussian mixture speaker models. IEEE Trans. Speech Audio Process. **3**, 72–83 (1995)
41. Riva, A., Bellazzi, R.: Learning temporal probabilistic causal models from longitudinal data. Artif. Intell. Med. **8**, 217234 (1996)
42. Saad, A., Dalla Man, C., Nandy, D.K., Levine, J.A., Bharucha, A.E., Rizza, R.A., Basu, R., Carter Rickey, E., Cobelli, C., Kudva, Y.C., Basu, A.: Diurnal pattern to insulin secretion and insulin action in healthy individuals. Diabetes **61**(11), 2691–2700 (2012)
43. Söderström, T., Stoica, P.: System Identification. Prentice Hall International, Hemel Hempstead (1989)
44. Sorensen, J.T.: A physiologic model of glucose metabolism in man and its use to design and assess improved insulin therapies for diabetes. Ph.D. thesis, Massachusets Institute of Technology (1978)
45. Trogmann, H., Kirchsteiger, H., Castillo Estrada, G., del Re, L.: Fast estimation of meal/insulin bolus effects in t1dm for in silico testing using hybrid approximation of physiological meal/insulin model. 70th Scientific Sessions of the American Diabetes Association (2010). Abstract 504-P
46. Van Cauter, E., Polonsky, K.S., Scheen, A.J.: Role of circadian rythmicity and sleep in human glucose regulation. Endocr. Rev. **18**(5), 716–738 (1997)
47. Visentin, R., Dalla Man, C., Kovatchev, B., Cobelli, C.: The university of virginia/padova type 1 diabetes simulator matches the glucose traces of a clinical trial. Diabetes Technol. Ther. **16**(7), 428–434 (2014)

Accuracy of BG Meters and CGM Systems: Possible Influence Factors for the Glucose Prediction Based on Tissue Glucose Concentrations

Guido Freckmann, Stefan Pleus, Manuela Link and Cornelia Haug

Abstract The goal of this paper is to describe the metrics used for the evaluation of accuracy of blood glucose (BG) meters for self-monitoring of blood glucose (SMBG) and continuous-glucose monitoring (CGM) system and their limitations and to discuss the current status of SMBG and CGM accuracy. SMBG measurement is used by patients for therapy control and for calculation of appropriate insulin doses for approximately 30 years. The minimum accuracy criteria for SMBG meters are currently defined by ISO 15197:2003 (at least 95 % of results within ±20 % or ±15 mg/dL of the comparison method measurement results for BG concentrations above or below 75 mg/dL, respectively). In 2013, these accuracy limits were revised in the standard ISO 15197:2013: at least 95 % of results within ±15 % or ±15 mg/dL for BG above or below 100 mg/dL, respectively. SMBG systems are also used by patients for calibration of CGM systems. Therefore, precision and trueness of the SMBG system are influencing the accuracy of the CGM results. The timing of the BG measurement used for calibration has to be taken into account because, during rapid glucose changes, a time lag exists between BG and the tissue glucose that is measured by CGM systems. The accuracy of CGM devices is often reported by the mean absolute relative deviation (MARD) between CGM results and BG comparison results. This parameter is influenced by different factors like study procedures, glucose fluctuations during the study, and distribution of comparison BG measurements. It is important to define standard study procedures and evaluations to be able to compare MARD results from different studies. For the correct prediction of glucose concentrations, the specific prediction method as well as the accuracy of the CGM system, which may be affected by the accuracy of the SMBG system used for calibration, and the timing of the calibration are important aspects.

G. Freckmann (✉) · S. Pleus · M. Link · C. Haug
Institut für Diabetes-Technologie Forschungs- und Entwicklungsgesellschaft mbH an der Universität Ulm, Helmholtzstrasse 20, 89081 Ulm, Germany
e-mail: guido.freckmann@uni-ulm.de

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*,
Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_2

1 Introduction

Methods for prediction of future glucose concentration based on previous glucose using continuous subcutaneous glucose monitoring (CGM) data are currently under development. A prediction of future glucose could be beneficial to patients; early detection of glucose concentration falling below certain low glucose thresholds could, for example, help to prevent hypoglycemia. Glucose prediction can be combined with algorithms for closed-loop control to optimize glycemic control.

In current therapy, most patients treated with multiple daily insulin injections (MDI) or continuous subcutaneous insulin infusion (CSII) use self-monitoring of blood glucose (SMBG). SMBG measurement is used by patients for therapy control and for calculation of appropriate insulin doses for more than 30 years. MDI and CSII patients measure 3–8 single spot blood glucose values per day. For adequate glucose prediction, a much higher frequency of glucose values is needed. This higher frequency of glucose values can, for example, be achieved by using CGM systems.

In the last decade, different CGM systems became available that provide a glucose value every 1–10 min. The current CGM systems are measuring the glucose in the subcutaneous tissue and are calibrated against SMBG values. Thus, SMBG accuracy influences the accuracy of CGM systems and, consequently, the accuracy of glucose predictions based on CGM results. Additionally, the timing of the calibration is an important aspect, since time delays between BG and tissue glucose, which is measured by CGM results, can be observed when glucose concentration is changing rapidly.

The goal of this paper is to describe the metrics used for the evaluation of accuracy of SMBG meters and CGM systems and their limitations and to discuss the current status of SMBG and CGM accuracy.

2 SMBG Accuracy and CGM Calibration with SMBG Results

All currently available CGM systems are calibrated on blood glucose values provided by SMBG. The accuracy of the BG value and the timing of the BG measurement used for calibration are important for the accuracy of the CGM results.

2.1 SMBG Accuracy

SMBG measurement is currently used by patients for therapy control and for calculation of appropriate insulin doses. Therefore, the measured blood glucose values should be accurate to avoid miscalculation of insulin doses or failure to detect hypo- and hyperglycemia. The American Diabetes Association stated in 1987, that a total

accuracy of 10 % should be achieved [6]. In 1994, this goal was revised to an analytical accuracy of 5 % [1]. Today, more than 20 years later, these accuracy goals are still not achieved by most SMBG meters.

The minimum acceptable accuracy for results produced by a SMBG system is currently defined by ISO 15197:2003: at least 95 % of results have to be within ±20 % or ±0.83 mmol/L (15 mg/dL) of the comparison method measurement results (manufacturers measurement procedure) for BG concentrations above or below <4.2 mmol/L (<75 mg/dL), respectively [15]. In addition, ISO 15197 calls for the display of the system accuracy results of SMBG systems for glucose concentrations <4.2 mmol/L (75 mg/dL), as the percentage of values falling within the following intervals: ±0.28 mmol/L (±5 mg/dL), ±0.56 mmol/L (±10 mg/dL), and ±0.83 mmol/L (±15 mg/dL). For glucose concentrations >4.2 mmol/L (75 mg/dL), results shall be expressed as the percentage of values falling within the following intervals: ±5, ±10, ±15 and ±20 %.

In 2013, this ISO standard was revised (ISO 15197:2013) with the following accuracy criteria: 95 % of results have to be within ±15 % or ±15 mg/dL for BG concentrations above or below 100 mg/dL, respectively [16]. The new ISO standard also asks for the stricter limits in evaluations where the BG meter is used by laypersons. It also requires the assessment of clinical accuracy using the consensus error grid [16, 25]. Only limited data is available showing the accuracy in the hands of users, especially according to the revised ISO 15197. In Table 1, an example of how the system accuracy results can be displayed as recommended in ISO 15197:2013 is shown.

The difference plot in Fig. 1 is an example that shows the system accuracy results of three reagents lots of an SMBG system that was evaluated following ISO 15197 with all results within the limits of ISO 15197:2003 and 15197:2013 [29]. Unfortunately, the performance of the SMBG devices often is lower than shown in the example.

Table 1 Example for presentation of system accuracy results as recommended in ISO 15197:2013

System accuracy results for glucose concentration <5.55 mmol/L (<100 mg/dL)		
Within ±0.28 mmol/L (Within ±5 mg/dL)	Within ±0.56 mmol/L (Within ±10 mg/dL)	Within ±0.83 mmol/L (Within ±15 mg/dL)
68/150 (45.3 %)	105/150 (70.0 %)	143/150 (95.3 %)
System accuracy results for glucose concentration ≥ 5.55 mmol/L (≥ 100 mg/dL)		
Within ±5 %	Within ±10 %	Within ±15 %
221/450 (49.1 %)	383/450 (85.1 %)	439/450 (97.6 %)
System accuracy results for glucose concentrations between X.XX mmol/L (XX mg/dL) and YY.Y mmol/L (YYY mg/dL)		
Within ±0.83 mmol/L or ±15 % (Within ±15 mg/dL or ±15 %)		
582/600 (97.0 %)		

Note X.XX mmol/L (XX mg/dL) and YY.Y mmol/L (YYY mg/dL) stand for the lowest and highest glucose concentration measured with the comparison method, respectively

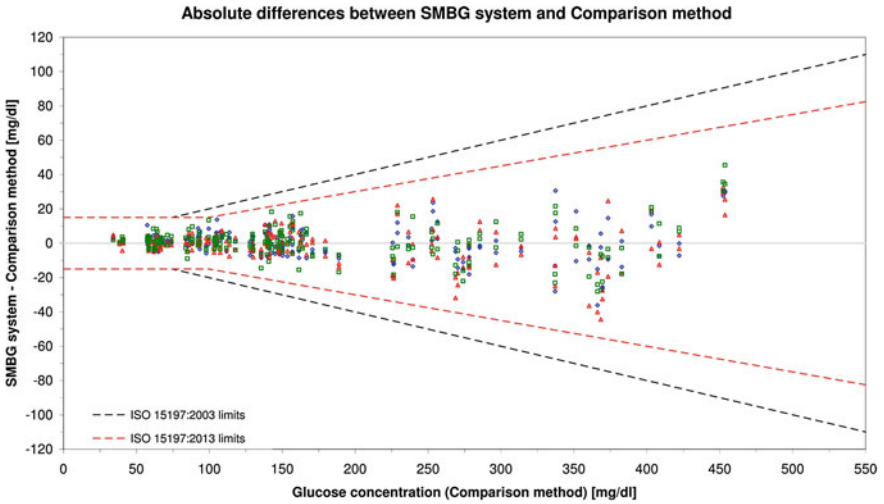

Fig. 1 Difference plot showing system accuracy results of an SMBG system with three reagent system lots. System accuracy limits of ISO 15197:2003 (*dashed black line*) and ISO 15197:2013 (*dashed red line*) were applied

Multiple studies were published during the last few years, testing meters applying the limits of the ISO 15197:2003 and ISO 15197:2013 standards. [3, 4, 9, 11, 12, 20, 26]. The quality of such studies differs, as only some studies are following the ISO 15197:2003 study protocol, while others have major deviations from the protocol [31]. All of these studies were comparing the system accuracy when applying the ISO 15197:2003 limits and they found that between 60 and 100 % of the investigated SMBG systems have at least 95 % of results within the ISO 15197:2003 limits. In some of these studies, the ISO 15197:2013 limits were applied as well, which were fulfilled by less than half of the investigated test strip lots. The observed differences in system accuracy can be attributed to or is influenced by a list of reasons including, but not limited to, the production process, the type and quality of test strip coding, user handling [30], and the manufacturers measurement method used for calibration. Measurement methods are reported to have systematic differences of up to 8 % [32]. Such factors can lead to systematic or random measurement errors.

2.2 CGM Calibration with SMBG Results

Tissue glucose is not readily available for measurement, thus the more easily available BG is measured by patients with SMBG systems in order to be used for calibration of CGM devices. However, when calibrating against SMBG results, not only does the accuracy of the SMBG system influence the quality of the calibration, but also the timing of the BG measurement subsequently used for CGM calibration.

The accuracy, i.e., precision and trueness, of the SMBG system used for calibration are influencing the accuracy of the CGM results. In 2009, Kamath and colleagues reported that the MARD of the CGM system they used could be nearly halved (16.0 % vs. 8.5 %) when switching from calibration against SMBG results to calibration against results from a laboratory method [17]. It is unclear whether the effect could be reduced with top-of-the-line SMBG systems, thus reducing the initial MARD.

Additionally, a time delay between the glucose values provided by BG meters and CGM systems is observed during rapid glucose changes. This time delay is composed by a physiologic time delay, which is independent from a specific CGM system used, and a technical time delay caused by the specific CGM system. The physiologic time delay reflects the time required for the diffusion of glucose from the blood capillaries into the subcutaneous tissue [18]. The technical time delay is caused by two main factors. First, the glucose has to diffuse through the CGM sensor membrane and onto the sensor; and second, the raw signal from the CGM sensors often is smoothened by an algorithm, which further increases the time delay. This time delay occurs at all times and is not limited to rapid glucose changes; however, it is most clearly visible during rapid glucose changes and its contribution to a measurement error is more pronounced during these rapid glucose changes than when glucose only changing slowly. Rapid changes of glucose are observed frequently after the ingestion of meals or during exercise. Subsequently, the timing of the calibration is important and CGM systems should be calibrated at times of minimal BG change. Zueger and colleagues found lower MARD when using preprandial calibration as compared to using postprandial calibration [35].

SMBG accuracy and CGM calibration are the two most common influences on the accuracy of CGM systems. However, there are other aspects as well. In one study, it was suspected that the SMBG system showed a pronounced variability between vials of test strips (see Fig. 2). In other cases, contaminated hands were found to heavily affect the accuracy of an otherwise good SMBG system [13, 14]. However, these influence factors are not common and they can be avoided by taking appropriate steps.

In summary, calibration of CGM systems should only be performed with high-accuracy SMBG systems, and phases of rapid glucose changes, i.e., within 1–2 h after carbohydrate intake, should be avoided for calibration of CGM systems.

In this context it should be noted that, recently, a factory-calibrated tissue glucose monitoring system was introduced in the European market. The term factory-calibrated is used for systems that do not have to be calibrated while wearing, as they have an internal calibration implemented during the system's production. It is possible that in the near future, more and more systems will be factory-calibrated. As such factory-calibrations can be performed in a controlled environment, it is highly likely that the error introduced by inappropriate calibration will be markedly lower than that introduced by online, prospective calibration with SMBG systems (Fig. 2).

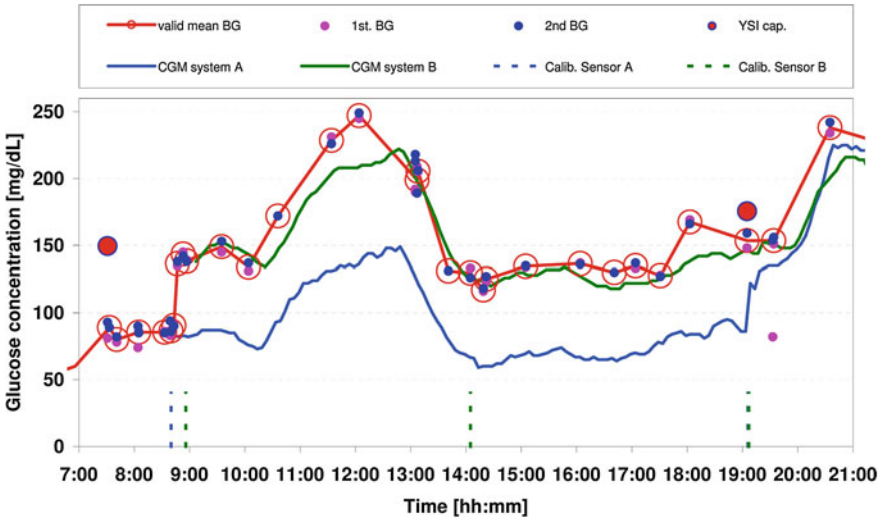

Fig. 2 Effect of SMBG accuracy on calibration of CGM systems. The increase in BG concentration at approximately 9:00 was presumably caused by vial-to-vial differences of the SMBG system. CGM system A was calibrated using test strips from one vial, CGM system B was calibrated using test strips from another vial

3 Accuracy of CGM Systems

In the literature, the accuracy of CGM systems is assessed by a number of different parameters. In the following section, two of these parameters are described in more detail, namely the mean absolute relative difference (MARD) and the precision absolute difference (PARD).

The MARD indicates how closely CGM results and BG results match. The PARD on the other hand indicates how closely two sensor traces of the same CGM system worn by the same subject at the same time follow each other.

A third parameter, the continuous-glucose error grid analysis (CG-EGA), while not described in detail, is worth mentioning. The CG-EGA provides a clinical assessment of the CGM systems accuracy. However, BG results have to be obtained at least every 15 min in order to adequately perform CG-EGA, which places an additional burden on both the study participants and the study personnel.

3.1 Mean Absolute Relative Difference

As stated above, the MARD indicates how closely CGM results and BG results match. The MARD is commonly used to assess CGM accuracy, and it is calculated

as the average of the absolute values of the individual relative differences between CGM results and BG results (see Eq. 1).

$$MARD = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{CGM_i - BG_i}{BG_i}\right| \times 100\,\% \tag{1}$$

In Eq. 1, BG_i is the result of the ith BG measurement, CGM_i is the corresponding CGM result and N is the total number of pairs of BG results and CGM results.

Despite being commonly used, MARD results from different studies may not necessarily be comparable, since the MARD is influenced by a number of factors [24]: The MARD can be affected by the selection of the study participants, since their specific glucose values can influence MARD, and the study protocol, namely the rates of change (because of the time delay at rapid glucose chances, especially if glucose excursions are induced) and the duration of hypo- and hyperglycemic episodes. Nielsen and colleagues, for example, found that the MARD if the same CGM system was lower in type-2 diabetes subjects than in type-1 diabetes subjects, who are known to show higher glycemic variability [23].

Since these influence factors may make comparisons between different studies difficult, it is worth mentioning that the Clinical and Laboratory Standards Institute (CLSI) has issued a guideline, in which some specifications toward study protocols and data analysis are made [5]. In 2013, data from two clinical trials were published, in which the performance of, in total, five CGM systems was assessed [10, 28, 34]. Three CGM systems were used in the same study, whereas the other two systems were used in another study with very similar study procedures, thus allowing for a comprehensive comparison of all five CGM systems.

As stated above, an SMBG system used for calibration should provide adequate analytical performance especially if SMBG is used for calculating the MARD values. In many studies, devices are calibrated using SMBG values, but the evaluation of MARD is based on values obtained by a laboratory method. This seems to be the superior method, but this may lead to additional error if there is a systematic measurement error between the laboratory method and the SMBG system. In one of the abovementioned comparison studies, a systematic measurement error of 14 % between the SMBG system used for calibration and the laboratory method was found [10]. Since most evaluations in that study were made against SMBG data, this was not a problem, but it may have lead to problems for evaluations in which CGM data were compared to measurement data from the laboratory method.

While the MARD is a parameter that is easy to compute and to interpret, the causes of low performance (i.e., high MARD) cannot be identified in detail. While a low MARD indicates close approximation of the CGM results to the BG results, higher MARD values give no indication about systematic measurement error or its sign (positive or negative error) or random measurement error. Figure 3 shows an example of a CGM trace and the corresponding BG measurements.

CGM systems that are currently available or were available until recently, show MARD results between approximately 10 and 20 % [7, 10, 28, 34].

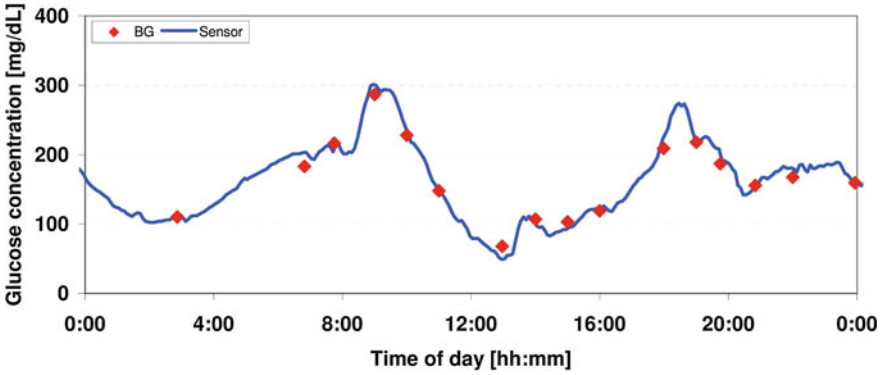

Fig. 3 Example of a CGM trace (*blue line*) and BG measurements (*red diamonds*) used for MARD calculation. Note that at approximately 07:00, the BG result is smaller than the CGM result and at approximately 13:00 the BG result is higher than the CGM result. In the calculation, both incidents contribute towards a higher MARD result, i.e., lower performance, independent from the direction of the deviation

3.2 Precision Absolute Relative Difference

The PARD was introduced in 2009 as percentage ARD by Zisser and colleagues [33]. While this parameter is not as widely used as the MARD, is has regularly been used since its introduction to assess sensor-to-sensor difference of CGM systems.

In order to calculate the PARD, the same subject has to wear at least two sensors of the same CGM system. It is then calculated as the average of the absolute difference between the two sensor signals divided by their mean value (see Eq. 2).

$$PARD = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{CGM_{i,1} - CGM_{i,2}}{(CGM_{i,1} + CGM_{i,2})/2}\right| \times 100\,\% \tag{2}$$

In Eq. 2, $CGM_{i,1}$ is the ith result of the first CGM sensor, $CGM_{i,2}$ is the ith result of the second CGM sensor and N is the total number of CGM result pairs.

The major advantage of the PARD is that it ideally includes all CGM results obtained during the study, whereas the MARD is limited to those CGM results that have corresponding BG results. For CGM systems, which store one result per 1–10 min, this means that if there is no data loss, between 144 and 1440 results per day can be used in the assessment.

As with the MARD, the PARD is easy to compute and to interpret, but again, it lacks in detailed information. While low PARD results indicate that the sensor signals follow each other closely, high PARD results give no indication about the systematic or random measurement errors. Figure 4 shows two CGM traces obtained from sensors of the same CGM system worn by the same subject. CGM systems that are currently available or that were available until recently show PARD results between approximately 7 and 18 % [2, 10, 28, 33, 34].

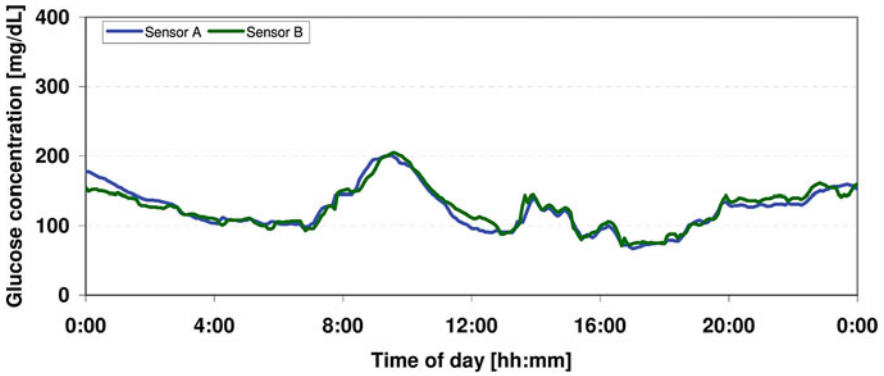

Fig. 4 Traces from two sensors of the same CGM system in the same subject. Note that between 00:00 and approximately 03:00, the Sensor B results are lower than the Sensor A results and between approximately 09:00 and 13:00 the Sensor B result is higher than the Sensor A result. Independent of the direction of this deviation, both incidents contribute to a higher PARD, i.e., lower performance

4 Glucose Prediction Based on Tissue Glucose Concentrations

Glucose prediction requires adequate accuracy of CGM systems. This does not only include the commonly used MARD, but also the PARD. Ideally, CGM systems exhibit both low MARD numbers (i.e., CGM results and BG results match closely) as well as low PARD numbers (i.e., the CGM results from different sensors of the same CGM system follow each other closely). It is quite obvious that adequate agreement between CGM results and BG results and minimal time lag between the tissue and blood glucose concentrations are required for predicting the future BG from CGM results. On the other hand, the differences between CGM results of one sensor and another sensor of the same type should be sufficiently small, in order to allow for reproducible glucose prediction.

Glucose prediction has to take into account not only the accuracy of the CGM system itself, but also the accuracy of the measurement system that is used for calibration and the timing of the calibration.

Current CGM systems seem sufficiently accurate to allow for adequate glucose prediction with currently available prediction models [8, 19, 21, 22, 27].

References

1. American Diabetes Association: self-monitoring of blood glucose. Diabetes Care **17**(1), 81–86 (1994). doi:10.2337/diacare.17.1.81. http://care.diabetesjournals.org/content/17/1/81.short
2. Bailey, T., Zisser, H., Chang, A.: New features and performance of a next-generation SEVEN-day continuous glucose monitoring system with short lag time. Diabetes Technol. Ther. **11**(12), 749–755 (2009)
3. Baumstark, A., Pleus, S., Schmid, C., Link, M., Haug, C., Freckmann, G.: Lot-to-lot variability of test strips and accuracy assessment of systems for self-monitoring of blood glucose according to ISO 15197. J. Diabetes Sci. Technol. **6**(5), 1076–1086 (2012)
4. Brazg, R.L., Klaff, L.J., Parkin, C.G.: Performance variability of seven commonly used self-monitoring of blood glucose systems: clinical considerations for patients and providers. J. Diabetes Sci. Technol. **7**(1), 144–152 (2013)
5. Clinical and Laboratory Standards Institute: performance metrics for continuous interstitial glucose monitoring; approved guideline (2008). http://shopping.netsuite.com/c.1253739/site/Sample_pdf/POCT05A_sample.pdf. Accessed 24 June 14 A.D
6. Consensus statement on self-monitoring of blood glucose. Diabetes Care **10**(1), pp. 95–99 (1987)
7. Damiano, E.R., El-Khatib, F.H., Zheng, H., Nathan, D.M., Russell, S.J.: A comparative effectiveness analysis of three continuous glucose monitors. Diabetes Care **36**(2), 251–259 (2013)
8. Del Favero, S., Bruttomesso, D., Di Palma, F., Lanzola, G., Visentin, R., Filippi, A., Scotton, R., Toffanin, C., Messori, M., Scarpellini, S., Keith-Hynes, P., Kovatchev, B.P., Devries, J.H., Renard, E., Magni, L., Avogaro, A., Cobelli, C.: First use of model predictive control in outpatient wearable artificial pancreas. Diabetes Care **37**(5), 1212–1215 (2014)
9. Freckmann, G., Baumstark, A., Jendrike, N., Zschornack, E., Kocher, S., Tshiananga, J., Heister, F., Haug, C.: System accuracy evaluation of 27 blood glucose monitoring systems according to DIN EN ISO 15197. Diabetes Technol. Ther. **12**(3), 221–231 (2010)
10. Freckmann, G., Pleus, S., Link, M., Zschornack, E., Klotzer, H.M., Haug, C.: Performance evaluation of three continuous glucose monitoring systems: comparison of six sensors per subject in parallel. J. Diabetes Sci. Technol. **7**(4), 842–853 (2013)
11. Freckmann, G., Baumstark, A., Schmid, C., Pleus, S., Link, M., Haug, C.: Evaluation of 12 blood glucose monitoring systems for self-testing: system accuracy and measurement reproducibility. Diabetes Technol. Ther. **16**(2), 113–122 (2014)
12. Hasslacher, C., Kulozik, F., Platten, I.: Accuracy of self monitoring blood glucose systems in a clinical setting: application of new planned ISO- standards. Clin. Lab. **59**(7–8), 727–733 (2013)
13. Hirose, T., Mita, T., Fujitani, Y., Kawamori, R., Watada, H.: Glucose monitoring after fruit peeling: pseudohyperglycemia when neglecting hand washing before fingertip blood sampling: wash your hands with tap water before you check blood glucose level. Diabetes Care **34**(3), 596–597 (2011)
14. Hortensius, J., Slingerland, R.J., Kleefstra, N., Logtenberg, S.J., Groenier, K.H., Houweling, S.T., Bilo, H.J.: Self-monitoring of blood glucose: the use of the first or the second drop of blood. Diabetes Care **34**(3), 556–560 (2011)
15. International Organization for Standardization: in vitro diagnostic test systems—requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus. EN ISO 15197 (2003)
16. International Organization for Standardization: in vitro diagnostic test systems—requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus. ISO 15197 (2013)
17. Kamath, A., Mahalingam, A., Brauker, J.: Analysis of time lags and other sources of error of the DexCom SEVEN continuous glucose monitor. Diabetes Technol. Ther. **11**(11), 689–695 (2009)
18. Koschinsky, T., Heinemann, L.: Sensors for glucose monitoring: technical and clinical aspects. Diabetes Metab. Res. Rev. **17**(2), 113–123 (2001)

19. Kovatchev, B.P., Renard, E., Cobelli, C., Zisser, H.C., Keith-Hynes, P., Anderson, S.M., Brown, S.A., Chernavvsky, D.R., Breton, M.D., Mize, L.B., Farret, A., Place, J., Bruttomesso, D., Del Favero, S., Boscari, F., Galasso, S., Avogaro, A., Magni, L., Di Palma, F., Toffanin, C., Messori, M., Dassau, E., Doyle, F.J.: Safety of outpatient closed-loop control: first randomized crossover trials of a wearable artificial pancreas. Diabetes Care **37**(7), 1789–1796 (2014)
20. Kuo, C.Y., Hsu, C.T., Ho, C.S., Su, T.E., Wu, M.H., Wang, C.J.: Accuracy and precision evaluation of seven self-monitoring blood glucose systems. Diabetes Technol. Ther. **13**(5), 596–600 (2011)
21. Leelarathna, L., Thabit, H., Allen, J.M., Nodale, M., Wilinska, M.E., Powell, K., Lane, S., Evans, M.L., Hovorka, R.: Evaluating the performance of a novel embedded closed-loop system. J. Diabetes Sci. Technol. **8**(2), 267–272 (2014)
22. Luijf, Y.M., DeVries, J.H., Zwinderman, K., Leelarathna, L., Nodale, M., Caldwell, K., Kumareswaran, K., Elleri, D., Allen, J.M., Wilinska, M.E., Evans, M.L., Hovorka, R., Doll, W., Ellmerer, M., Mader, J.K., Renard, E., Place, J., Farret, A., Cobelli, C., Del Favero, S., Dalla Man, C., Avogaro, A., Bruttomesso, D., Filippi, A., Scotton, R., Magni, L., Lanzola, G., Di Palma, F., Soru, P., Toffanin, C., De Nicolao, G., Arnolds, S., Benesch, C., Heinemann, L.: Day and night closed-loop control in adults with type 1 diabetes: a comparison of two closed-loop algorithms driving continuous subcutaneous insulin infusion versus patient self-management. Diabetes Care **36**(12), 3882–3887 (2013)
23. Nielsen, J.K., Freckmann, G., Kapitza, C., Ocvirk, G., Koelker, K.H., Kamecke, U., Gillen, R., Amann-Zalan, I., Jendrike, N., Christiansen, J.S., Koschinsky, T., Heinemann, L.: Glucose monitoring by microdialysis: performance in a multicentre study. Diabet. Med. **26**(7), 714–721 (2009)
24. Obermaier, K., Schmelzeisen-Redeker, G., Schoemaker, M., Klotzer, H.M., Kirchsteiger, H., Eikmeier, H., del Re, L.: Performance evaluations of continuous glucose monitoring systems: precision absolute relative deviation is part of the assessment. J. Diabetes Sci. Technol. **7**(4), 824–832 (2013)
25. Parkes, J.L., Slatin, S.L., Pardo, S., Ginsberg, B.H.: A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. Diabetes Care **23**(8), 1143–1148 (2000). doi:10.2337/diacare.23.8.1143. http://care.diabetesjournals.org/content/23/8/1143.abstract
26. Pfutzner, A., Mitri, M., Musholt, P.B., Sachsenheimer, D., Borchert, M., Yap, A., Forst, T.: Clinical assessment of the accuracy of blood glucose measurement devices. Curr. Med. Res. Opin. **28**(4), 525–531 (2012)
27. Phillip, M., Battelino, T., Atlas, E., Kordonouri, O., Bratina, N., Miller, S., Biester, T., Stefanija, M.A., Muller, I., Nimri, R., Danne, T.: Nocturnal glucose control with an artificial pancreas at a diabetes camp. N. Engl. J. Med. **368**(9), 824–833 (2013)
28. Pleus, S., Schmid, C., Link, M., Zschornack, E., Klotzer, H.M., Haug, C., Freckmann, G.: Performance evaluation of a continuous glucose monitoring system under conditions similar to daily life. J. Diabetes Sci. Technol. **7**(4), 833–841 (2013)
29. Pleus, S., Schmid, C., Link, M., Baumstark, A., Haug, C., Stolberg, E., Freckmann, G.: Accuracy assessment of two novel systems for self-monitoring of blood glucose following ISO 15197:2013. J. Diabetes Sci. Technol. **8**(4), 906–908 (2014). doi:10.1177/1932296814536030. http://dst.sagepub.com/content/8/4/906.short
30. Schmid, C., Haug, C., Heinemann, L., Freckmann, G.: System accuracy of blood glucose monitoring systems: impact of use by patients and ambient conditions. Diabetes Technol. Ther. **15**(10), 889–896 (2013)
31. Thorpe, G.H.: Assessing the quality of publications evaluating the accuracy of blood glucose monitoring systems. Diabetes Technol. Ther. **15**(3), 253–259 (2013)
32. Twomey, P.J.: Plasma glucose measurement with the yellow springs glucose 2300 STAT and the Olympus AU640. J. Clin. Pathol. **57**(7), 752–754 (2004)
33. Zisser, H.C., Bailey, T.S., Schwartz, S., Ratner, R.E., Wise, J.: Accuracy of the SEVEN continuous glucose monitoring system: comparison with frequently sampled venous glucose measurements. J. Diabetes Sci. Technol. **3**(5), 1146–1154 (2009)

34. Zschornack, E., Schmid, C., Pleus, S., Link, M., Klotzer, H.M., Obermaier, K., Schoemaker, M., Strasser, M., Frisch, G., Schmelzeisen-Redeker, G., Haug, C., Freckmann, G.: Evaluation of the performance of a novel system for continuous glucose monitoring. J. Diabetes Sci. Technol. **7**(4), 815–823 (2013)
35. Zueger, T., Diem, P., Mougiakakou, S., Stettler, C.: Influence of time point of calibration on accuracy of continuous glucose monitoring in individuals with type 1 diabetes. Diabetes Technol. Ther. **14**(7), 583–588 (2012)

CGM—How Good Is Good Enough?

Michael Schoemaker and Christopher G. Parkin

Abstract Continuous glucose monitoring (CGM) systems are more informative than the traditional self-monitoring of capillary blood glucose (BG). Although advances in CGM technology have significantly improved the clinical utility of CGM devices compared with earlier versions, it is often difficult to assess the accuracy and precision of current devices due to differences in assessment protocols and reporting of results. Because CGM sensor accuracy can impact both the clinical utility and patient acceptance of CGM use, it is important to consider the performance characteristics seen in the current systems when assessing the clinical value of this technology. Moreover, standardization of the metrics used to assess CGM accuracy and precision are needed to help developers, clinicians, and patients make informed decisions regarding the CGM systems they are considering. In this chapter, we discuss the most commonly used methods for the assessment of CGM system performance, the accuracy and reliability of current CGM systems, and the remaining unsolved technological and physiological hurdles.

1 Background

Continuous glucose monitoring (CGM) technology provides significant advantages to individuals treated with intensive insulin therapy compared with traditional self-monitoring of blood glucose (SMBG). Unlike SMBG, which measures the current glucose level at a single point in time, CGM presents a constant stream of data (measured every 1–15 min), indicating not only the current interstitial glucose level but also trends in glucose direction and velocity of glucose change. Moreover, CGM systems feature alarms that alert users when glucose is or predicted to be above or below programmed glucose thresholds, which makes the technology clinically useful

M. Schoemaker (✉)
Roche Diagnostics GmbH, Mannheim, Germany
e-mail: michael.schoemaker@roche.com

C.G. Parkin
CGParkin Communications, Boulder City, USA

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*,
Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_3

in identifying postprandial hyperglycemia, nocturnal hypoglycemia, and in assisting individuals with hypoglycemia unawareness [43].

Numerous clinical trials have shown that use of CGM improves glycemic control and reduces hypoglycemia in children and adults with type 1 diabetes [6, 15, 20, 27, 38, 41, 48]. However, the clinical benefit of CGM use is directly related to the frequency of use of the technology. In many studies, significant clinical improvements were seen only in those patients who regularly wore their CGM devices 60–70 % of the time [2, 9, 20, 27, 41, 46–48].

As the technology continues to evolve, use of CGM is emerging as a standard of care for diabetes patients managed with intensive insulin therapy [8, 30, 39, 44]. Additionally, CGM is an integral component of the "artificial pancreas" (AP) systems under development, which utilize continuous subcutaneous insulin infusion (CSII) linked with CGM and an automated control algorithm to communicate between the two components to mimic physiologic insulin delivery.

Despite its demonstrated benefits, adoption and routine utilization of CGM has been relatively slow. A large US registry of individuals with type 1 diabetes showed that only 6 % of individuals used CGM as a regular part of their diabetes management [7]. It is also known that many patients who are started on CGM discontinue it shortly thereafter [42].

Although cost is often identified as a hindrance to CGM use; [4, 26] inadequate sensor accuracy may be a more influential obstacle to widespread adoption of CGM for current use in clinical diabetes management and future use in AP systems [13, 28]. In clinical trials, the number of patients who discontinued the studies due to accuracy or sensor-related issues was significant [20, 26, 41]. Therefore, it is important that both clinicians and patients are able to easily assess the performance of the CGM systems they are considering.

2 CGM Performance Assessment

2.1 Sensor Signal

The effectiveness of CGM devices in measuring current glucose and predicting future glucose levels is dependent upon the accuracy and reliability of the CGM signal. The raw (unfiltered) signal of a CGM sensor is the basis for meaningful data output. A raw signal that exhibits strong noise and artifacts requires robust filter algorithms in order to equalize the signal. The characteristics of noise and artifacts will dictate the type of filter algorithm needed. Regardless of the filter algorithm selected, all filters introduce a time lag in that the filtered signal lags behind the raw signal; the stronger the filter algorithm, the longer the time lag. A raw signal of high quality requires minimal filtering and data processing, resulting in a much cleaner sensor response with little delay between the shown data and the raw signal.

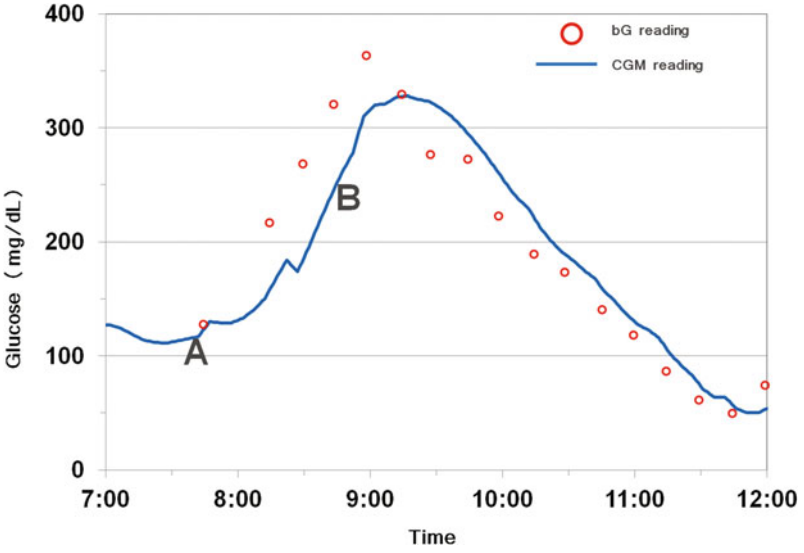

Fig. 1 Example of "time lag" between interstitial and blood glucose levels. *There is virtually no time lag when glucose is stable (A), whereas the lag is apparent when glucose levels are changing rapidly (B).*

2.2 Reference Methodology

Blood glucose measurement is most commonly used as reference data for the analysis of CGM accuracy. However, the limitations of using this reference method must be considered. It is important to consider that this method is comparing glucose concentrations in two body compartments: blood and interstitial fluid. As demonstrated by Basu and colleagues, [5] who was the first to directly measure the transport of glucose from the vascular compartment to subcutaneous tissue, demonstrated that the mean time to appearance of tracer glucose in the abdominal subcutaneous tissue after intravenous bolus is between 5 and 6 min in the resting, overnight fasted state. Moreover, the glucose transport lag time may vary when the patient is not in the fasting state (e.g., glucose levels are rising or declining). An example of this lag is presented in Fig. 1. As a consequence, CGM and blood glucose points that are paired according to their measurement time stamp may be erroneously paired from the physiological point of view. Additionally, the blood glucose readings are paired with a very small fraction of CGM readings, most CGM data are neglected, and deviations of CGM readings from the blood glucose occurring between the blood glucose measurements are not detected at all.

2.3 Accuracy and Precision

A common metric for assessment of CGM accuracy is the aggregate mean absolute relative difference (MARD) between all temporally matched sensor data and reference measurements across all subjects of a study. Reported as a percentage, MARD is the average of the absolute error between all CGM values and matched reference values. A small percentage indicates that the CGM readings are close to the reference glucose value, whereas a larger MARD percentage indicates greater discrepancies between the CGM and reference glucose values.

Another metric is the *median* absolute relative difference, which is also expressed as a percentage. Both the MARD and median absolute relative difference are often reported in CGM accuracy studies [29, 34, 50]; however, some studies [10, 22, 49] report only the median difference, which is misleading because it diminishes the impact of outliers. When there are significant numbers of outliers, accuracy appears to be greater.

Although the MARD is a more stringent metric and is easy to compute and interpret for succinct summarization of CGM properties, certain limitations should be considered when comparing reported findings from different accuracy studies [37]. Because the MARD uses blood glucose readings as a reference, the influence of time lags is introduced and the majority of CGM values are neglected. Additionally, the MARD makes no distinction between positive and negative errors or between systematic and random errors.

It is also important to consider the composition of the study cohort and study setting. Are the study subjects prone to severe and/or frequent fluctuations in glucose? How often and how long are they in the hyperglycemic and hypoglycemic range? All of these factors can significantly influence the calculated MARD. For example, in study subjects with relatively stable glucose, the MARD may show close agreement with blood glucose readings. However, in subjects whose glucose is predominantly in the hypoglycemic range or fluctuating between the hypoglycemic and hyperglycemic states, the MARD would likely be less than desirable. Therefore, it is advisable to conduct separate accuracy evaluations for different glucose concentration ranges (e.g., 40–70 mg/dL) and different rates of glucose change (e.g., stable glucose vs. rapidly changing glucose). Additionally, when comparing the accuracy of two or more different CGM systems, it is advisable to perform head-to-head assessments in which the two systems are running simultaneously in each study patient. This will neutralize the potential impact study cohort or study setting differences.

Although use of the MARD provides easily interpreted information about the accuracy of CGM devices compared with reference blood glucose, its utility is impacted by the limitations previously discussed. However, when used in conjunction with calculation of sensor precision, the limitations are diminished (Table 1).

Table 1 MARD and PARD at various glucose ranges [21]

Metric	Abbott	Medtronic	Dexcom
	Navigator	Guardian	Seven Plus
OVERALL MARD % (40–400 mg/dl)	**12.4 ± 3.6**	**16.4 ± 6.9**	**16.7 ± 3.8**
MARD % (<70 mg/dl)	22.6 ± 7.8	32.2 ± 18.1	38.3 ± 24.3
MARD % (70–180 mg/dl)	11.9 ± 3.5	15.4 ± 6.9	17.0 ± 4.0
MARD % (>180 mg/dl)	11.0 ± 4.9	15.1 ± 8.1	11.7 ± 4.2
OVERALL PARD % (40–400 mg/dl)	**10.1 ± 4.1**	**18.1 ± 6.5**	**15.4 ± 4.2**
PARD % (<70 mg/dl)	10.8 ± 4.8	22.3 ± 17.5	24.3 ± 7.5
PARD % (70–180 mg/dl)	10.0 ± 4.2	17.7 ± 5.7	14.9 ± 4.4
PARD % (>180 mg/dl)	9.9 ± 4.1	18.5 ± 9.0	16.4 ± 5.1

Precision of the absolute relative difference (PARD) is a metric used to compare glucose readings from two identical CGM sensors working simultaneously in the same patient; a lower percentage indicates greater precision (less variance) [3, 37, 50, 51]. It is calculated according to Eq. (1):

$$PARD = 100 \cdot \frac{Gluc_{sens1} - Gluc_{sens2}}{Mean(Gluc_{sens1}, Gluc_{sens2})} \tag{1}$$

Using this metric, all CGM data are considered for performance evaluation. However, there is no correlation to blood glucose concentration. Nevertheless, the comparison of two concurrent CGM sensors provides additional and complementary insights into CGM sensor properties and performance.

CGM performance studies that present both MARD and PARD at various glucose ranges facilitate a more reliable assessment of a given CGM sensors true accuracy and precision characteristics. Data from a study that assessed the accuracy of three different CGM systems that were worn simultaneously (in duplicate) are provided as an example of how use of MARD and PARD allows for more straightforward comparisons of various CGM systems [21].

Although significant differences in accuracy and precision between the devices are readily apparent, the data also highlight the differences when comparing overall accuracy/precision across the full glucose range studied (40–400 mg/dL) and within the hypoglycemic range (40–70 mg/dL). This is important because accuracy and precision in the hypoglycemic range is required for the technology to provide timely warnings of immediate or impending hypoglycemia. Failure to provide reliable warnings can impact user acceptance; fear of hypoglycemia is widely recognized as a major barrier to achieving and sustaining optimal glycemic control [16].

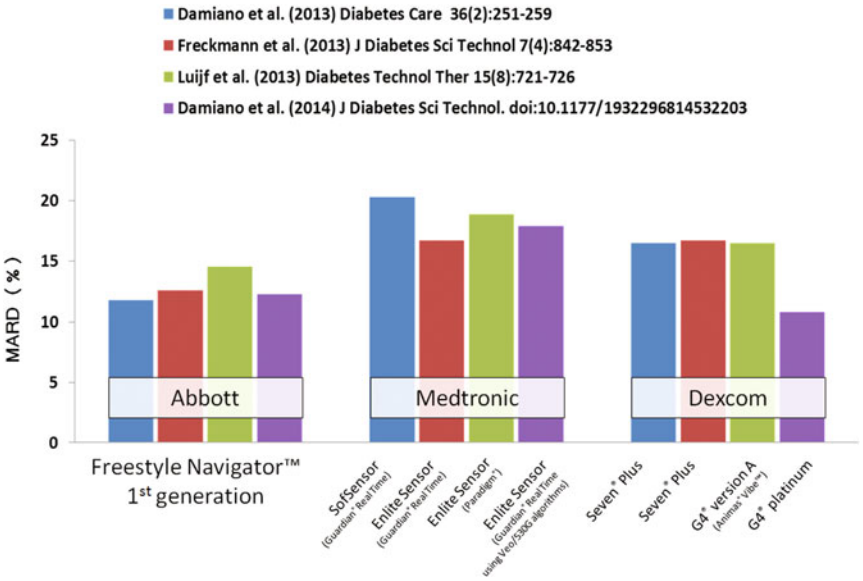

Fig. 2 Overall MARD from four head-to-head studies

3 State of the Art

Despite advances in technology, CGM devices remain less accurate than capillary blood glucose measurements. Some of these deviations are physiological, such as the time lags that occur when glucose levels are rapidly changing. At other times, CGM sensor data may deviate from measured blood glucose values due to inaccurate calibration of the CGM device, sensor dislodgment, and other artifacts such as local decrease of blood flow [25]. If these unpredictable deviations are of sufficient length or magnitude, they could result in erroneous and potentially harmful patient responses (e.g., insulin under-dosing or over-dosing) [32].

It should be noted that the "accuracy gap" between current CGM devices and blood glucose monitoring systems is shrinking. Whereas current blood glucose monitoring systems that fulfill the ISO 1597:2003 criteria exhibit overall MARD values ranging from 4.9 to 6.8 % [45], a recent study of CGM devices showed MARD values as low as 9.2 % [51].

When assessing CGM performance, it is also important to consider differences between the currently available CGM systems. Figure 2 presents comparisons in accuracy (MARD) from four recent studies [17, 18, 21, 33] between the first generation CGM system (FreeStyle Navigator™, Abbott Diabetes, Alameda, USA) and more recent CGM systems from Medtronic (Northridge, USA) and Dexcom (San Diego, USA).

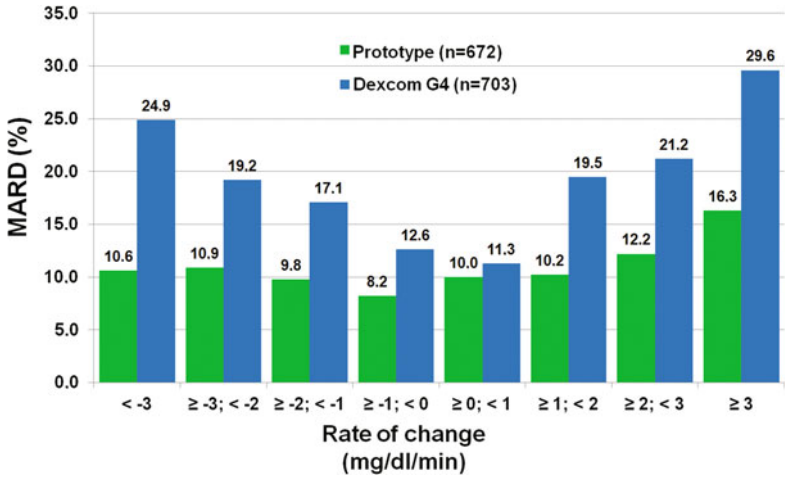

Fig. 3 Stratified MARD for different rates of glucose change

Advances in CGM development have also facilitated significant improvements in accuracy at various rates of glucose change as assessed by MARD. Figure 3 presents data from a head-to-head comparison between a commercially available device (Dexcom G4® PLATINUM Dexcom) and a prototype system (Roche Diagnostics GmbH) [40].

4 Unresolved Issues

In addition to their continuing efforts to improve the accuracy and precision of CGM devices, developers are challenged to address factors that can impact sensor reliability and performance. These factors include motion (including micromotion within the subcutaneous tissue) and pressure, which can affect tissue physiology (interstitial stresses) and local blood flow around the sensor, resulting in signal disruption and/or erroneous glucose readings [25].

4.1 Transient Sensor Signal Disruption

Numerous experiments and clinical studies have revealed occasional, spurious sensor signal dropouts. When this occurs, the glucose values reported by the CGM device are usually lower than the actual values. An example of this is presented in Fig. 4.

Signal disruption most commonly occurs during sleep, due to individuals lying on top of their sensor. These signal dropouts tend to last approximately 15–30 min.

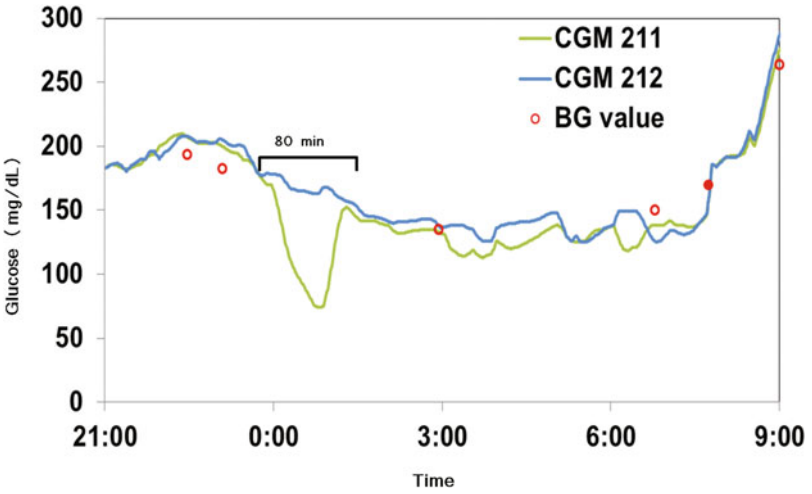

Fig. 4 Example of nocturnal signal drop-out

Although nocturnal signal drops are not likely to prompt an erroneous response in patients using standalone CGM devices, resulting "false alarms" may provoke patients to turn off their CGM and/or discontinue use [19]. However, signal disruptions are of great concern within the context of sensor augmented insulin pump use in which insulin administration can be automatically suspended when CGM measure values fall below the programmed low-glucose threshold and (eventually) in closed-loop insulin delivery system (artificial pancreas) use. Signal disruption also poses a significant challenge in artificial pancreas development efforts.

4.2 Transient Significant CGM Inaccuracies

Whereas signal disruptions are often the result of acute forces such as pressure which can temporarily impact the sensor signal, chronic forces that trigger the foreign body response, and inflammation around the sensor site can impact sensor performance [25]. These forces can include the impact of motion (shear forces) that occur due to normal physiologic activity (e.g., walking, running) [23] and foreign body response [1, 36], which is influenced by the shape [35], size [31] and surface typography [11] of the implanted sensor. Reporting of erroneous interstitial glucose values are unpredictable and may continue over several hours.

5 Next Steps in CGM Development

Increasing standardization of performance metrics will enable CGM technology developers and patients to better assess the accuracy and precision of future CGM devices. The better understanding of the sensor-to-tissue interface will help developers to improve accuracy. As a result, adoption and sustained use of CGM as standalone devices will likely increase. However, the challenges posed by intermittent signal disruptions and transient inaccuracies must be addressed in order for closed-loop insulin delivery development to move forward.

Development of CGM sensors that incorporate intelligent failure detection algorithms, which may include sensors for activity, pressure, temperature, and heart rate, is a long-range goal in CGM development. The addition of glucose pattern recognition capability linked with global position data would enhance the utility of CGM in the future. However, use of multiple sensors presents a potential short-term solution [12]. This approach requires that the two sensors are redundant in that they are not housed in the same probe. Otherwise, tissue compression or other local effects that can impact both sensors and signal disruptions would remain undetected. Use of duplicate sensors can identify occurrence of signal disruptions and erroneous glucose data.

Another proposed solution is to use sensors that are soft and small with microtextured surfaces in order to match tissue modulus [24]. This would both enhance biomechanical biocompatibility and help minimize reduce interfacial stress concentrations [24].

6 Conclusion

Despite several years of expert discussions, neither industry developers nor independent investigators have been able to standardize methodologies for evaluating CGM performance [37]. The one guideline that does exist only covers performance metrics, how studies should be designed and the data analyzed, but no acceptance criteria are provided [14]. Consequently, studies continue to use differing designs and metrics for assessing CGM accuracy, which makes comparisons of currently available CGM devices difficult.

Although use of CGM has been shown to provide significant clinical benefits compared to SMBG [6, 15, 20, 27, 38, 41, 48], sensor performance has slowed widespread adoption of this technology and still is a limiting factor in the development of closed-loop insulin delivery systems.

Because CGM system performance is currently evaluated against blood glucose measurements, these assessments do not reflect the true nature of CGM and may lead to misleading results. Nevertheless, because BG values are still regarded as the "gold standard," accuracy must be assessed in comparison to BG values.

As discussed, the MARD is the most relevant accuracy metric for CGM, not only when computed for the overall glucose range but also separately for different glucose concentration ranges and different rates of change. However, because the MARD does not fully detect sensor performance due to the limited number of paired data points and/or the inherent physiological differences between the two body compartments sampled (i.e., blood glucose vs. interstitial glucose), calculation of the PARD becomes important as a measure for CGM precision and in supporting the accuracy assessment by the MARD.

The question "how good is good enough?" can only be answered as it relates to the intended use. When used as a standalone, adjunctive device for monitoring current and trending glucose levels, with no control over insulin pump infusion, current CGM devices are adequate for their intended use. However, when the CGM data are used for therapy decisions or the CGM device directly controls insulin delivery, accuracy, and precision become increasingly important depending upon the degree of control. For example, when CGM is linked to a low-glucose suspend (LGS) insulin pump, which automatically stops administering basal insulin when current or impending low glucose is detected, the risk of patient harm due to CGM device malfunction is relatively low. Conversely, when CGM is functioning as a component of a closed-loop insulin delivery system, patient risk is significantly increased because detection of an erroneous high-glucose reading would automatically prompt the insulin pump to administer unneeded insulin, resulting in hypoglycemia. This highlights the importance of effectively addressing the critical biomechanical factors (micromotion and pressure) and sensor-to-tissue interface to ensure accuracy, precision, and reliability in clinical use.

In summary, simple, straightforward and clinically significant standards and guidelines are needed to establish common rules for the development of safe and effective products using CGM technology. These rules should include standardized metrics for accuracy and precision, guidance for study designs and data analyses schemata, and acceptance criteria based on intended CGM use.

References

1. Anderson, J.M., Rodriguez, A., Chang, D.T.: Foreign body reaction to biomaterials. Semin. Immunol. **20**(2), 86–100 (2008)
2. Bailey, T.S., Zisser, H.C., Garg, S.K.: Reduction in hemoglobin A1C with real-time continuous glucose monitoring: results from a 12-week observational study. Diabetes Technol. Ther. **9**(3), 203–210 (2007)
3. Bailey, T., Zisser, H., Chang, A.: New features and performance of a next-generation SEVEN-day continuous glucose monitoring system with short lag time. Diabetes Technol. Ther. **11**(12), 749–755 (2009)
4. Bartelme, A., Bridger, P.: The role of reimbursement in the adoption of continuous glucose monitors. J. Diabetes Sci. Technol. **3**(4), 992–995 (2009)
5. Basu, A., Dube, S., Slama, M., Errazuriz, I., Amezcua, J.C., Kudva, Y.C., Peyser, T., Carter, R.E., Cobelli, C., Basu, R.: Time lag of glucose from intravascular to interstitial compartment in humans. Diabetes **62**(12), 4083–4087 (2013)

6. Battelino, T., Phillip, M., Bratina, N., Nimri, R., Oskarsson, P., Bolinder, J.: Effect of continuous glucose monitoring on hypoglycemia in type 1 diabetes. Diabetes Care **34**(4), 795–800 (2011)
7. Beck, R.: T1D exchange overview - 25.000 opportunities, few limitations. American Diabetes Association 72nd Scientific Sessions, Philadelphia (2012)
8. Blevins, T.C., Bode, B.W., Garg, S.K., Grunberger, G., Hirsch, I.B., Jovanovič, L., Nardacci, E., Orzeck, E.A., Roberts, V.L., Tamborlane, W.V., Rothermel, C.: Statement by the American association of clinical endocrinologists consensus panel on continuous glucose monitoring. endocr pract **16**(5), 730–745 (2010)
9. Buckingham, B., Beck, R.W., Tamborlane, W.V., Xing, D., Kollman, C., Fiallo-Scharer, R., Mauras, N., Ruedy, K.J., Tansey, M., Weinzimer, S.A., Wysocki, T.: Continuous glucose monitoring in children with type 1 diabetes. J. Pediatr. **151**(4), 388–393 (2007)
10. Calhoun, P., Lum, J., Beck, R.W., Kollman, C.: Performance comparison of the medtronic sof-sensor and enlite glucose sensors in inpatient studies of individuals with type 1 diabetes. Diabetes Technol. Ther. **15**(9), 758–761 (2013)
11. Campbell, C.E., von Recum, A.F.: Microtopography and soft tissue response. J. Investig. Surg. **2**(1), 51–74 (1989)
12. Castle, J.R., Ward, W.K.: Amperometric glucose sensors: sources of error and potential benefit of redundancy. J. Diabetes Sci. Technol. **4**(1), 221–225 (2010)
13. Chia, C.W., Saudek, C.D.: Glucose sensors: toward closed loop insulin delivery. Endocrinol. Metab. Clin. North Am. **33**(1), 175–195 (2004)
14. Clinical and Laboratory Standards Institute: POCT05A: performance metrics for continuous glucose monitoring – approved guideline (2008)
15. Continuous glucose monitoring and intensive treatment of type 1 diabetes (2008). www.nejm.org
16. Cryer, P.E.: Hypoglycemia: still the limiting factor in the glycemic management of diabetes. Endocr. Pract. **14**(6), 750–756 (2008)
17. Damiano, E.R., El-Khatib, F.H., Zheng, H., Nathan, D.M., Russell, S.J.: A comparative effectiveness analysis of three continuous glucose monitors. Diabetes Care **36**(2), 251–259 (2013)
18. Damiano, E.R., McKeon, K., El-Khatib, F.H., Zheng, H., Nathan, D.M., Russell, S.J.: A comparative effectiveness analysis of three continuous glucose monitors: the Navigator, G4 Platinum, and Enlite. J. Diabetes Sci. Technol. **8**(4), 699–708 (2014)
19. Davey, R.J., Jones, T.W., Fournier, P.A.: Effect of short-term use of a continuous glucose monitoring system with a real-time glucose display and a low glucose alarm on incidence and duration of hypoglycemia in a home setting in type 1 diabetes mellitus. J. Diabetes Sci. Technol. **4**(6), 1457–1464 (2010)
20. Deiss, D., Bolinder, J., Riveline, J.P., Battelino, T., Bosi, E., Tubiana-Rufi, N., Kerr, D., Phillip, M.: Improved glycemic control in poorly controlled patients with type 1 diabetes using real-time continuous glucose monitoring. Diabetes Care **29**(12), 2730–2732 (2006)
21. Freckmann, G., Pleus, S., Link, M., Zschornack, E., Klotzer, H.M., Haug, C.: Performance evaluation of three continuous glucose monitoring systems: comparison of six sensors per subject in parallel. J. Diabetes Sci. Technol. **7**(4), 842–853 (2013)
22. Garg, S.K., Voelmle, M.K., Gottlieb, P.: Feasibility of 10-day use of a continuous glucose-monitoring system in adults with type 1 diabetes. Diabetes Care **32**(3), 436–438 (2009)
23. Geerligs, M., Peters, G.W., Ackermans, P.A., Oomens, C.W., Baaijens, F.P.: Does subcutaneous adipose tissue behave as an (anti-)thixotropic material? J. Biomech. **43**(6), 1153–1159 (2010)
24. Helton, K.L., Ratner, B.D., Wisniewski, N.A.: Biomechanics of the sensor-tissue interface-effects of motion, pressure, and design on sensor performance and foreign body response-part II: examples and application. J. Diabetes Sci. Technol. **5**(3), 647–656 (2011)
25. Helton, K.L., Ratner, B.D., Wisniewski, N.A.: Biomechanics of the sensor-tissue interface-effects of motion, pressure, and design on sensor performance and the foreign body response-part I: theoretical framework. J. Diabetes Sci. Technol. **5**(3), 632–646 (2011)
26. Hermanides, J., Phillip, M., DeVries, J.H.: Current application of continuous glucose monitoring in the treatment of diabetes: pros and cons. Diabetes Care **34**(Suppl 2), 197–201 (2011)

27. Hirsch, I.B., Abelseth, J., Bode, B.W., Fischer, J.S., Kaufman, F.R., Mastrototaro, J., Parkin, C.G., Wolpert, H.A., Buckingham, B.A.: Sensor-augmented insulin pump therapy: results of the first randomized treat-to-target study. Diabetes Technol. Ther. **10**(5), 377–383 (2008)
28. Hovorka, R.: Continuous glucose monitoring and closed-loop systems. Diabet. Med. **23**(1), 1–12 (2006)
29. Keenan, D.B., Mastrototaro, J.J., Zisser, H., Cooper, K.A., Raghavendhar, G., Lee, S.W., Yusi, J., Bailey, T.S., Brazg, R.L., Shah, R.V.: Accuracy of the Enlite 6-day glucose sensor with guardian and Veo calibration algorithms. Diabetes Technol. Ther. **14**(3), 225–231 (2012)
30. Klonoff, D.C., Buckingham, B., Christiansen, J.S., Montori, V.M., Tamborlane, W.V., Vigersky, R.A., Wolpert, H.: Continuous glucose monitoring: an endocrine society clinical practice guideline. J. Clin. Endocrinol. Metab. **96**(10), 2968–2979 (2011)
31. Kvist, P.H., Iburg, T., Aalbaek, B., Gerstenberg, M., Schoier, C., Kaastrup, P., Buch-Rasmussen, T., Hasselager, E., Jensen, H.E.: Biocompatibility of an enzyme-based, electrochemical glucose sensor for short-term implantation in the subcutis. Diabetes Technol. Ther. **8**(5), 546–559 (2006)
32. Leelarathna, L., Nodale, M., Allen, J.M., Elleri, D., Kumareswaran, K., Haidar, A., Caldwell, K., Wilinska, M.E., Acerini, C.L., Evans, M.L., Murphy, H.R., Dunger, D.B., Hovorka, R.: Evaluating the accuracy and large inaccuracy of two continuous glucose monitoring systems. Diabetes Technol. Ther. **15**(2), 143–149 (2013)
33. Luijf, Y., Mader, J., Doll, W., Pieber, T., Farret, A., Place, J., Renard, E., Bruttomesso, D., Filippi, A., Avogaroand, A., Arnolds, S., Benesch, C., Heinemann, L., DeVries, J.: Accuracy and reliability of continuous glucose monitoring systems: a head-to-head comparison. Diabetes Technol. Ther. **15**(8), 721–726 (2013)
34. Mastrototaro, J., Shin, J., Marcus, A., Sulur, G., Abelseth, J., Bode, B.W., Buckingham, B.A., Hirsch, I.B., Kaufman, F.R., Schwartz, S.L., Wolpert, H.A.: The accuracy and efficacy of real-time continuous glucose monitoring sensor in patients with type 1 diabetes. Diabetes Technol. Ther. **10**(5), 385–390 (2008)
35. Matlaga, B.F., Yasenchak, L.P., Salthouse, T.N.: Tissue response to implanted polymers: the significance of sample shape. J. Biomed. Mater. Res. **10**(3), 391–397 (1976)
36. Morhenn, V.B., Lemperle, G., Gallo, R.L.: Phagocytosis of different particulate dermal filler substances by human macrophages and skin cells. Dermatol. Surg. **28**(6), 484–490 (2002)
37. Obermaier, K., Schmelzeisen-Redeker, G., Schoemaker, M., Klotzer, H.M., Kirchsteiger, H., Eikmeier, H., del Re, L.: Performance evaluations of contiuous glucose monitoring systems: precision absolute relative deviation is part of the assessment. J. Diabetes Sci. Technol. **7**(4), 824–832 (2013)
38. O'Connell, M.A., Donath, S., O'Neal, D.N., Colman, P.G., Ambler, G.R., Jones, T.W., Davis, E.A., Cameron, F.J.: Glycaemic impact of patient-led use of sensor-guided pump therapy in type 1 diabetes: a randomised controlled trial. Diabetologia **52**(7), 1250–1257 (2009)
39. Phillip, M., Battelino, T., Rodriguez, H., Danne, T., Kaufman, F.: Use of insulin pump therapy in the pediatric age-group: consensus statement from the European Society for Paediatric Endocrinology, the Lawson Wilkins Pediatric Endocrine Society, and the International Society for Pediatric and Adolescent Diabetes, endorsed by the American Diabetes Association and the European Association for the Study of Diabetes. Diabetes Care **30**(6), 1653–1662 (2007)
40. Pleus, S., Haug, C., Link, M., Zschomack, E., Freckmann, G., Obermaier, K., Schoemaker, M.: Rate-of-change dependence of the performance of two CGM systems during induced glucose excursions. Diabetes **63**(Suppl1), A216 (2014)
41. Raccah, D., Sulmont, V., Reznik, Y., Guerci, B., Renard, E., Hanaire, H., Jeandidier, N., Nicolino, M.: Incremental value of continuous glucose monitoring when starting pump therapy in patients with poorly controlled type 1 diabetes: the RealTrend study. Diabetes Care **32**(12), 2245–2250 (2009)
42. Ramchandani, N., Arya, S., Ten, S., Bhandari, S.: Real-life utilization of real-time continuous glucose monitoring: the complete picture. J. Diabetes Sci. Technol. **5**(4), 860–870 (2011)
43. Ramchandani, N., Heptulla, R.A.: New technologies for diabetes: a review of the present and the future. Int. J. Pediatr. Endocrinol. **2012**(1), 28 (2012)
44. Standards of medical care in diabetes–2012. Diabetes Care **35**(1), 11–63 (2012)

45. Tack, C., Pohlmeier, H., Behnke, T., Schmid, V., Grenningloh, M., Forst, T., Pfutzner, A.: Accuracy evaluation of five blood glucose monitoring systems obtained from the pharmacy: a European multicenter study with 453 subjects. Diabetes Technol. Ther. **14**(4), 330–337 (2012)
46. Tamborlane, W.V., Beck, R.W., Bode, B.W., Buckingham, B., Chase, H.P., Clemons, R., Fiallo-Scharer, R., Fox, L.A., Gilliam, L.K., Hirsch, I.B., Huang, E.S., Kollman, C., Kowalski, A.J., Laffel, L., Lawrence, J.M., Lee, J., Mauras, N., O'Grady, M., Ruedy, K.J., Tansey, M., Tsalikian, E., Weinzimer, S., Wilson, D.M., Wolpert, H., Wysocki, T., Xing, D., Chase, H.P., Fiallo-Scharer, R., Messer, L., Gage, V., Burdick, P., Laffel, L., Milaszewski, K., Pratt, K., Bismuth, E., Keady, J., Lawlor, M., Buckingham, B., Wilson, D.M., Block, J., Benassi, K., Tsalikian, E., Tansey, M., Kucera, D., Coffey, J., Cabbage, J., Wolpert, H., Shetty, G., Atakov-Castillo, A., Giusti, J., O'Donnell, S., Ghiloni, S., Hirsch, I.B., Gilliam, L.K., Fitzpatrick, K., Khakpour, D., Wysocki, T., Fox, L.A., Mauras, N., Englert, K., Permuy, J., Bode, B.W., O'Neil, K., Tolbert, L., Lawrence, J.M., Clemons, R., Maeva, M., Sattler, B., Weinzimer, S., Tamborlane, W.V., Ives, B., Bosson-Heenan, J., Beck, R.W., Ruedy, K.J., Kollman, C., Xing, D., Jackson, J., Steffes, M., Bucksa, J.M., Nowicki, M.L., Van Hale, C., Makky, V., O'Grady, M., Huang, E., Basu, A., Meltzer, D.O., Zhao, L., Lee, J., Kowalski, A.J., Laffel, L., Tamborlane, W.V., Beck, R.W., Kowalski, A.J., Ruedy, K.J., Weinstock, R.S., Anderson, B.J., Kruger, D., LaVange, L., Rodriguez, H.: Continuous glucose monitoring and intensive treatment of type 1 diabetes. N. Engl. J. Med. **359**(14), 1464–1476 (2008)
47. Weinzimer, S., Xing, D., Tansey, M., Fiallo-Scharer, R., Mauras, N., Wysocki, T., Beck, R., Tamborlane, W., Ruedy, K., Chase, H.P., Fiallo-Scharer, R., Messer, L., Tallant, B., Gage, V., Tsalikian, E., Tansey, M.J., Larson, L.F., Coffey, J., Cabbage, J., Wysocki, T., Mauras, N., Fox, L.A., Bird, K., Englert, K., Buckingham, B.A., Wilson, D.M., Clinton, P., Caswell, K., Weinzimer, S.A., Tamborlane, W.V., Doyle, E.A., Mokotoff, H., Steffen, A., Ives, B., Beck, R.W., Ruedy, K.J., Kollman, C., Xing, D., Stockdale, C.R., Jackson, J., Steffes, M.W., Bucksa, J.M., Nowicki, M.L., Van Hale, C.A., Makky, V., Grave, G.D., Horlick, M., Teff, K., Winer, K.K., Becker, D.M., Cleary, P., Ryan, C.M., White, N.H., White, P.C.: Prolonged use of continuous glucose monitors in children with type 1 diabetes on continuous subcutaneous insulin infusion or intensive multiple-daily injection therapy. Pediatr. Diabetes **10**(2), 91–96 (2009)
48. Weinzimer, S., Miller, K., Beck, R., Xing, D., Fiallo-Scharer, R., Gilliam, L.K., Kollman, C., Laffel, L., Mauras, N., Ruedy, K., Tamborlane, W., Tsalikian, E.: Effectiveness of continuous glucose monitoring in a clinical care environment: evidence from the Juvenile Diabetes Research Foundation continuous glucose monitoring (JDRF-CGM) trial. Diabetes Care **33**(1), 17–22 (2010)
49. Wilson, D.M., Beck, R.W., Tamborlane, W.V., Dontchev, M.J., Kollman, C., Chase, P., Fox, L.A., Ruedy, K.J., Tsalikian, E., Weinzimer, S.A.: The accuracy of the FreeStyle Navigator continuous glucose monitoring system in children with type 1 diabetes. Diabetes Care **30**(1), 59–64 (2007)
50. Zisser, H.C., Bailey, T.S., Schwartz, S., Ratner, R.E., Wise, J.: Accuracy of the SEVEN continuous glucose monitoring system: comparison with frequently sampled venous glucose measurements. J. Diabetes Sci. Technol. **3**(5), 1146–1154 (2009)
51. Zschornack, E., Schmid, C., Pleus, S., Link, M., Klotzer, H.M., Obermaier, K., Schoemaker, M., Strasser, M., Frisch, G., Schmelzeisen-Redeker, G., Haug, C., Freckmann, G.: Evaluation of the performance of a novel system for continuous glucose monitoring. J. Diabetes Sci. Technol. **7**(4), 815–823 (2013)

Can We Use Measurements to Classify Patients Suffering from Type 1 Diabetes into Subcategories and Does It Make Sense?

Florian Reiterer, Harald Kirchsteiger, Guido Freckmann and Luigi del Re

Abstract We propose two ideas of how recorded signals from continuous glucose monitoring systems could be used to derive information about the patient-specific characteristics of the glucose dynamics for individuals suffering from type 1 diabetes and how these characteristics of the glucose dynamics could be linked to basic patient data (sex, age,…). Ultimately, these relationships could be used in the future in order to classify patients based on these basic patient data. In the first approach a simple transfer function model was used to fit recorded signals from continuous glucose monitoring systems. Using this approach on data from a recent clinical study, a statistically significant relationship between the model parameters and sex, body mass index, weight and age of the corresponding patients could be identified. The observed relationships could be verified with findings in the clinical studies that were documented in the previous publications. In the second approach a moving average filter with a varying filter width was applied on the data and the variance between filtered and unfiltered signal as a function of the filter width was analysed. From the analysed data a relationship between the low blood glucose index and the high frequency content of signals from continuous glucose monitoring systems seems likely.

1 Introduction

In the healthy human body, meal ingestion is followed by insulin secretion of the β cells of the pancreas. The main source of energy directly available for energy producing chemical processes inside cells is glucose. The carbohydrate portion of

F. Reiterer (✉) · H. Kirchsteiger · L. del Re
Institute for Design and Control of Mechatronical Systems,
Johannes Kepler University Linz, 4040 Linz, Austria
e-mail: florian.reiterer@jku.at
URL: http://desreg.jku.at/

G. Freckmann
Institut für Diabetes-Technologie, Forschungs- und Entwicklungsgesellschaft mbH,
University of Ulm, Ulm, Germany

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*,
Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_4

meals is almost entirely transformed to glucose, thereby raising the blood glucose (BG) concentration. It is insulin that promotes cell uptake of blood glucose, and hereby lowers the BG concentration. Once the concentration is close to a basal state, insulin secretion is normalized and the remaining insulin is quickly cleared from the circulation. As a consequence, BG stays within a narrow range of ≈80 to 140 mg/dl throughout the day [9].

When there is no insulin available—as in type 1 diabetes subjects—much higher BG than usual (called *hyperglycemia*) result and the associated adverse long-term effects were studied thoroughly [19]. It was indeed shown in [19] that intensive insulin therapy through several daily insulin injections helps to prevent and slow down the progress of diabetes-related complications. However, intensive insulin treatment implies the risk of insulin overdose which results in dangerously low BG (called *hypoglycemia*) and immediately impairs functionality of glucose-dependent tissues such as the brain.

The safety issues connected with the BG fluctuations following meal intakes and insulin injections and the difficult question for the proper dosing of insulin for patients suffering from type 1 diabetes mellitus (T1DM) result in a high demand of an accurate description of the dynamical behaviour of BG for subjects suffering from T1DM, including the description of the rise in BG after meals and the pharmacokinetics of insulin after injection in the human body. In recent years ongoing improvements in the reliability and accuracy of continuous glucose monitoring systems (CGMS) have led to a widespread effort in the development of models for the glucose dynamics in patients suffering from T1DM. Models described in the relevant literature have been derived from both, first principles (physiological models, see e.g. [4]) as well as from measurements (data-based models, see e.g. [2] and [15]). For reproducing and describing the patient-specific evolutions of the BG level, data-based models and methods from system identification have proven especially powerful. However, the large intra- and inter-patient variability as well as the complexity of the problem pose serious limitations to finding a universal model in order to describe the BG dynamics of all patients and at all times. Therefore, the effort is directed to finding models derived from patient-specific CGMS recordings that are able to describe a limited section of the BG dynamics (e.g. the variations in the BG level following the breakfast intake, see e.g. [2] and [12]) in an accurate manner. These models, however, only have a limited applicability and acquiring the necessary data for the system identification is expensive. To account for the full range of inter- and intra-patient variability, basically, a separate model would have to be derived for each person and each time frame under consideration, with a large scattering of model parameters. This complicates the effort of finding "the perfect strategy" to keep the BG level of subjects with T1DM within the target range.

A kind of intermediary solution between the task of finding a single model that fits all patient data at all times (which seems almost impossible seeming the complexity of this task) and the limited applicability and high expense of derivation of specific models based on the CGMS recordings of a single individual for a selected time frame would be to derive a limited set of models for a selected class of patients. This approach would not only lower the number of models down to a manageable size, but

would also incentive to rethink the current effort of finding one perfect approach that is suitable for managing the BG level of all T1DM patients. Instead, one could think of using a dedicated strategy for each of the classes of patients. Ideally, it should also be possible to assign patients to one of the classes not using extensive CGMS recordings, but by easily available data like sex, age and body mass index (BMI) which would again result in a considerable drop in the expenses of describing the patient-specific BG variations by means of modelling.

A literature research regarding the classification of patients with T1DM into subcategories only yields a very limited amount of matches. Most of these publications were written with the intent to classify the patients w.r.t. one single aspect of their disease. In [18], for example the aim is to classify patients regarding their risk for severe hypoglycemia (SH), whereas [3] focusses on the classification of patients w.r.t. their awareness of autonomic and neuroglycopenic symptoms associated with severe hypoglycemia. Other publications aim at classifying patients regarding the effects of T1DM. In [17], for example patients are classified according to their stage of maculopathy, whereas [16] aims at classifying patients regarding the presence of diabetic retinopathy and diabetic nephropathy. All of the aforementioned publications investigate the presence or absence of statistically significant correlations between basic patient data (sex, age,…) and the belonging of the patients to a certain class in the corresponding studies. Another approach was found in [6]. In this publication physiological models are fitted to recorded CGMS data, using a separate model for each patient and each time frame of 300 min. The resulting models are then classified in different categories according to the qualitative appearance of the BG dynamics in each time frame of the study. To the best knowledge of the authors there have not been any studies that aim at classifying patients according to their overall BG dynamics with the purpose of building dedicated models of the BG dynamics for each class of patients.

The purpose of this chapter is to perform a feasibility study for the modelling of specific patient classes. Therefore, the following questions are investigated:

- Is it possible to classify patients according to certain characteristics of their BG dynamics?
- If so, are there statistically significant correlations between basic patient data (sex, BMI,…) and the observed classes of patients?
- Is it possible to assign a patient to a specific class based only on basic data?
- Knowing about a patient belonging to a certain class, is it possible to infer useful information about the BG dynamics of this patient?
- Is the knowledge about a patient belonging to a specific class may be even sufficient to approximate the entire BG dynamics of this patient?

In the end these questions boil down to the basic question that is asked in the title: Can we use measurements to classify patients suffering from type 1 diabetes into subcategories and does it make sense? As a first step towards answering these questions a preliminary investigation was performed using CGMS data from a recent clinical study [8]. Two approaches have been chosen in order to evaluate the CGMS data and to classify the patients according to these evaluations. In the first approach

a simple transfer function model has been used in order to model the BG dynamics of the patients suffering from T1DM and the observed model parameters have been further analysed in order to classify the patients. It has been checked whether or not there are statistically significant correlations between the different categories and basic patient data and whether or not these findings correspond to what would be expected from a physiological point of view. In the second approach the high frequency content of the CGMS data has been used in order to classify patients. For this purpose the CGMS data have been filtered using a moving average filter with a varying filter width and the variance between filtered and unfiltered signal as a function of filter width has been used to categorize the patients. Again, the different classes of patients have been checked for statistically significant correlations with basic patient data.

It should already be pointed out clearly at this point that the current paper does not provide clear "yes" or "no" answers to any of the questions from the list. Instead, the intention of this study is to check the feasibility and benefit of putting patients in different categories and to investigate statistical relationships between classes of patients and basic patient data in order to identify easily available categories of data that could later on be used for an a priori classification of patients. Furthermore, it was tried to verify these statistical findings with well-known physiological relationships that have been reported in the literature.

2 Database of CGMS Recordings

For the current work CGMS recordings from a recent clinical study performed at the Institute of Diabetes Technology, Germany, have been used [8]. During these studies 12 subjects suffering from T1DM spent 7 days hospitalized. During this time each of them wore six CMGS in parallel: two FreeStyle Navigator™by Abbott Diabetes Care, Alameda, CA (Navigator), two Guardian® REAL-Time by Medtronic MiniMed, Northridge, CA (Guardian) and two DexCom™Seven® Plus by DexCom, San Diego, CA (Dexcom). The recording with all CGMS started in the morning of the first day of the clinical trial. Each individual wore all CGMS sensors until the specified time of removal as recommended by the corresponding manufacturer. The individuals therefore ended up wearing the Navigator for five clinical days, the Guardian for six clinical days and the Dexcom for seven clinical days, respectively. The core period of the study is defined as from 7:00 of the second day until 7:00 of the sixth day. During this period all CGMS should already be well calibrated and adjusted to the correct measurement of the transient interstitial glucose (IG) values. During the entire period of the study all CGMS signals have been recorded, together with regular BG measurements by means of self-monitoring of blood glucose (SMBG) and documentation about meal intakes (together with the corresponding carbohydrate content of the meals), insulin bolus injections and basal insulin rates. All individuals but one (who performed one daily injection of long-term insulin) used insulin pumps for the supply of basal insulin.

On days 2 and 5 of the trial glucose excursions have been induced by serving fast-absorbing meals with a high carbohydrate content and by delaying the insulin injections by 15 min leading to considerable temporary peaks in the BG values. Furthermore, the injected insulin quantities have been increased by 15 % compared to the recommended values.

All analyses described in this document focus on the core period of the study. Furthermore, it is assumed for simplicity that the CGMS are perfectly calibrated and read the transient values of the BG without any additional time delays. These, however, are simplifications that have been made for the current work. Seen that the character of this work is only preliminary in nature and seen that the error associated with these simplifications is rather small, it seems justified. For future, more detailed studies, a post-processing of the CGMS signals together with the BG measurements as described in [5] could be performed in order to account for the loss of calibration of the CGMS as well as for delay times between BG and IG. For consecutive studies on the topic the use of such a post-processing of CGMS data is intended to be used.

3 Modelling Using a Simple Transfer Function Model

In the following section we use a rather simple transfer function model to fit the recorded CGMS data and to classify the patients according to the resulting model parameters. Furthermore, the relationship between basic patient data and the different model parameters is investigated.

3.1 Description of the Model and System Identification

In the first approach a simple transfer function model was identified in order to fit the CGMS data of the patients. The resulting parameters of the model are then used to cluster the patients into different groups.

For each patient and each day the CGMS signal was chosen for the modelling that best fitted the recorded BG values. This was done by choosing for each day the CGMS with the minimal MARD as calculated with the well-known formula:

$$MARD = \frac{100}{N} \sum_{i=1}^{N} \frac{|y_{CGMS,i} - y_{SMBG,i}|}{y_{SMBG,i}} \tag{1}$$

For the modelling the entire core period of the study was splitted into time frames from just before the start of the breakfast until 10:00 pm of the same day. This means that the effect of all three major meals of the day (and of the corresponding

insulin bolus injections) as well as intermediary snacks are captured by the CGMS recordings. The fluctuations of the BG values during the night are not considered in the modelling.

A third-order transfer function including an integrator term is chosen to model the response of the BG to carbohydrate intakes as well as to bolus insulin injections:

$$BG(s) = \frac{K_1}{(1+sT_1)^2 s} Carbs(s) + \frac{K_2}{(1+sT_2)^2 s} Insulin(s) \tag{2}$$

This model has already been used extensively (and successfully) in previous studies, see e.g. [12, 13]. The chosen model structure has the advantage that the model parameters have a direct physiological meaning. The constants K_1 and K_2 can be interpreted as the sensitivity of the BG level to carbohydrate intakes and insulin injections, whereas the time constants T_1 and T_2 determine the time for meal intakes/insulin injections to manifest themselves in a changed BG value. For each patient and each day the parameters K_1, T_1, K_2 and T_2 have been chosen to be constant and have been adjusted in a way to achieve an optimal alignment between recorded CGMS data and the BG values as predicted by the model. Choosing constant values for insulin and glucose sensitivity for the entire day is basically incorrect because these sensitivities are not constant factors, but change significantly depending on the time of the day and on other influencing factors (physical activity, insulin injections site, etc.—see e.g. [10]). For example, the measured insulin sensitivity in the morning is lower than in the evening. Assuming constant values for the entire day period means that the identified parameters correspond to averages over the entire day period.

For the identification of the model parameters it was taken advantage of the fact that by approximating the insulin bolus injections and the carbohydrate intakes in the model as impulses there exists an analytical solution for the system response in the time domain. Assuming the system to be in steady-state at the time before the breakfast, the solution is given by:

$$\begin{aligned} BG(t) = BG_0 &+ K_1 \sum_{i=1}^{K} A_i \left(1 - \exp\left(-\frac{t-t_i}{T_1}\right)\left(1 + \frac{t-t_i}{T_1}\right)\right) \mathscr{H}(t-t_i) \\ &+ K_2 \sum_{j=1}^{L} B_j \left(1 - \exp\left(-\frac{t-t_j}{T_2}\right)\left(1 + \frac{t-t_j}{T_2}\right)\right) \mathscr{H}(t-t_j) \end{aligned} \tag{3}$$

In this formula BG_0 corresponds to the initial BG concentration (before the intake of the breakfast), t is the time (in min, with $t = 0$ min at the start of the time frame of the identification for each day), A_i are the carbohydrate contents of each meal (in g), t_i corresponds to the time of the different meals, K is the total number of meal intakes in the studied time frame, B_j correspond to the insulin amounts of each bolus injection (in IU), t_j corresponds to the time of the corresponding bolus injection, L is the total number of insulin bolus injections in the studied time frame and $\mathscr{H}$ is the Heaviside function.

Using the analytical solution for the BG concentration as a function of time 3, the constants K_1, T_1, K_2 and T_2 were identified using the standard optimization routines LSQCURVEFIT and NLMEFIT in MATLAB. The algorithm LSQCURVEFIT performs a direct minimization of the quadratic error, whereas NLMEFIT fits the model by maximizing an approximation to the marginal likelihood. For each patient either the results obtained using LSQCURVEFIT or NLMEFIT have been used, depending on which routine resulted in the minimum value for the average error between CGMS signal and fitted results. For the evaluation of the quality of the fitted curves the VAF and FIT values have been used. These are defined according to the well-known formulae:

$$VAF = 100 \left(1 - \frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \right) \tag{4}$$

$$FIT = 100 \left(1 - \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2}} \right) \tag{5}$$

In these formulae the measured values from the CGMS are denoted with y_i, whereas the model outputs are called $\hat{y}_i$. The average value of the CGMS signal over the entire time frame is named $\bar{y}$. Both, the VAF and the FIT result in a value of 100 % for the case of a perfect fit (all estimated values are exactly the same as the measured values) and a value of 0 % if the error of the estimated signal is just the same as for a horizontal line at the mean value of the measured signal.

Based on the VAF and FIT of the fitted models, it was distinguished between successful and failed fits. For the subsequent analyses only fits with a sufficient quality have been chosen. For this purpose a threshold of VAF = 80 % (corresponds to FIT = 55.28 %) was defined as a minimum quality of successful fits. This approach is consistent to [6]. An example for a successful fit can be seen in Fig. 1. Suitable fits could be found in 67 % of the cases (overall value for all patients and all days). The detailed results for all patients and all days of the core period of the clinical study can be found in Table 1. This table contains all FIT values obtained using MATLAB's LSQCURVEFIT and NLMEFIT algorithms (except for patient B2_12 for whom no results could be obtained using NLMEFIT). Identified curves with a FIT value higher than the threshold of 55 % are marked with a checkmark. It can be seen that for some patients the NLMEFIT algorithm leads to higher FIT values, whereas for others better results could be obtained using LSQCURVEFIT. For the further evaluations for each patient the results from the algorithm that yielded higher FIT values have been used. In case of very comparable FIT values the results obtained from NLMEFIT were chosen. Therefore, the identified parameters with LSQCURVEFIT have been

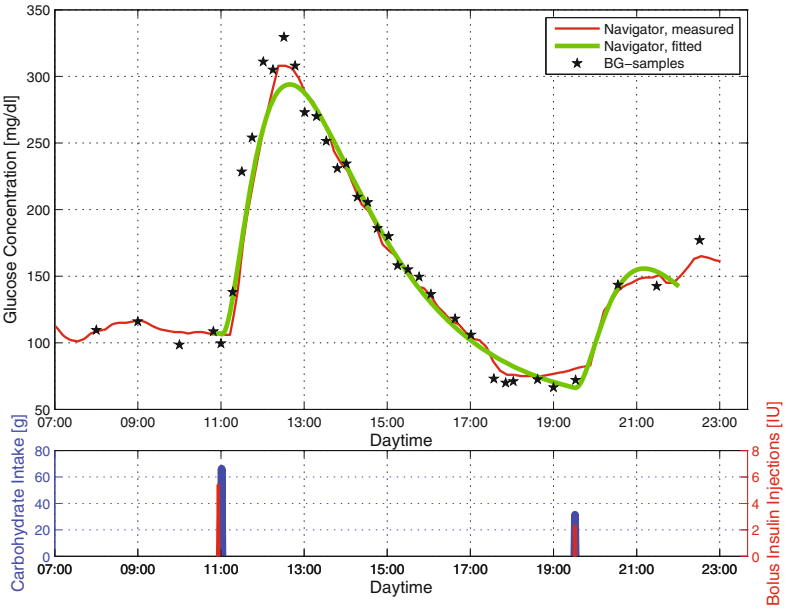

Fig. 1 Example: Measured CGMS signal and identified model for patient B2_06 at day 5

used for patients B1_01, B2_06 and B3_12, whereas for all other patients the results obtained with NLMEFIT have been used for all other patients.

3.2 Trends and Correlations

The subsequent clustering of the identified parameters has not been performed directly, but first a reduction of the dimensionality was performed from four parameters down to two parameters. Instead of searching for clusters is the four-dimensional space K_1, T_1, K_2 and T_2, the search was performed in the two-dimensional domain K_1/K_2 and T_2/T_1. Due to the fact that carbohydrate intakes and insulin injections almost always occur simultaneously, it is difficult to identify unique values for K_1 or K_2. However, what typically can be identified directly is the amount of insulin that is needed to counterbalance the intake of 1 g of carbohydrates. It is exactly this number that is expressed by K_1/K_2. The number therefore corresponds to a sort of inverse insulin sensitivity where big numbers of K_1/K_2 correspond to a low insulin sensitivity which means that a lot of insulin is necessary to counteract the blood glucose increase following a carbohydrate intake. The second number, T_2/T_1, expresses the time that the insulin needs to cause an effect on the BG level, compared to the time that the carbohydrates take. A large number of T_2/T_1 therefore corresponds to a slow insulin reaction compared to the time for the BG rise following a carbohydrate

Table 1 Results of the system identification: FIT values [%] for all patients and all days of the core period of the clinical study obtained using MATLAB's LSQCURVEFIT and NLMEFIT algorithms

Patient	Algorithm	Day2	Day3	Day4	Day5
B1_01	LSQCURVEFIT	63.39 ✓	64.23 ✓	76.17 ✓	66.89 ✓
	NLMEFIT	61.26 ✓	58.53 ✓	76.06 ✓	60.37 ✓
B1_02	LSQCURVEFIT	40.55 x	0.73 x	7.96 x	79.65 ✓
	NLMEFIT	40.88 x	19.36 x	14.32 x	79.65 ✓
B1_03	LSQCURVEFIT	73.88 ✓	−28.70 x	69.81 ✓	62.09 ✓
	NLMEFIT	76.89 ✓	55.89 ✓	75.69 ✓	62.09 ✓
B1_04	LSQCURVEFIT	23.58 x	18.71 x	46.87 x	81.58 ✓
	NLMEFIT	23.58 x	18.70 x	46.94 x	81.58 ✓
B2_05	LSQCURVEFIT	71.24 ✓	46.10 x	67.21 ✓	42.45 x
	NLMEFIT	71.23 ✓	46.09 x	67.20 ✓	45.04 x
B2_06	LSQCURVEFIT	83.80 ✓	66.58 ✓	51.29 x	90.77 ✓
	NLMEFIT	83.30 ✓	66.58 ✓	50.32 x	90.77 ✓
B2_07	LSQCURVEFIT	38.51 x	48.25 x	61.18 ✓	67.13 ✓
	NLMEFIT	38.51 x	62.47 ✓	61.18 ✓	74.89 ✓
B2_08	LSQCURVEFIT	57.86 ✓	52.49 x	36.58 x	43.78 x
	NLMEFIT	57.85 ✓	64.20 ✓	40.16 x	43.78 x
B3_09	LSQCURVEFIT	65.28 ✓	76.51 ✓	42.47 x	56.93 ✓
	NLMEFIT	65.27 ✓	76.50 ✓	51.99 x	67.18 ✓
B3_10	LSQCURVEFIT	66.51 ✓	30.97 x	68.59 ✓	56.86 ✓
	NLMEFIT	66.51 ✓	59.11 ✓	72.70 ✓	78.02 ✓
B3_11	LSQCURVEFIT	63.02 ✓	51.86 x	68.16 ✓	74.66 ✓
	NLMEFIT	63.02 ✓	52.84 x	68.16 ✓	74.66 ✓
B3_12	LSQCURVEFIT	66.98 ✓	4.82 x	4.45 x	65.81 ✓
	NLMEFIT	–	–	–	–

intake. Typically, the time for the insulin reaction is longer than the time for the carbohydrate reaction ($T_2 > T_1$) which means that for the case of a simultaneous carbohydrate intake and insulin injection there will be some temporal peak in the BG level (see e.g. Fig. 1). The magnitude of these peaks depends on both, the insulin sensitivity and the times for insulin and carbohydrate effects. Maximum peaks will therefore occur for big values of K_1/K_2 and T_2/T_1. So, the reduction of dimensionality from four to two not only helps to overcome problems with the uniqueness of the identified parameters and reduce the scattering of these parameters, but also ensures a simple physiological meaning of the two new parameters. However, it should be kept in mind that the two parameters K_1/K_2 and T_2/T_1 are not sufficient to reproduce the BG profiles as a function of time.

The results for the identified values of K_1/K_2 and T_2/T_1 for all patients and all days (only the successful fits) can be seen in Fig. 2. In this figure one symbol in a

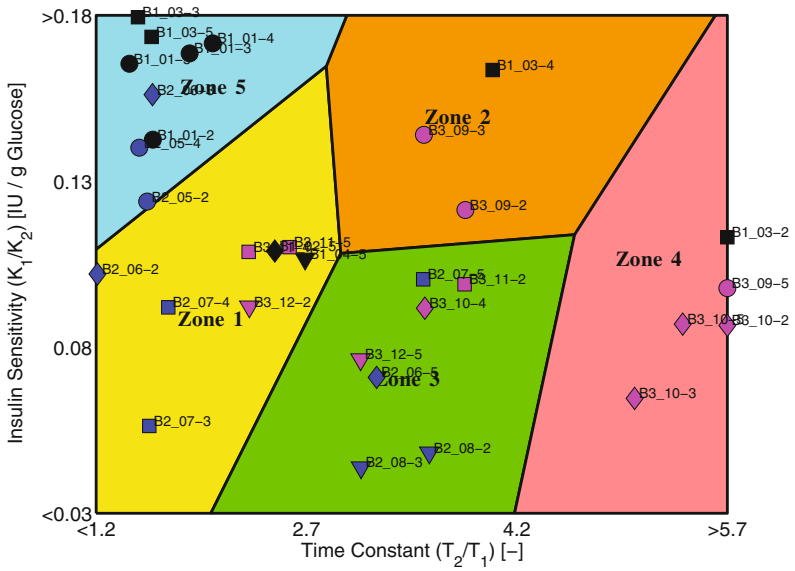

Fig. 2 Results of the system identification for all patients and all days projected into 2D and results of the subsequent clustering using the k-means algorithm

specific colour corresponds to one specific patient. The 12 patients are named B1_01 through B3_12. Connected to the name of the patient by a minus there is also the number of the day of core phase of the study (2 through 5). Not the entire data range is shown, but the values are clipped at a lower limit of $K_1/K_2 = 0.03$ and $T_2/T_1 = 1.2$ and an upper limit of $K_1/K_2 = 0.18$ and $T_2/T_1 = 5.7$.

It becomes evident from Fig. 2 that the number of successful fits varies from patient to patient. There are patients for whom only one suitable fit was found (e.g. for the patient B1_04), whereas for others the fits for all four days have an acceptable quality (e.g. for the patient B3_10). Furthermore, it can be seen that depending on the patient there could be a larger or a smaller scattering of the identified parameter sets. For example the values for all four days are very close together for patient B1_01, whereas there are tremendous differences for patient B1_03.

Using the results of the system identification (after the projection of the parameters into 2D) a clustering of data was performed in order to identify different classes of patients. It should be kept in mind that the system identification was performed in a way that the best fit is achieved for each patient and each specific day. Quite possibly, a different identification approach (as it was, e.g. used in [12]) might lead to fits with a quite similar quality and at the same time reduce part of the intra-patient variability that is only caused by the chosen approach for the system identification. Additionally, it should be mentioned that all the points from Fig. 2 were used for the further analyses presented in this publication. This means that depending on how well the identification worked, up to four points exist for the same patient. Therefore, the

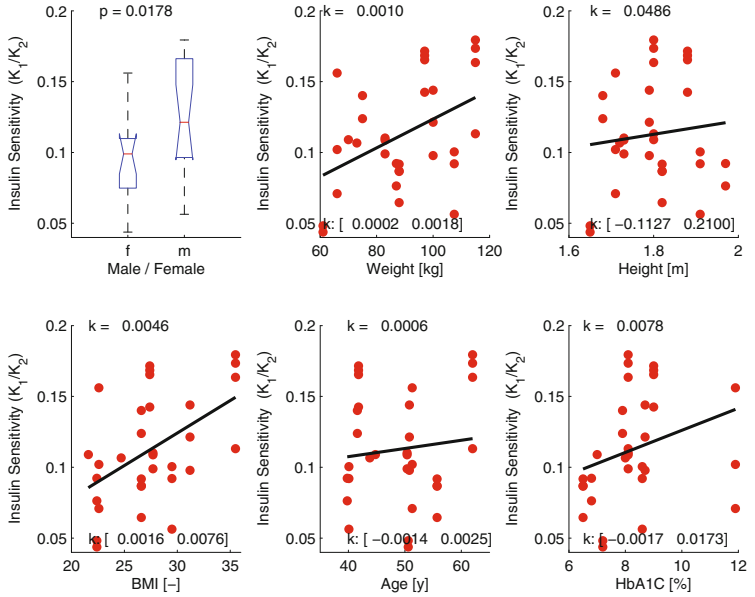

Fig. 3 Check for statistically significant trends between insulin sensitivity (K_1/K_2) and basic patient data. From *top left* to *bottom right* panel: sex, weight, height, BMI, age and HbA1C

data from patients with a higher number of successful fits get a stronger weight than those from patients for whom a lower number of suitable fits was found. However, a sensitivity study has been performed that confirmed that even for the case of only using one data point per patient (by using an average point for each patient for the case of multiple successful fits) would lead to very similar conclusions as presented in the subsequent paragraphs.

Even though there are a lot of limitations and shortcomings in the work that is presented here, the results could still be useful for further analyses on larger databases. One could hope to answer the following questions:

- Are there statistically significant correlations between basic patient data and the parameters of the identified model?
- Are these correlations consistent with what is known about the physiology, i.e. do these correlations make sense from a physiological point of view?
- Are there clusters of patients visible in the parameter space?

In order to answer the first question statistical checks were performed to identify correlations between both, the insulin sensitivity K_1/K_2 and the dimensionless time constant T_2/T_1, and basic patient data that are easily available, i.e. sex, weight, height, BMI, age and HbA1C of the patients. The results of these tests are shown in Figs. 3 and 4.

In order to check for trends w.r.t. the sex of the patients (which are categorical data) the one-way analysis of variance (ANOVA) was used (implemented in MATLAB's

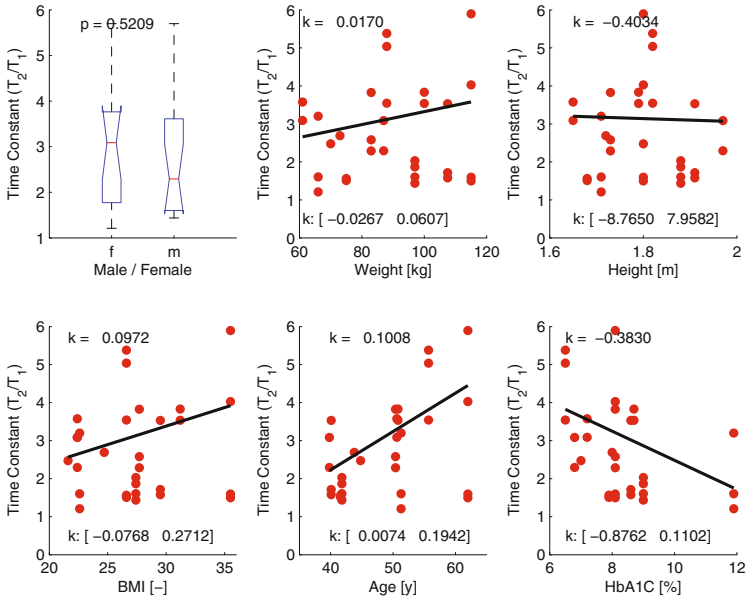

Fig. 4 Check for statistically significant trends between dimensionless time constant (T_2/T_1) and basic patient data. From *top left* to *bottom right* panel: sex, weight, height, BMI, age and HbA1C

ANOVA1 routine), whereas for the other categories (which are continuous data) simple linear regression was employed. A trend was defined as statistically significant if p in the one-way ANOVA is smaller than 0.05 or if the 95 % confidence interval of regression parameter k (in $y = k * x + d$) does not include 0. These values (the p value of the ANOVA and the confidence interval of k) are also printed in Figs. 3 and 4. It can be seen that statistically significant trends could be identified between the insulin sensitivity K_1/K_2 and the patients' sex, weight and BMI and the dimensionless time constant T_2/T_1 and the patients' age. For the relationship between insulin sensitivity and the patients' height, age and HbA1C, as well as for the relationship between dimensionless time constant and sex, weight, height, BMI and HbA1C no statistically significant correlations were found.

Having found some statistically significant correlations, question number two is whether or not these correlations match with what is known about the physiology and with what was found in previous clinical studies. In [11] the effect of sex, age and obesity on the insulin sensitivity was studied. According to the authors of [11] it was found that obese patients have a significantly lower insulin sensitivity than lean patients. Furthermore, it was found that insulin sensitivity is lower for men than for women. Additionally, the authors state that no statistically significant relationship between age and insulin sensitivity could be identified. These three findings match with what was found in the current work (see Fig. 3). Regarding trends w.r.t. the time constants it is stated in [1] that elderly subjects were found to have higher postprandial

glucose levels due to a higher rates of meal appearance (only for women) and a lower rate of glucose disappearance (for both, men and women). This corresponds to the finding from the current study that T_2/T_1 increases with age. Therefore, all principal statistically significant correlations as identified in the current study could be verified with the findings from clinical studies as described in the literature.

3.3 Clustering and Classification

In a further step it was tried to answer the question whether or not it is possible to identify clusters of patients in the parameter space. Based on the scattering as visible in Fig. 2 it becomes evident that assigning each patient to one specific group is a difficult task. Therefore, it was tried to identify clusters of parameter sets independent of the patient that the specific parameter set belongs to and to divide the area in Fig. 2 based on the location of these clusters. The task of clustering points has been performed in MATLAB using the k-means algorithm (routine KMEANS) and Euclidian distances (after z-score-transformation of the data). The best division in distinct groups (measured through the silhouette value) could be achieved using five clusters in the k-means algorithm. Based on the location of the clusters the area in the K_1/K_2 versus T_2/T_1 has been divided into five zones through straight lines using MATLAB's CLASSIFY routine.

The results of this division are shown in Fig. 2 with the five different zones plotted in distinct colours. It can be seen in this figure that for some patients the points at different days are not always assigned to the same zone. This is most probably due to the fact that there exists a huge intra-patient variability in the parameters which complicate the attempt to classify patients based on their model parameters. Nevertheless, it was checked whether the five different zones can be distinguished based on basic patient data. This check was performed using Fisher's Exact Test [7] for the sex of the patients and the one-way ANOVA (using MATLAB's ANOVA1) for the other patient data (weight, height, BMI, age, HbA1C). For these tests it has to be kept in mind that they do not evaluate whether or not all the different zones can be clearly distinguished. Instead they only test whether or not the null hypothesis that the data from all zones were basically samples of one common data pool can be rejected. The results of the statistical tests are shown in Fig. 5. It can be seen that the one-way ANOVA showed statistically significant differences (at a 5 % level) between the different zones for the patients' BMI and age. Furthermore, Fisher's Exact Test showed a statistically significant difference for the patients' sex ($p = 0.0013$). These results confirm that the parameters for which statistically significant trends could be identified versus the insulin sensitivity K_1/K_2 and the dimensionless time constant T_2/T_1 are also the most promising ones for a future classification of patients.

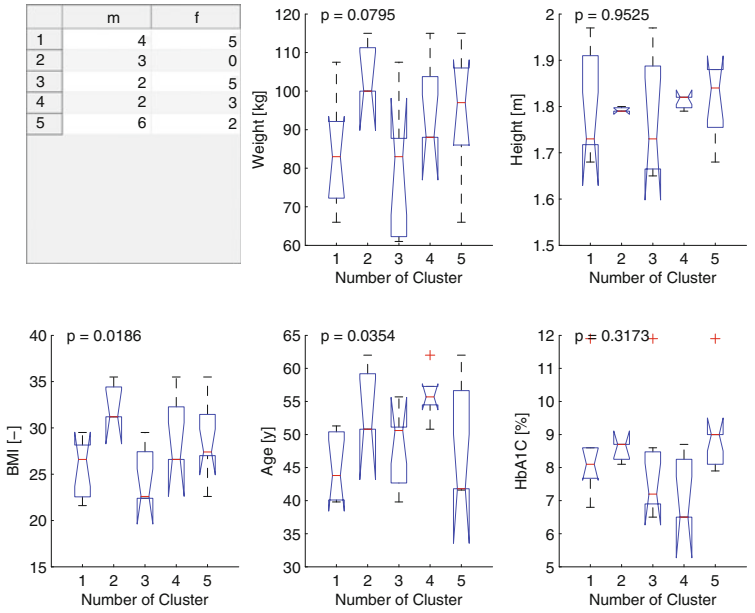

Fig. 5 One-way ANOVA, identified model parameters: Check for statistically significant differences in the basic patient data for the classes as found by the clustering. From *top left* to *bottom right* panel: sex, weight, height, BMI, age and HbA1C

3.4 Discussion of Results and Further Outlook

In this chapter the BG dynamics were modelled using a simple transfer function model. The basic idea is to perform a classification of the patients based on the identified parameters of the model. The classification is performed using the parameters K_1/K_2 that described the insulin sensitivity of the patient (high values correspond to a low sensitivity) and the dimensionless time constant T_2/T_1 (high values correspond to a slower insulin response of the BG level compared to the carbohydrate response). In the current work test calculations using data from a recent clinical study have been performed. Using these data it was found that there are statistically significant correlations between the model parameters and the patients' sex, weight, BMI and age. The physiological interpretation of the findings could also be confirmed by findings from dedicated clinical studies as documented in the literature.

An important question to be asked is whether or not it is possible to classify patients according to their overall BG dynamics based on easily available data like age or BMI. This question cannot be answered conclusively based on the findings of the current work. Instead, the study showed the challenges connected with the a priori classification of patients. The biggest challenge for such a classification is for sure the tremendous intra-patient variability. As can be seen in Fig. 2 the scattering of identified parameters for one patient can be bigger than the differences between

two different patients. This tremendous intra-patient variability is a well-known fact and has been documented extensively in the literature. A good overview of the topic can be found in [10]. In that paper it is stated that the intra-patient variability of the insulin action in clamp studies was found to be between 15 and 25 % for fast acting insulin. The inter-patient variability of the insulin action was stated to be around 10 % higher than the intra-patient variability. The insulin action as described in [10] seems to be closely connected with the insulin sensitivity K_1/K_2 from the current work. The intra-patient variability of K_1/K_2 was found to be in the same order of magnitude as the one for the insulin action stated in [10]. In our study, only two patients had a bigger value of the intra-patient variability of the insulin sensitivity (B2_06: 38 %, B2_07: 28 %). For the intra-patient variability of the dimensionless time constant T_2/T_1 no direct analogous could be found in [10]. The most similar parameter seems to be the time of the maximum insulin concentration in blood after subcutaneous injection (t_{max}). For this parameter an intra-patient variability, as found in two different studies, of 25 and 107 % is stated. In the current work a huge intra-patient variability of T_2/T_1 of up to 60 % was found which also matches the huge variability as stated in [10]. However, it should also be pointed out that the parameters as identified in the current work are not instantaneous values but rather average values over an entire day period. Therefore, one could expect that the effect of the intra-patient variability can be reduced considerably by fitting the data of an entire day instead of only for one specific meal. Maybe a smaller intra-patient variability could be achieved by using different methods for the system identification. One could, for example think of artificially restricting the intra-patient variability as described in [12]. Using the methodology from that paper it might be possible to achieve a similar quality of the fits with a significantly lower value for the intra-patient variability.

The next steps in the work would be to extend the study to a bigger database of patients. The current work was performed using only the data from 12 patients. It will be interesting to see if the same kind of statistically significant correlations can also be found for a larger database of CGMS recordings. Furthermore, it is planned to try to reduce the intra-patient variability to a minimum using the method described in [12]. Additionally, the way of evaluating fits should be reconsidered. For the current work the CGMS signals were assumed to be a perfect representation of the BG dynamics. This, of course, is not true. In the future a more physiological evaluation should performed through the assessment of transient BG profiles as they can be reconstructed from CGMS measurements and regular SMBG through the algorithm described in [5]).

Furthermore, the evaluation of the quality of fits should be rethought. For the current work the quality of a fit was assessed using the VAF/FIT value. Fits with a VAF smaller than 80 % were discarded as unsuccessful (as it was also done in [6]). It might be better, however, to switch to another approach that also takes into account the measurement errors of SMBG and CGMS. Furthermore, the comparison with a horizontal line at the average measurement value does not seem very intuitive because it might lead to the discard of reasonable fits for the case that the measured CGMS signal does not show enough inclinations. However, to the best knowledge of the

authors such an enhanced evaluation method as described here (and the corresponding cost function for the optimization of the fits according to this evaluation method) still remains to be developed.

4 Analysis of the High Frequency Content of CGMS Signals

In the following section a moving average filter with a varying filter width is applied on CGMS recordings and the variance between filtered and unfiltered signal as a function of the filter width is analysed. The variance as a function of filter width is used to classify patients.

4.1 Filtering of CGMS Signals

Additionally to a classification based on the overall BG dynamics of the patients (as described in the chapter "CGM—How Good Is Good Enough?") a second method is proposed in the current publication. The basic idea of this second method is to classify patients based on the differences in the high frequency content of CGMS signals. In order to do so the CGMS signal is filtered using a simple moving average filter with a varying filter width and the variance between filtered and unfiltered signals is analysed.

In order not to mix the effect of different sensors and different patients, this part of the study has been performed using only the signals from one of the Navigator systems. The Navigator was chosen because it is stated in [8] that in average (over all patients and all days) the Navigator showed a significantly lower MARD than the other two systems. As already mentioned before, the analysis focussed on the core phase of the clinical study from 7:00 on day 2 until 7:00 of day 6. For the filtering operation a simple, non-causal moving average filter with equal weights for each element was used. The filtered signal was thus calculated using the following formula:

$$y_{f,i} = \frac{1}{2q+1} \sum_{j=i-q}^{i+q} y_j \tag{6}$$

In this formula $y_{f,i}$ is the filtered signal at time step i, y_j corresponds to the recorded CGMS signal at time j and q is the filter width. A filter width of 0 corresponds to the unfiltered signal, whereas a filter width of, e.g. 5 would mean that the filtered signal at time step i is calculated using the signals from $i-5$ up to $i+5$ by simple averaging those signals. A larger filter width corresponds to a smoother signal and by progressively extending the filter width the filtered signal will become flatter and flatter. As a consequence the variance between filtered and unfiltered signal will increase with increasing filter width. This only holds up to a certain filter width at

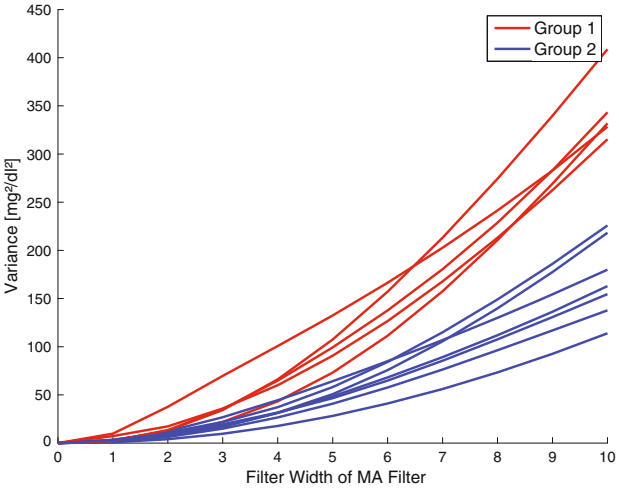

Fig. 6 Variance between filtered and unfiltered CGMS signal as a function of filter width and results of the subsequent clustering using the k-means algorithm

which a more or less flat filtered signal is generated. Starting from that point the variance between filtered and unfiltered signal will stay (almost) constant for all subsequent values of the filter width.

For the current study the focus is on the filtering behaviour in the high frequency part of the signals. The Navigator CGMS provides one measurement value of the IG every ten minutes. It was found that up to a filter width of around 10 (corresponding to $+/-$ 100 min) the high frequency part of the signals is smoothed out, whereas the general behaviour of the signals remains unchanged. For even larger filter width parts of the general signal trends start to disappear. Therefore, the behaviour of the variance between unfiltered and filtered signal has been studied only for filter widths q up to 10. A plot of the variance between filtered and unfiltered signal as a function of filter width can be seen in Fig. 6.

4.2 Trends and Classification

As a next step it was tried to identify clusters of patients for the functions of variance versus filter width. A clustering attempt was performed using the k-means algorithm (MATLAB's KMEANS routine) and yielded best results (measured by the silhouette value) using two clusters as is can be seen in Fig. 6 (called "Group 1" and "Group 2").

As already done in Sect. 3.2 it was again tried to identify statistically significant relationships between the clusters found using the k-means algorithm and basic patients data. To check for statistically significant differences between the two clusters regarding the patients' sex, Fisher's Exact Test [7] was used again.

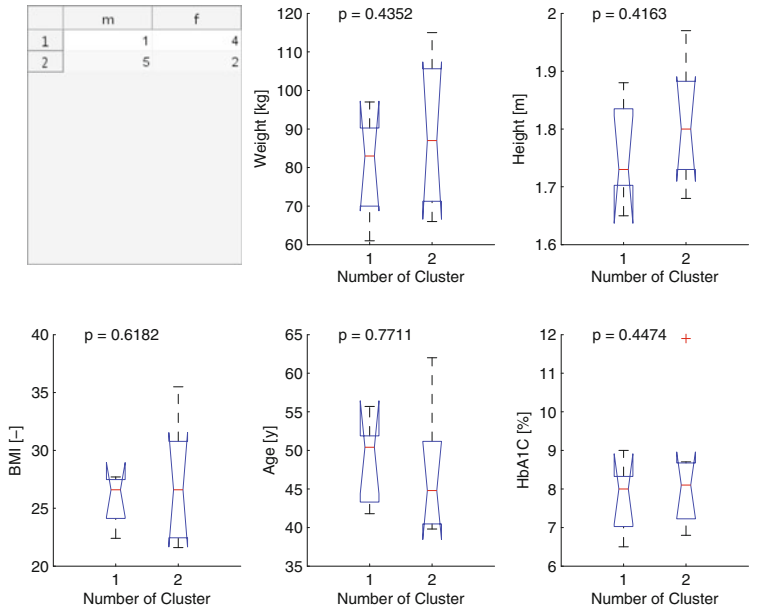

Fig. 7 One-way ANOVA, variance between filtered and unfiltered signal: Check for statistically significant differences in the basic patient data for the classes as found by the clustering. From *top left* to *bottom right* panel: sex, weight, height, BMI, age and HbA1C

No statistically significant difference was found ($p = 0.1136$). A check for differences in weight, height, BMI, age and HbA1C was performed using the one-way ANOVA (MATLAB's ANOVA1 routine).

The results of the statistical tests are shown in Fig. 7. It was found that there are no statistically significant differences between the two identified clusters regarding basic patient data. This fact could easily lead to the conclusion that clustering patients using the high frequency content of their CGMS signals does not seem a worthwhile procedure. However, interesting results were found when checking for statistically significant differences between the two identified clusters and low blood glucose index (LBGI) as well as high blood glucose index (HBGI). LBGI and HBGI have been introduced in [14] as a measure to evaluate a patient's risk for hypoglycemia and hyperglycemia based on the results of previous SMBG values. Both, LBGI and HBGI, are calculated from the so-called Kovatchev risk function 7. This function returns low values in the euglycemic range, but high values in both, the hypoglycemic and the hyperglycemic region:

$$r(BG) = 22.77 * (\ln(BG)^{1.084} - 5.381)^2 \tag{7}$$

Using this definition of the risk function, the LBGI and the HBGI can be calculated using the following formulae:

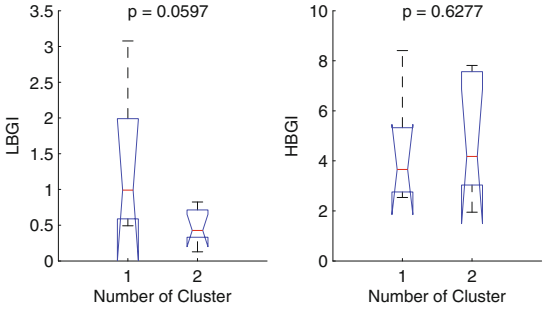

Fig. 8 One-way ANOVA, variance between filtered and unfiltered signal: Check for statistically significant differences for LBGI and HBGI

$$LBGI = \frac{1}{N}\sum_{i=1}^{N} rl(x_i) \tag{8}$$

$$HBGI = \frac{1}{N}\sum_{i=1}^{N} rh(x_i) \tag{9}$$

with:

$rl(BG) = r(BG)$ if $\ln(BG])^{1.084} - 5.381 < 0$ and 0 otherwise and
$rh(BG) = r(BG)$ if $\ln(BG)^{1.084} - 5.381 > 0$ and 0 otherwise.

Calculating LBGI and HBGI for all of the 12 patients in the clinical study using the CGMS data (from the Navigator) during the core phase of the clinical study, values for the LBGI between 0.13 and 3.08 and values for the HBGI between 1.95 and 8.41 were found. Again, a test for statistically significant differences between the two clusters using the one-way ANOVA was performed. The results of this test are shown in Fig. 8.

It was found that the patients in cluster 1 (showing a higher variance as a function of filter width) have a considerably higher value of LBGI (and therefore most likely also a higher risk of severe hypogycemia, see [14]) than patients from cluster 2. For the HBGI no significant differences were found between cluster 1 and cluster 2.

Even if the difference in LBGI between the two identified clusters is not statistically significant on a 5 % level (p = 0.0597), it seemed worthwhile to have a closer look at the relationship between the high frequency content of the CGMS signals and the LBGI. In the first step the functions for variance between filtered and unfiltered signals versus filter width were approximated using the following relationship:

$$\sigma = k * q \tag{10}$$

According to this formula the standard deviation between filtered and unfiltered signal (σ; with $\sigma = \sqrt{variance}$) is proportional to the filter width (q). It was found that this approximation is very good for almost all of the functions of variance versus filter width. This way the functions could be condensed to one single regression parameter k per patient.

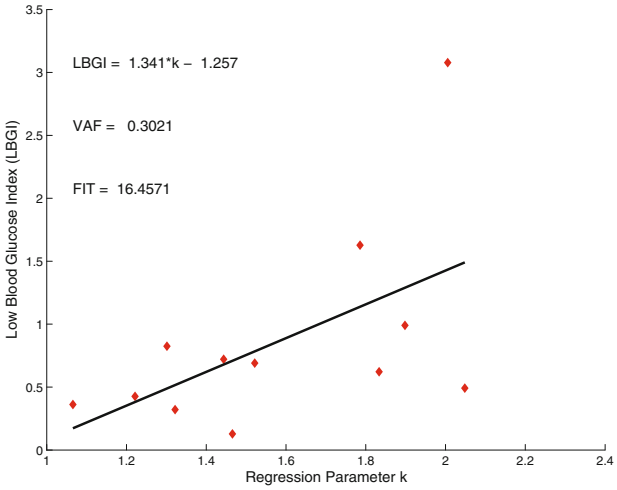

Fig. 9 Linear regression: Regression parameter k (that describes the linear relationship between the filter width and the standard deviation between filtered and unfiltered signal) versus LBGI

In the next step the values of the LBGI for all patients have been plotted as a function of the regression parameter k. This can be seen in Fig. 9. Even though there is a considerable scattering of data, a tendency of higher values for the LBGI for the case of higher values of the regression parameter k seems evident. To quantify this tendency a simple linear regression was performed. It was found that the relationship is not statistically significant on a 5 % level, but still likely ($p = 0.0642$).

4.3 Discussion of Results and Further Outlook

In the analyses described in this chapter the feasibility of using the high frequency content of CGMS signals to classify patients has been checked. For this purpose the CGMS signals obtained in a recent clinical study have been filtered using a simple moving average filter with a varying filter width and the variance between filtered and unfiltered signal was studied as a function of the filter width. It was found that the function "variance versus filter width" is different from patient to patient and could be used to assign patients to different groups. For the preliminary tests performed for this work two groups of patients have been used. No statistically significant difference between the two groups could be found w.r.t. any of the available categories of basic patient data. However, it was found that a relationship between the high frequency content of the CGMS signals and the LBGI is likely. It seems that the presence of more high frequency fluctuations in the CGMS signals are linked to a higher risk of hypoglycemia. Such a relationship, however, does not seem intuitively and requires further investigation in order to verify or falsify it.

In the near future it is intended to extend the analysis described in this chapter to a larger group of patients in order to be able to make more conclusive statements about the relationships found in the current study. If it is possible to verify a relationship between high frequency noise in the CGMS signals and a higher risk of hypoglycemia, it is also important to find a plausible mechanism/explanation for this relationship. Additionally, it would be interesting to see if similar/identical relationships can be found using different kinds of non-causal filters.

Acknowledgments The authors gratefully acknowledge the sponsoring of this work by the COMET K2 center "Austrian Center of Competence in Mechatronics (ACCM)". The COMET Program is funded by the Austrian federal government, the federal state Upper Austria and the scientific partners of ACCM.

References

1. Basu, R., Dalla Man, C., Campioni, M., Basu, A., Klee, G., Toffolo, G., Cobelli, C., Rizza, R.A.: Effects of age and sex on postprandial glucose metabolism. Diabetes **55**, 2001–2014 (2006)
2. Cescon, M., Johansson, R., Renard, E., Maran, A.: Identification of individualised empirical models of carbohydrate and insulin effects on T1DM blood glucose dynamics. Intern. J. Control **87**(7), 1438–1453 (2014)
3. Clarke, W.L., Cox, D.J., Gonder-Frederick, L.A., Julian, D., Schlundt, D., Polonsky, W.: Reduced awareness of hypoglycemia in adults with IDDM. Diabetes Care **18**(4), 517–522 (1995)
4. Cobelli, C., Dalla Man, C., Sparacino, G., Magni, L., De Nicolao, G., Kovatchev, B.P.: Diabetes: Models, signals, and control. IEEE Rev. Biomed. Eng. **2**, 54–96 (2009)
5. Del Favero, S., Facchinetti, A., Sparacino, G., Cobelli, C.: Improving accuracy and precision of glucose sensor profiles: Retrospective fitting by constrained deconvolution. IEEE Trans. Biomed. Eng. **61**(4), 1044–1053 (2014)
6. Fernandez, M., Villasana, M., Streja, D.: Glucose dynamics in type I diabetes: Insights from the classical and linear minimal models. Comput. Biol. Med. **37**, 611–627 (2007)
7. Fisher, R.A.: On the interpretation of χ^2 from contingency tables, and the calculation of p. J. R. Stat. Soc. **85**(1), 87–94 (1922)
8. Freckmann, G., Pleus, S., Link, M., Zschornack, E., Klötzer, H.M., Haug, C.: Performance evaluation of three continuous glucose monitoring systems: Comparison of six sensors per subject in parallel. J. Diabetes Sci. Technol. **7**(4), 842–853 (2013)
9. Guyton, A.C.: Textbook of Medical Physiology, 11th edn. Elsevier Inc., Philadelphia (2006)
10. Heinemann, L.: Variablility of insulin absorption and insulin action. Diabetes Technol. Ther. **4**(5), 673–682 (2002)
11. Kerakelides, H., Irving, B.A., Short, K.R., O'Brian, P., Sreekumaran Nair, K.: Age, obesity, and sex effects on insulin sensitivity and skeletal muscle mitochondrial function. Diabetes **59**(89–97), 1207–1214 (2010)
12. Kirchsteiger, H., Johansson, R., Renard, E., del Re, L.: Continuous time interval model identification of blood glucose dynamics for type 1 diabetes. Intern. J. Control (2014). doi:10.1080/00207179.2014.897004
13. Kirchsteiger, H., Pölzer, S., Johansson, R., Renard, E., del Re, L.: Direct continuous time system identification of miso transfer function models applied to type 1 diabetes. In: Proceedings 49th IEEE Conference on Decision and Control (CDC), pp. 5176–5181 (2011)

14. Kovatchev, B.P., Straume, M., Cox, D.J., Farhy, L.S.: Risk analysis of blood glucose data: A quantitative approach to optimizing the control of insulin dependent diabetes. J. Theor. Med. **3**, 1–10 (2000)
15. Naumova, V., Pereverzyev, S.V., Sampath, S.: A meta-learning approach to the regulatrized learning-case study: Blood glucose prediction. Neural Netw. **33**, 181–193 (2012)
16. Rodrigues, T.C., Pecis, M., Canani, L.H., Schreiner, L., Kramer, C.K., Biavatti, K., Macedo, B., Esteves, J.F., Azevedo, M.J.: Characterization of patients with type 1 diabetes mellitus in southern brazil: Chronic complications and associated factors. Rev. Assoc. Med. Bras. **56**(1), 67–73 (2010)
17. Tariq, A., Usman Akram, M., Shaukat, A., Khan, S.A.: Automated detection and grading of diabetic maculopathy in digital retinal images. J. Digital Imaging **26**(4), 803–812 (2013)
18. Ter Braak, E.W., Appelman, A.M., Van De Laak, M.F., Stolk, R.P., Van Haeften, T.W., Erkelens, D.W.: Clinical characteristics of type 1 diabetic patients with and without severe hypoglycemia. Diabetes Care **23**(10), 1467–1471 (2000)
19. The Diabetes Control And Complications Trial Research Group: The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. New England Journal of Medicine **329**(14), 977–986 (1993). http://www.nejm.org/doi/full/10.1056/NEJM199309303291401

Prevention of Severe Hypoglycemia by Continuous EEG Monitoring

Claus Bogh Juhl, Jonas Duun-Henriksen, Jens Ahm Sørensen, Anne Sophie Sejling and Rasmus Elsborg Madsen

Abstract Background: The brain is dependent on constant glucose supply and hypoglycaemia results in reduced cognition, unconsciousness, seizures, and possible death. Prevention of hypoglycaemia is accordingly a key point in diabetes treatment. The effect of hypoglycemia on the electrical activity of the brain is well described. We propose an alarm device for severe hypoglycaemia based on continuous electroencephalography (EEG), real-time data analysis by an automated algorithm and an auditory alarm. Methods and results: People with type 1 diabetes (T1D) were exposed to hypoglycaemia by excess insulin administration. EEG was scored visually by neurophysiologists for deviations compatible with neuroglycopenia. From these initial experiments, a multiparameter algorithm was developed by applying an artificial neural network. Subsequently series of experiments were conducted in T1D patients in order to improve the algorithm and to test its clinical applicability. Thus people with T1D with normal, reduced, or absent awareness of hypoglycaemia were exposed to hypoglycaemia both during daytime and during sleep. EEG was analyzed by the automated algorithm. All patients developed hypoglycaemia-associated EEG changes. During the sleep experiments, these changes occurred irrespective of sleep stage, and the majority of patients woke up when they received an auditory alarm at the onset of hypoglycaemia-associated EEG changes. We found that the EEG changes were independent of diabetes duration, awareness status, and counter-regulatory hormone response. In subsequent studies, a miniaturized partly implanted EEG recorder was tested. The device consists of a small implant

C.B. Juhl (✉) · J. Duun-Henriksen · R. Elsborg Madsen
Hyposafe A/S, 2800 Lyngby, Denmark
e-mail: cj@hyposafe.dk

C.B. Juhl
Department of Endocrinology, Sydvestjysk Hospital, 6700 Esbjerg, Denmark

J.A. Sørensen
Department of Plastic Surgery, Odense University Hospital, 5000 Odense C, Denmark

C.B. Juhl · A.S. Sejling
Faculty of Health Science, University of Southern Denmark, 5000 Odense, Denmark

A.S. Sejling
Department of Cardiology, Nephrology and Endocrinology, NordsjæLands Hospital, 3400 Hillerød, Denmark

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*,
Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_5

recording the EEG and an external device which wirelessly receives EEG and gives an alarm when hypoglycaemia-associated EEG changes are apparent. In preliminary studies of one month duration, we found that the device was well tolerated and successfully warned the patients in time to take appropriate action before severe hypoglycaemia was present. Conclusion: The results obtained so far hold promises for the development of an EEG-based alarm for severe hypoglycaemia in people with T1D. Hypoglycaemia-associated EEG changes seem to be a general feature which makes a common algorithm applicable. Further studies are needed to define the sensitivity and tolerability of the partly implanted EEG-based device for severe hypoglycaemia during long-term use.

1 Background

Hypoglycaemia is defined as blood glucose lower than 3.9 mmol/l. The patient may or may not be aware of its occurrence, and if glucose continues to fall the patient may need help from a third person to recover. This potentially dangerous condition is termed severe hypoglycaemia. Hypoglycaemia remains a limiting factor in the treatment of type 1 diabetes (T1D). The brain is dependent on constant glucose supply and hypoglycaemia results in reduced cognition, unconsciousness, seizures, and possible death. People with diabetes fear the occurrence of hypoglycaemia equally to diabetes complications such as end-stage renal disease and blindness [1]. Prevention of hypoglycaemia is accordingly a key point in diabetes treatment. This goal is sought by a panel of different approaches. Structured outpatient diabetes training results in a significant reduction in the occurrence of severe hypoglycaemia [15]. Use of insulin analogs with superior pharmacokinetics as opposed to human insulin reduces the frequency of nocturnal hypoglycaemia in particular [14]. Continuous glucose monitoring (CGM) is increasingly used in the treatment of T1D. When combined with continuous subcutaneous insulin infusion, this allows a temporary cessation of insulin infusion in case of low-glucose measure. Recent studies have shown that the use of CGM with low-glucose suspend is able to reduce the risk of nocturnal hypoglycaemia in T1D [10, 11]. In clinical studies as well as from real-life observations it is, however, clear that adherence to the use of CGM may often be inadequate to achieve clinical benefits. Some people find the devices problematic or uncomfortable to use, others are annoyed by the inaccuracy of the measurements, and even after short-term use the coverage in subgroups of patients is down to 50 % of the time while others quit the use completely [17, 24].

Electroencephalography (EEG) is a state of the art procedure to obtain information about the metabolic and electric status of the brain. The technique was described in 1929 by Hans Berger [2] and has continuously been improved with respect to sampling frequency and electrode density, and computer techniques allow advanced mathematical evaluation of data beyond visual inspection of the EEG.

Hypoglycaemia-associated EEG changes were described in 1951 by two independent research groups [3, 19]. Ross et al. recorded EEG during induced hypoglycaemia

and subsequent visual examination of the EEG and demonstrated a shift of the dominant frequency toward a slower electrical activity of the brain [19]. Boudin on the other hand, recorded EEG during spontaneous hypoglycaemia in patients diagnosed with insulinoma with analogous finding of theta and delta activity in the EEG [3]. Since then a number of studies have assessed the effect of hypoglycaemia on EEG both in healthy individuals and in diabetes patients. The cause of hypoglycaemia included spontaneous events as well as pharmacologically induced hypoglycaemia, and the shift toward a slower frequency is a general feature. Also the EEG changes seem to be identical irrespective of diabetes duration and the patient's ability to sense hypoglycemia [18] and is unaffected by recent antecedent hypoglycaemia [21].

Pramming and co-workers studied people with T1D during graded hypoglycaemia induced by insulin infusion and suggested that the brain has an individual glucose threshold for hypoglycaemia-associated changes [16]. This threshold has been shown to be reproducible within each subject with type 1 diabetes [8]. Assuming that these EEG changes are sufficiently specific and general, it has been suggested to construct an alarm device based on continuous EEG recording and real-time mathematical processing of the data. Thus if EEG changes occur before cognitive decline in patients with insulin-treated diabetes or other conditions associated with risk of hypoglycaemia, such a device may be a useful approach for warning against severe hypoglycaemia [8, 23].

Based on this, we propose an alarm device for severe hypoglycaemia based on continuous electroencephalography (EEG) and real-time data analysis by an automated algorithm and an auditory alarm.

2 Clinical Studies—Proof of Concept

In this section, we describe the experiments we have performed in order to test the feasibility of an EEG-based alarm for severe hypoglycaemia. The results summarized here are achieved with commercially available equipment. In a later paragraph, we will describe experiments conducted with an implantable hypoglycaemia alarm device.

In our initial experiments, we tested the spatial localization of the hypoglycaemia-associated EEG changes, and these experiments have been described elsewhere [5]. In accordance with the literature, we found that EEG changes were most abundant in the temporal region and on the cranial vertex. Since muscular activity may give rise to signal artifacts, it is preferential to perform the EEG recording in the areas with minimal muscular activity. Placing the superior electrode in the area between P3 and C3 and the inferior electrode at approximately T5 according to the international 10/20 system, it is possible to achieve an optimal signal to artifact ratio in the EEG.

To achieve knowledge about the timing and the nature of the EEG changes, we induced graded hypoglycaemia in 15 patients with T1D, duration (mean (range)) 26.5 (6–53) years, including both aware and unaware patients. Cognitive function was evaluated by repeated cognitive tests (counting backward and subtraction of

seven from a given number). Insulin infusion was replaced by glucose infusion when patients showed obvious cognitive or physical signs of hypoglycaemia or when plasma glucose reached 1.8 mmol/l. Hypoglycaemia-associated EEG changes were found by the automated algorithm in all 15 subjects at a plasma glucose ranging from 2.0 to 3.4 mmol/l and occurred 2928 min (meanSD)(range 3–113 min) before termination of insulin infusion. Since a main purpose of this experiment was to achieve data for further development of the algorithm, the patients did not receive a real-time alarm, and therefore, it is not possible to state if they would have been able to react following an alarm. In 12 of 15 patients, however, EEG changes occurred before severe neuroglycopenia was apparent as evaluated by the cognitive tests. In three cases, the patients were moderately cognitively impaired at the time of EEG changes [8].

Nocturnal hypoglycaemia constitutes a special problem. Approximately 50 % of all hypoglycaemia episodes occur during sleep, and prolonged episodes of hypoglycaemia occur in both children and adults in 8.5 % of all nights as measured by continuous glucose monitoring [9]. Sleep EEG differs significantly from daytime EEG. Stages of deep sleep is characterized and defined by the presence of slow wave dominance, and accordingly it is a specific challenge to differentiate between deep sleep and hypoglycaemia in the development of a robust algorithm for hypoglycaemia detection.

To address this, we conducted a trial in which ten patients with T1D, duration (mean (range)) 23.7 (10–37) years and impaired awareness of hypoglycaemia were exposed to insulin-induced hypoglycemia during daytime and during sleep. EEG was recorded by a single electrode (Foramen Electrode AD-Tech Medical (WI, USA), length 300 mm, diameter 1.1 mm with 3 contact points with a center-to-center distance of 30 mm) positioned subcutaneously over the temporal region of the brain behind the ear. The electrode was connected to an EEG recorder (g-USBamp, G-TEC, Austria) and analyzed real time by an automated multiparameter algorithm. This algorithm was based on the original daytime algorithm and optimized for hypoglycaemia detection during sleep by applying data from eight pilot-experiments of insulin-induced hypoglycaemia during sleep in people with T1D. Participants received an auditory alarm when EEG changes met a predefined threshold, and were instructed to consume a sandwich and a juice.

During the night experiments, nine out of ten developed EEG changes compatible with hypoglycaemia at a mean blood glucose of 2.0 mmol/l. Eight were awakened by the alarm. Four corrected hypoglycemia (mean BG 2.2 mmol/l) themselves by ingestion of carbohydrates, while five needed glucose infusion. There were two events of false alarm. Figure 1 shows a representative example of an experiment where the patient woke up following the auditory alarm and avoided impending severe hypoglycaemia.

Importantly EEG changes occurred irrespective of sleep stage as evaluated according to the American Academy of Sleep Medicine score [6, 23]. From the data achieved during this trial, we were able to improve the algorithm to achieve an earlier detection of hypoglycaemia, and this will accordingly be tested in future trials.

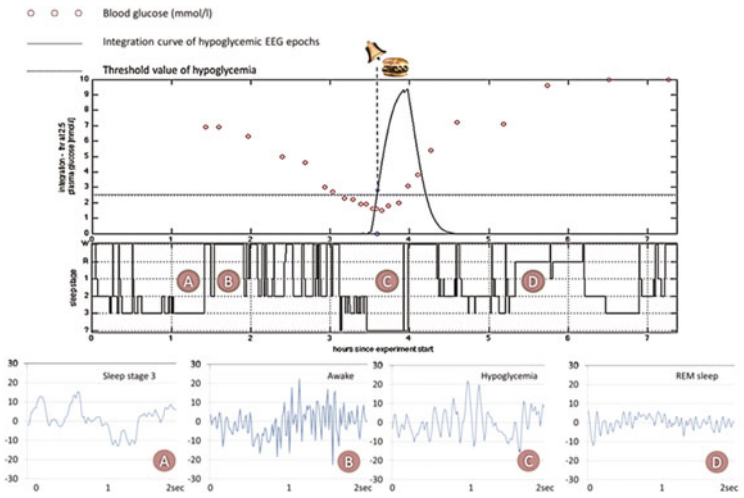

Fig. 1 Representative example of a night experiment. At progressing hypoglycaemia, the curve of integrated hypoglycaemia events exceeds the predefined threshold of hypoglycaemia-associated EEG changes. Although occurring at a rather low-blood glucose value, the patient was still able to correct impending severe hypoglycaemia by carbohydrate ingestion. The panels at the *bottom* show examples of one-channel EEG during different sleep-stages corresponding to the hypnogram in the middle

3 The Device

The experiments described in the paragraphs above were conducted with standard equipment for EEG recording. In this paragraph, we will briefly describe the device developed for warning against severe hypoglycaemia.

The alarm for severe hypoglycaemia consists of two main parts: An implant positioned subcutaneously behind the ear and an external part placed behind the ear. The implant consists of a sealed coil and an electrode with three measurement points, a length of 100 mm and a diameter of 1.1 mm pointing toward the top of the head. The implant captures the EEG and transmits the signal through a near field inductive communication link to the external part. The two parts have to be closely aligned for the system to function. The external part of the device is designed as an ear hanger (Fig. 2) which is easy to wear with a minimum of discomfort for the user. It contains a sound generator a light indicator to inform the user of critical events, e.g., impending severe hypoglycaemia and a battery which also powers the inner device through the inductive communication link.

Fig. 2 The two parts of the hypoglycaemia device. The coil and the electrode are placed subcutaneously, while the ear hanger is placed as shown on figure

4 Quantitative Evaluation of EEG Recorded with the Partly Implanted EEG Recorder

To assess correlation between surface and subcutaneous EEG, we performed simultaneous EEG from the subcutaneous implant and from surface electrodes recorded with standard EEG equipment. This was performed in five healthy individuals during three periods of one minute duration with eyes closed. The normalized correlation coefficient, $\bar{R}_{xy}$, was calculated as:

$$\bar{R}_{xy}(m) = \frac{\sum_{n=max(0,m)}^{min(N+m-1,N-1)} x_{n+m} \cdot y_n}{\sqrt{\sum_{n=0}^{N-1} x_n^2 \cdot \sum_{n=0}^{N-1} y_n^2}} \tag{1}$$

where x is the subcutaneous EEG, y is the surface EEG, N is the number of samples in each calculation, and m is the lag in the range $-N+1$ to $N-1$. If $\bar{R}_{xy} = 0$, there is no correlation between the signals while if $\bar{R}_{xy} = 1$ the signals are identical and synchronous.

The resulting cross-correlogram is shown in Fig. 3. The mean normalized correlation coefficient at lag = 0 was 0.73 which is highly significant. Also the correlation coefficient at the first-order side lobes, found at ± 22 lags, was highly significant. This verifies a continuous wave in each EEG with a frequency of 9.4 Hz and substantiates a high accordance between the two synchronously but independently recorded time series of EEG data.

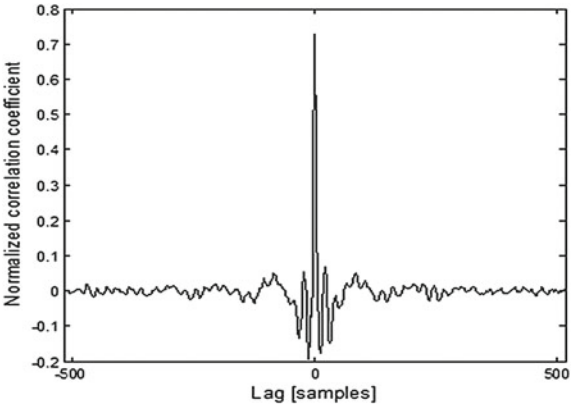

Fig. 3 Normalized cross-correlation between surface and subcutaneous EEG, while subjects have closed eyes. The maximum correlation at lag 0 is 0.73

5 Development of an Algorithm for Detection and Warning of Severe Hypoglycaemia in Type 1 Diabetes

In this paragraph, we describe the development approach and structure of the algorithm that detects impending severe hypoglycemia.

It has been shown that EEG changes appear prior to severe hypoglycemia in most cases [8]. A simple generalized model that describes the changes can be formulated as increased theta and delta activity, and decreased alpha activity. Although this simple model can be useful in controlled clinical trials to explain the changes, it is not a robust model for detecting real-life hypoglycemic situations, where we have to take intra- and inter-individual variations and environmental factors into account. These variations include differences in EEG response to hypoglycemia (frequency and amplitude), spatial differences of the EEG, day-to-day EEG response variations, EEG artifacts, and background EEG differences.

Our intention is to build a robust algorithm that can detect and warn the user about the situation prior to loss of cognitive function (neuroglycopenia), giving the user time to intervene and prevent severe hypoglycemia. Due to the design of the device and the fact that the device must be implanted, we are constrained to use only a single-bipolar EEG channel that is located in the area between P3, C3, and T5 in the 10–20 system.

Given the constraints for an algorithm, we have developed the algorithm structure that is shown in Fig. 4.

The structure of the algorithm is closely tied to our understanding of what is happening in the EEG prior to severe hypoglycopenia, namely that short (few seconds) episodes of hypoglycemia-related EEG changes (HREC) emerge in the EEG, see Fig. 5. A single HREC does not in itself provide enough evidence that severe hypoglycaemia is impending since the brain occasionally generate phenomena that are similar to HREC. The last part of the algorithm integrates the detected HREC, and

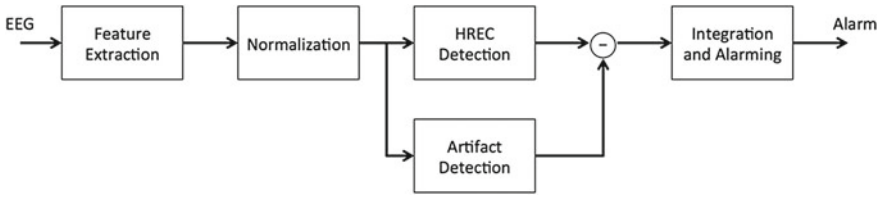

Fig. 4 Algorithm structure. For further information about the development of the algorithm see [5]. HREC: Hypoglycaemia-related EEG changes

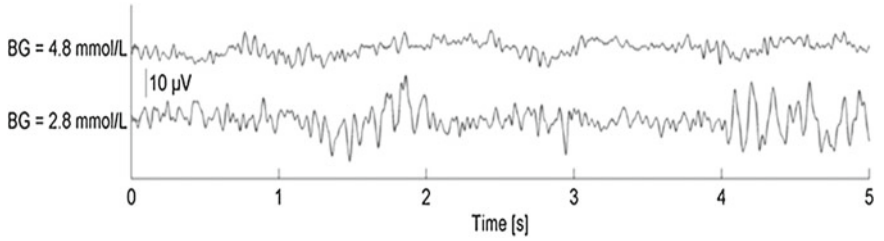

Fig. 5 Example of EEG recorded during euglycaemia (*upper panel*) and hypoglycemia (*lower panel*) where two events of hypoglycaemia-related EEG changes is seen from seconds 1 to 2 and from seconds 4 to 5

when enough evidence within a certain timeframe (3 min) is collected, the integration function releases an alarm.

Taking a step back in the algorithm structure, we have created a noise and artifact detector that removes falsely detected HREC events. Given the location of the single EEG channel, few artifact and noise situations occur. Artifacts originating from, e.g., jar movement (when chewing) are, however, strong enough to provide difficulties for most feature extractors. The artifact and noise detector is similar to the HREC detector that is described below and is trained in a similar manner with a lot of noise and artifact examples.

In parallel with the artifact detector, the HREC detector receives normalized EEG features and determines if they are positives. The classifier structure used is an approximation of an artificial neural network, allowing ultralow power processors to do the calculations in real time [5]. The performance of this classifier is similar to those of neural networks [22] or support vector machines [7] and is highly dependent on the training approach, and the data and labels that are available. A neurophysiologist reviewed the EEG time series together with the blood glucose values and labeled the presence of HREC phenomena. These labels have been used together with a learning scheme that rewards or punishes the classifier when it finds nonlabeled HREC, depending on both the glucose level and the time to hypoglycemia. The learning scheme is shown in Fig. 6.

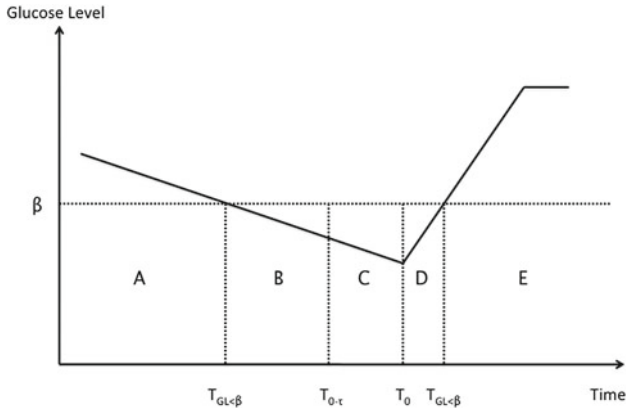

Fig. 6 Classifier learning scheme for hypoglycaemia-associated EEG changes. Data from section A and E are used as negative hypoglycemia data for the algorithm. Data from the transition section D are discarded. The algorithm is rewarded when finding nonlabeled positives in section B and C

The learning scheme has a positive effect on the whole algorithm performance, since it allows the classifier to find EEG parts that appear prior to severe hypoglycemia but goes undetected by the human eye. The approach also allows nonlabeled data to be used in the classifier training. The whole classifier-training framework is based on the leave-one-out principle, i.e., when testing on a subject that subjects data were not used for training.

The features that are fed to the classifiers are normalized to reduce the effects of differences in scull thickness, spatial location, and other parameters that shape the EEG signal distribution. The normalization is made on the individual features, since we assume that the shaping of the EEG signal is different for different frequencies. The features are normalized by subtraction of the median followed by a division of the 80 percentile. The median and percentile measures give robust measures of the underlying distribution as opposed to the mean and standard deviation measures. The normalization parameters are obtained from the first 30 min of each individual trial, thus only normal data.

The actual feature extraction is based on the raw input measure from the brain. The most essential features for the HREC classifier are frequency power estimates, where especially the delta, theta, and alpha bands are of interest. The frequency bands are of varying size and are most narrow in the range from 2–12 Hz where the most significant changes take place prior to severe hypoglycemia. The band powers are averaged and sampled at 1 Hz. From band power averages, we also calculate the relatives, i.e., alpha power divided by theta power and feed all the features to the HREC and artifact classifiers.

The 5 steps described above complete our algorithm, and the performances in the clinical studies are presented in the following.

6 Clinical Studies—Preliminary Results with Implanted Device

In this section, we will describe our experience with the implanted alarm for severe hypoglycaemia. Since only few experiments have been conducted up to now, we will show representative examples of preliminary results.

Patients with type 1 diabetes and impaired awareness of hypoglycaemia were equipped with the alarm device and were encouraged to wear the device continuously during everyday activities and during sleep for 30 days. Via a logging device, it was possible to store all EEG recorded. The patients were encouraged to measure blood glucose at least four times daily and wore a continuous glucose monitor (Guardian, Medtronic, MiniMed, CA) to be able to retrieve information about glucose values and trends aside from SMBG.

Figure 7 shows one example of 72 h of continuous EEG recording. In addition, blood glucose values and values from the continuous glucose monitoring are shown. The lower panel shows the curve of integrated EEG derived events compatible with neuroglucopenia. During the three days with documented normoglycaemia (red dots), the curve of integrated events of EEG compatible with hypoglycaemia (bottom) was continuously below the predefined threshold value of hypoglycaemia alarm.

Figure 8 gives an example of continuous EEG recording with real-time alarms at the time of impending severe hypoglycaemia. At lunch-time (13:00), the continuous glucose curve shows a fast drop in glucose. The patient receives an alarm and reacts appropriately by ingesting his lunch followed by an immediate rise in glucose. Also before supper time, both interstitial glucose and blood glucose as measured by CGM and repeated subcutaneous measurements, respectively, shows a fall. Again the patient receives a timely alarm and corrects the impending severe hypoglycaemia by ingesting his supper. No events of false alarm were seen in the present experiment.

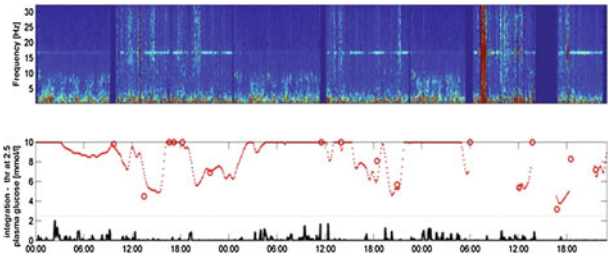

Fig. 7 72 h of normoglycaemia documented by SMBG and CGM and continuous EEG recording. *Lower panel* shows outcome of automated real-time analysis without false hypoglycaemia alarms. *Red circles* blood glucose, *red dotted line* interstitial glucose as measured by CGM, *black line* hypoglycaemia-associated EEG changes, *blue dotted line* threshold for hypoglycaemia alarm

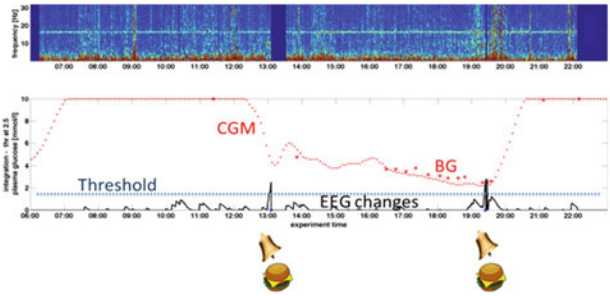

Fig. 8 Example of hypoglycaemia alarm at impending low glucose both at lunch and at supper time in a patient with T1D

7 Discussion and Perspectives

Despite the use of insulin analogs, continuous insulin delivery by insulin pumps, and continuous glucose monitoring, hypoglycaemia remains an important issue in the treatment of diabetes. Only a minority of the diabetic patients achieves the intended glycaemic target. To get closer to this target, an effective hypoglycaemia protection is essential.

Although continuous glucose monitoring is used both as a stand-alone system, as a part of sensor augmented pump therapy and integrated with continuous insulin delivery with low-glucose suspend, only a minority of patients are currently using these devices. Since the brain and the heart are the organs most vulnerable to restricted glucose supply, leading to neuroglycopenia and risk of cardiac arrhythmias, respectively, it is tempting to apply continuous monitoring of these organs as a part of a hypoglycaemia warning system. Electrocardiography as a basis of a biosensor alarm has been tested in smaller clinical trials and the prolongation of QT interval during hypoglycaemia is well described [12]. It has, however, not yet been possible to use this as a reliable hypoglycaemia warning system.

Hypoglycaemia-associated EEG changes are well described and seem to be independent of diabetes awareness status [18], the course of hypoglycaemia and of recent antecedent episodes of hypoglycaemia [20].

Some pitfalls should be taken into consideration in the development of a hypoglycaemia alarm based on EEG rather than direct measures of glucose concentration. First, a number of conditions aside from hypoglycaemia are associated with changes in the EEG and may accordingly induce either false-positive or false-negative readings. These conditions include e.g., epilepsy and other cerebral pathologies and certain drugs. The impact of these conditions should be tested in future trials. Second, the timing of EEG changes in relation to hypoglycaemia-associated cognitive impairment is crucial as the patients must be able to take appropriate action at the time of alarm. We found that most, but not all, studied subjects were cognitively intact at the time of significant EEG changes. The term hypoglycaemia-associated autonomic failure describes the situation where lack of hypoglycaemia sensation and

insufficient counter-regulatory hormone response render the patient in an increased risk of severe hypoglycaemia [4]. The mechanisms are not fully clarified. A theory of post-hypoglycemic brain glycogen super-compensation, causing a delayed cerebral response in case of progressing hypoglycaemia has been suggested but not substantiated [13]. In accordance with this, Sejling et al. found a preserved glucose threshold of EEG changes during repeated hypoglycaemia events [21]. Still we cannot rule out that EEG changes in some occasions may occur later in patients with recurrent events of hypoglycaemia. The present concept involves an operation. Although this is a minor procedure and we have not experienced any per-operative complications so far, this might be a hurdle for some patients. Finally the adherence to therapy over a prolonged period of time is to be tested in future trials.

On the other hand, there will also be advantages of the concept described. As opposed to continuous glucose monitoring the invasive procedure is a once-only event. The mounting of the external device is straight forward, and the size is identical to a modern hearing aid, and thus likely to be acceptable for most patients. Since the alarm threshold reflects the brain's reaction to hypoglycaemia, this can be individualized in order to achieve the highest possible sensitivity. The sensitivity and positive predictive value of alarms are yet to be determined in larger clinical studies. However, from the data available from our experiments, we have calculated a sensitivity of approximately 80 % with the expense of two false daytime alarms and one false nighttime alarm per week.

8 Conclusion

Although preliminary, our data support that an EEG-based alarm prediction of impending severe hypoglycaemia might be an option for diabetes patients with hypoglycaemia unawareness. Longer duration studies are needed to give information about the sensitivity and positive predictive value of the device and to achieve knowledge about patient satisfaction and usability issues of using the device.

References

1. Banck-Petersen, P., Larsen, T., Pedersen-Bjergaard, U., Bie-Olsen, L., Hoi-Hansen, T., Thorsteinsson, B.: Concerns about hypoglycaemia and late complications in patients with insulin-treated diabetes. Eur. Diabetes Nurs. **4**(3), 113–118 (2007). doi:10.1002/edn.91
2. Berger, H.: Über das Elektrenkephalogramm des Menschen. Archiv für Psychiatrie und Nervenkrankheiten **87**(1), 527–570 (1929). doi:10.1007/BF01797193
3. Boudin, G., Remond, A., Castaigne, P.: Electroencephalographic study of spontaneous hypoglycemia due to adenoma of the islands of Langerhans; considerations on the electroencephalographic modifications of hypoglycemia. Rev. Neurol. (Paris) **84**(2), 171–174 (1951)
4. Cryer, P.E.: Mechanisms of hypoglycemia-associated autonomic failure in diabetes. N. Engl. J. Med. **369**(4), 362–372 (2013)

5. Elsborg, R., Remvig, L., Beck-Nielsen, H., Juhl, C.: Detecting hypoglycemia by using the brain as a biosensor, vol. 2. Biosensors for Health, Environment and Biosecurity (2011)
6. Iber, C., Ancoli-Israel, S., Chesson Jr, A.L., Quan, S.: The AASM Manual for the Scoring of Sleep and Associated Events. American Academy of Sleep Medicine, Westchester (2007)
7. Joachims, T.: Making large-scale SVM learning practical. In: Schlkopf, B., Burges, C., Smola, A. (eds.) Advances in Kernel Methods - Support Vector Learning. MIT Press, Cambridge (1999)
8. Juhl, C.B., Hojlund, K., Elsborg, R., Poulsen, M.K., Selmar, P.E., Holst, J.J., Christiansen, C., Beck-Nielsen, H.: Automated detection of hypoglycemia-induced EEG changes recorded by subcutaneous electrodes in subjects with type 1 diabetes-the brain as a biosensor. Diabetes Res. Clin. Pract. **88**(1), 22–28 (2010)
9. Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group: prolonged nocturnal hypoglycemia is common during 12 months of continuous glucose monitoring in children and adults with type 1 diabetes. **33**(5), 1004–1008 (2010)
10. Ly, T.T., Nicholas, J.A., Retterath, A., Lim, E.M., Davis, E.A., Jones, T.W.: Effect of sensor-augmented insulin pump therapy and automated insulin suspension vs standard insulin pump therapy on hypoglycemia in patients with type 1 diabetes: a randomized clinical trial. JAMA **310**(12), 1240–1247 (2013)
11. Maahs, D.M., Calhoun, P., Buckingham, B.A., Chase, H.P., Hramiak, I., Lum, J., Cameron, F., Bequette, B.W., Aye, T., Paul, T., Slover, R., Wadwa, R.P., Wilson, D.M., Kollman, C., Beck, R.W.: A randomized trial of a home system to reduce nocturnal hypoglycemia in type 1 diabetes. Diabetes Care **37**(7), 1885–1891 (2014)
12. Nguyen, H.T., Ghevondian, N., Jones, T.W.: Real-time detection of nocturnal hypoglycemic episodes using a novel non-invasive hypoglycemia monitor. Conf. Proc. IEEE Eng. Med. Biol. Soc. **2009**, 3822–3825 (2009)
13. Oz, G., Tesfaye, N., Kumar, A., Deelchand, D.K., Eberly, L.E., Seaquist, E.R.: Brain glycogen content and metabolism in subjects with type 1 diabetes and hypoglycemia unawareness. J. Cereb. Blood Flow Metab. **32**(2), 256–263 (2012)
14. Pedersen-Bjergaard, U., Kristensen, P.L., Beck-Nielsen, H., Norgaard, K., Perrild, H., Christiansen, J.S., Jensen, T., Hougaard, P., Parving, H.H., Thorsteinsson, B., Tarnow, L.: Effect of insulin analogues on risk of severe hypoglycaemia in patients with type 1 diabetes prone to recurrent severe hypoglycaemia (HypoAna trial): a prospective, randomised, open-label, blinded-endpoint crossover trial. Lancet Diabetes Endocrinol **2**(7), 553–561 (2014)
15. Plank, J., Kohler, G., Rakovac, I., Semlitsch, B.M., Horvath, K., Bock, G., Kraly, B., Pieber, T.R.: Long-term evaluation of a structured outpatient education programme for intensified insulin therapy in patients with type 1 diabetes: a 12-year follow-up. Diabetologia **47**(8), 1370–1375 (2004)
16. Pramming, S., Thorsteinsson, B., Stigsby, B., Binder, C.: Glycaemic threshold for changes in electroencephalograms during hypoglycaemia in patients with insulin dependent diabetes. Br. Med. J. (Clin Res Ed) **296**(6623), 665–667 (1988)
17. Ramchandani, N., Arya, S., Ten, S., Bhandari, S.: Real-life utilization of real-time continuous glucose monitoring: the complete picture. J. Diabetes Sci. Technol. **5**(4), 860–870 (2011)
18. Remvig, L.S., Elsborg, R., Sejling, A.S., Sorensen, J.A., Sonder Snogdal, L., Folkestad, L., Juhl, C.B.: Hypoglycemia-related electroencephalogram changes are independent of gender, age, duration of diabetes, and awareness status in type 1 diabetes. J. Diabetes. Sci. Technol. **6**(6), 1337–1344 (2012)
19. Ross, I.S., Loeser, L.H.: Electroencephalographic findings in essential hypoglycemia. Electroencephalogr. Clin. Neurophysiol. **3**(2), 141–148 (1951). doi:10.1016/0013-4694(51)90003-X
20. Sejling, A.S., Kjaer, T.W., Pedersen-Bjergaard, U.: S., R.L., Duun-Henriksen, J., Tarnow, L., Faber, J., Holst, J.J., Hilsted, L., Thorsteinsson, B., Juhl, C.B.: Cerebral electrical activity during recurrent hypoglycemia in patients with type 1 diabetes and normal hypoglycemia awareness or hypoglycemia unawareness (abstract). Diabetes **63**, A66 (2014)

21. Sejling, A.S., Kjær, T.W., Pedersen-Bjergaard, U., Diemar, S.S., Frandsen, C.S., Hilsted, L., Faber, J., Holst, J.J., Tarnow, L., Nielsen, M.N., Remvig, L.S., Thorsteinsson, B., Juhl, C.B.: Hypoglycemia-associated changes in the electroencephalogram in patients with type 1 diabetes and normal hypoglycemia awareness or unawareness. Diabetes **64**(5), 1760–1769 (2015)
22. Sigurdsson, S., Larsen, J., Hansen, L., Philipsen, P., Wulf, H.: Outlier estimation and detection application to skin lesion classification. In: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 1, pp. I–1049–I–1052 (2002). doi:10.1109/ICASSP.2002.5743975
23. Snogdal, L.S., Folkestad, L., Elsborg, R., Remvig, L.S., Beck-Nielsen, H., Thorsteinsson, B., Jennum, P., Gjerstad, M., Juhl, C.B.: Detection of hypoglycemia associated EEG changes during sleep in type 1 diabetes mellitus. Diabetes Res. Clin. Pract. **98**(1), 91–97 (2012)
24. Tamborlane, W.V., Beck, R.W., Bode, B.W., Buckingham, B., Chase, H.P., Clemons, R., Fiallo-Scharer, R., Fox, L.A., Gilliam, L.K., Hirsch, I.B., Huang, E.S., Kollman, C., Kowalski, A.J., Laffel, L., Lawrence, J.M., Lee, J., Mauras, N., O'Grady, M., Ruedy, K.J., Tansey, M., Tsalikian, E., Weinzimer, S., Wilson, D.M., Wolpert, H., Wysocki, T., Xing, D., Chase, H.P., Fiallo-Scharer, R., Messer, L., Gage, V., Burdick, P., Laffel, L., Milaszewski, K., Pratt, K., Bismuth, E., Keady, J., Lawlor, M., Buckingham, B., Wilson, D.M., Block, J., Benassi, K., Tsalikian, E., Tansey, M., Kucera, D., Coffey, J., Cabbage, J., Wolpert, H., Shetty, G., Atakov-Castillo, A., Giusti, J., O'Donnell, S., Ghiloni, S., Hirsch, I.B., Gilliam, L.K., Fitzpatrick, K., Khakpour, D., Wysocki, T., Fox, L.A., Mauras, N., Englert, K., Permuy, J., Bode, B.W., O'Neil, K., Tolbert, L., Lawrence, J.M., Clemons, R., Maeva, M., Sattler, B., Weinzimer, S., Tamborlane, W.V., Ives, B., Bosson-Heenan, J., Beck, R.W., Ruedy, K.J., Kollman, C., Xing, D., Jackson, J., Steffes, M., Bucksa, J.M., Nowicki, M.L., Van Hale, C., Makky, V., O'Grady, M., Huang, E., Basu, A., Meltzer, D.O., Zhao, L., Lee, J., Kowalski, A.J., Laffel, L., Tamborlane, W.V., Beck, R.W., Kowalski, A.J., Ruedy, K.J., Weinstock, R.S., Anderson, B.J., Kruger, D., LaVange, L., Rodriguez, H.: Continuous glucose monitoring and intensive treatment of type 1 diabetes. N. Engl. J. Med. **359**(14), 1464–1476 (2008)

Meta-Learning Based Blood Glucose Predictor for Diabetic Smartphone App

Valeriya Naumova, Lucian Nita, Jens Ulrik Poulsen and Sergei V. Pereverzyev

Abstract The obvious and highly accepted convenience of smartphone apps will, already in the nearest future, bring new opportunities for diabetes therapy management. In particular, it is expected that smartphones will be able to read, store, and display the blood glucose concentration from the continuous glucose monitoring systems. Using our knowledge and experience gained in the framework of the large-scale European Union FP7 funded project "DIAdvisor: personal glucose predictive diabetes advisor" (2008–2012), we explore a possibility to develop a novel smartphone app for diabetes patients that provides estimations of the future blood glucose concentration from current and past blood glucose readings. In addition to reliable clinical accuracy, a prediction algorithm implemented in such an app should satisfy multiple requirements, such as easily and quickly implementable on any mobile operating system, portability from individual to individual without readjustment or retraining procedure, and a low battery usage feature. In this study, we present a description of the prediction algorithm, developed in the course of the DIAdvisor project, and its version on Android OS that meets the above-mentioned requirements. Additionally, we compare the clinical accuracy of the algorithm with the state of the art in terms of the "gold standard" metric, Clarke error grid analysis, and the recently introduced metric, prediction error grid analysis.

V. Naumova (✉)
Center for Biomedical Computing, Simula Research Laboratory, Martin Linges vei 25, Fornebu, Norway
e-mail: valeriya@simula.no

L. Nita
RomSoft SRL, Chimiei 2 Bis Excel Business Center Iasi, Iasi, Romania

J.U. Poulsen
Novo Nordisk A/S, Tuborg Havnevej 19, Hellerup, Denmark

S.V. Pereverzyev
Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, AltenbergerstraßE 69, Linz, Austria

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*, Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_6

1 Introduction

In recent years, mobile applications have become a key driver of mobile Health (mHealth) deployment, esp. as a complementary way for self-monitoring. It is foreseen that the market of mHealth app users will reach 2 billion by 2017 [7].

Rapid advances in glucose monitoring devices, wearable devices like smartwatches with dedicated data-gathering apps, and predictive modeling algorithms, combined with the ongoing explosive development and availability of modern in nutritional and lifestyle data analysis tools, make it possible to provide people with diabetes with a dizzying array of information on which to base their decisions. Since blood glucose fluctuations can be very rapid and drastic, the strict glycemic control requires much effort, round-the-clock monitoring and self-motivation. At the same time, the increased usage of mobile apps and patient portals to engage people with diabetes provide researchers with access to a plethora of useful data for analysis. For all these reasons, diabetes management is well placed for taking advantage of recent technological innovations in data collection and analysis towards better self-management.

Currently more than 1,100 diabetes-related mobile applications (apps) are available for iOS and Android devices. However, only 1.2 % of the diabetic smartphone owners use an app to manage their condition. The reasons for such a low participation rate are the high dependence on manual entries, lack of integrity and predictive value, lack of personalization and feedback, and low user engagement [14].

Despite the criticism and low patient involvement, recent studies have shown that diabetes-related mobile applications possibly combined with telemedicine support could lead to significant improvement in HbA1c levels of type-1 diabetes patients [4]. In addition, it is expected that quite soon smartphones will be able to receive, read, and store measurements from continuous glucose monitoring (CGM) systems, which is of paramount importance for a better diabetes therapy management. In short, the CGM systems provide ongoing monitoring of glucose level on an automated basis throughout the day and night, blood glucose estimation every 5 or 10 (min), depending on a system manufacturer.

In view of this, the partners of the large-scale EU FP7-funded project "DIAdvisor: personal glucose predictive diabetes advisor" (2008–2012) [6] explore an idea of producing a new smartphone app that predicts the future blood glucose (BG) concentration of a diabetic patient from current and past BG measurements provided by CGM.

As briefly mentioned above, in order to ensure long-term usability of a diabetes app, several features have to be fulfilled. In the following, we discuss some of more technical features, which are necessary for the viability of a smartphone app with a prediction algorithm. First, its workload should not lock the smartphone from handling other tasks. Second, a prediction algorithm should be portable from individual to individual without being readjusted or retrained, such that no manual input is required after the app has been downloaded. Such feature is important for attracting potential users and customers. Hanna Mählen, Global Digital Marketing Consult at

Roche Diagnostic, US, for instance, indicated that the usability of the diabetes apps can be improved by eradicating needs for manual entries.

Finally, the third and the most important requirement is that a clinical accuracy provided by the BG-prediction app should be at least at the level of the state-of-the-art algorithms for the blood glucose prediction or even at the level of a CGM system. Note that in the majority of the literature on the blood glucose prediction, the clinical accuracy is measured in terms of the Clarke error grid analysis (EGA) [5], which is accepted as one of the "gold standards" for determining the accuracy of blood glucose meters. Another assessment metric, we consider in the paper, is the prediction error grid analysis (PRED-EGA) [15] that has been designed specifically for the assessment of blood glucose predictors.

Thus, the third requirement can mean that in terms of the above-mentioned metrics a prediction algorithm implemented in the app should perform at the level of the best blood glucose predictors reported in the literature and at the level of accuracy guaranteed by a CGM device, which means more than 95 % of accurate and benign predictions (zones A+B in terms of EGA) [16]. Our thorough investigations over the last years [13] have shown that the fully adaptive regularized learning (FARL) algorithm [9] is the only one among a variety of the known algorithms that potentially can meet all the above-mentioned requirements.

In [9], it has been demonstrated that the FARL algorithm can also be trained to predict the BG concentration from the extended inputs that include not only current and past BG measurements, but also information about meals and insulin administration. However, in order to lighten the burden of tedious manual input, this feature of the algorithm (i.e., extension to the additional input) could be neglected at the current stage of the development.

The study [12] demonstrates the ability and the functionality of the FARL algorithm to predict simultaneously the BG concentration for a variety of the prediction horizons (PH). It is also shown in [12] and in Fig. 2 below, the algorithm outperforms the existing state-of-the-art algorithms. On the other hand, keeping in mind the onset time of insulin and meal responses on the BG level, it is sufficient to know the BG concentration at least 20 min ahead of time to take preventive measures against out-of-range glucose excursions. This hints that PH = 20 (min) is of the main importance for the discussed application.

The scope of the next sections is to show that the FARL predictor, presented in more details in Sect. 2, could be effectively built and run well on Android devices. In Sect. 4 we report the results of the predictors performance assessment and compare them with the ones known from the literature. The results presented in the paper support the idea of the FARL predictor implementation as a diabetes app.

2 Fully Adaptive Regularized Learning Algorithm for the Blood Glucose Prediction

At first, we introduce some notations to be frequently used throughout this section. In the sequel, we denote a generic function by g that may approximate the evolution of the BG concentration in time; $g(t_i)$ denotes a value of the function g at time $t = t_i$; $g_{\lambda,\omega}(t)$ denotes a function of time t that also depends on parameters $\lambda \in \mathbf{R}$ and $\omega = (\omega_1, \omega_2, \omega_3) \in \mathbf{R}^3$; g_i denotes a value of the patient's BG concentration measured at time $t = t_i$. Moreover, letters in boldface are used to denote vectors or matrices with the corresponding components.

The FARL algorithm extrapolates m preceding values $g_0, g_{-1}, g_{-2}, \ldots, g_{-m+1}$ of the patient's BG concentration sampled correspondingly at the time moments $t_0 > t_{-1} > t_{-2} > \cdots > t_{-m+1}$ within a given sampling horizon $SH = t_0 - t_{-m+1}$ to forecast the evolution of the BG concentration for the future time period $[t_0, t_0 + PH]$, where PH is a given prediction horizon. The forecast is produced as a function of time $g_{\lambda,\omega}(t)$ at which the following penalized least squares functional

$$m^{-1} \sum_{i=-m+1}^{0} (g_i - g(t_i))^2 + \lambda \|g - g^{(0)}\|^2_{\mathscr{H}_K}, \ \lambda \in (0, 1), \tag{1}$$

attains its minimum over all functions g from a reproducing kernel Hilbert space (RKHS) $\mathscr{H}_K$, generated by the kernel

$$K(t, \tau) = K_\omega(t, \tau) = (t\tau)^{\omega_1} + \omega_2 e^{-\omega_3 (t-\tau)^2}. \tag{2}$$

Here $g^{(0)} = g^{(0)}(t)$ is an initial approximation that can be chosen in a similar manner as the minimizer of the functional

$$m^{-1} \sum_{i=-m+1}^{0} (g_i - g(t_i))^2 + \lambda \|g\|^2_{\mathscr{H}_K}. \tag{3}$$

Note that due to the representer theorem [17] for a given λ and ω the function $g_{\lambda,\omega}(t)$ has the form

$$g_{\lambda,\omega}(t) = \sum_{i=-m+1}^{0} c_i^\lambda K_\omega(t, t_i), \tag{4}$$

where a real vector $\mathbf{c}_\lambda = (c_0^\lambda, c_{-1}^\lambda, \ldots, c_{-m+1}^\lambda)$ of coefficients is defined as follows

$$\mathbf{c}_\lambda = (\lambda m \mathbf{I} + \mathbf{K})^{-1} (\mathbf{g} - \mathbf{g}^{(0)}). \tag{5}$$

Here $\mathbf{I}$ is the unit matrix of the size $m \times m$, $\mathbf{K} = \{K(t_i, t_j), \ i, j = 0, -1, \ldots, -m + 1\}$ is the so-called Gram matrix, and $\mathbf{g} = (g_0, g_{-1}, \ldots, g_{-m+1})$,

$\mathbf{g}^{(0)} = (g^{(0)}(t_0), g^{(0)}(t_{-1}), \ldots, g^{(0)}(t_{-m+1}))$, and

$$g_i^{(0)} = \sum_{j=-m+1}^{0} c_{j,0}^{\lambda} K_{\omega}(t_i, t_j), \tag{6}$$

with $(c_{0,0}^{\lambda}, c_{-1,0}^{\lambda}, \ldots, c_{-m+1,0}^{\lambda}) = (\lambda m \mathbf{I} + \mathbf{K})^{-1}\mathbf{g}$.

Of course, the penalized least squares method (1), (3) has been used for decades for reconstructing functions from given (possibly noisy) data. *The main novelty of the FARL algorithm* compared to what was previously known (the state-of-the-art) is that the space $\mathscr{H}_K$ is not a priori fixed but *chosen adaptively depending on the input data* $(t_i, g_i)_{i=-m+1}^{0}$. More precisely, in the FARL algorithm the kernel parameters $\omega_1, \omega_2, \omega_3$, as well as the regularization parameter λ, are the functions $\omega_k = \omega_k(u),\ k = 1, 2, 3,\ \lambda = \lambda(u)$ of the vector $u = (u_1, u_2)$ of coefficients of the linear least squares fit $g_{\text{lin}}(t) = u_1 t + u_2,\ t \in [t_{-m+1}, t_0]$, to the data $(t_i, g_i)_{i=-m+1}^{0}$. These coefficients determined as follows

$$u_1 = \frac{\sum_{i=-m+1}^{0}(t_i - \bar{t})(g_i - \bar{g})}{\sum_{i=-m+1}^{0}(t_i - \bar{t})^2}, \qquad u_2 = \bar{g} - u_1 \bar{t}, \tag{7}$$

where $\bar{a}$ denotes the average of components of a vector $(a_0, a_{-1}, \ldots, a_{-m+1})$.

To find the functions $\omega_k(u),\ k = 1, 2, 3,\ \lambda(u)$ that return the values of the kernel and the regularization parameters, we use the concept of *meta-learning*, which presupposes that parameters of an algorithm are selected on the base of the previous experience with similar input data.

At this point, it is worthwhile mentioning that CGM systems provide estimations of the BG concentration every 5 or 10 min, depending on a system manufacturer, which allows us to form a training set $(x_\mu, y_\mu),\ \mu = 1, 2, \ldots, M$, of data segments

$$x_\mu = \{(t_{-m+1}^{\mu}, g_{-m+1}^{\mu}), \ldots, (t_0^{\mu}, g_0^{\mu})\}, \tag{8}$$

$$y_\mu = \{(t_1^{\mu}, g_1^{\mu}), \ldots, (t_n^{\mu}, g_n^{\mu})\}, \tag{9}$$

where $t_{-m+1}^{\mu} < t_{-m+2}^{\mu} < \cdots < t_0^{\mu} < t_1^{\mu} < \cdots < t_n^{\mu}$ are the moments of time at which the patient's BG concentration were estimated by the CGM system as $g_{-m+1}^{\mu}, g_{-m+2}^{\mu}, \ldots, g_0^{\mu}, g_1^{\mu}, \ldots, g_n^{\mu}$. Moreover, the moments $\{t_i^{\mu}\}_{i=-m+1}^{n}$ can be chosen such that the data segments x_μ and y_μ are, respectively, similar to the prediction input $x = \{(t_{-m+1}, g_{-m+1}), \ldots, (t_0, g_0)\}$ and the desired prediction output $y = \{(t_1, g_1), \ldots, (t_n, g_n)\}$ in the sense that for each $\mu = 1, 2, \ldots, M,\ t_0^{\mu} - t_{-m+1}^{\mu} = SH$ and $t_n^{\mu} - t_0^{\mu} = PH$. Then using formulas (1), (3) with $g_i = g_i^{\mu},\ i = -m+1, \ldots, 0$, we obtain the function $g_{\lambda,\omega}(t)$ and can tune the parameters $\lambda, \omega_1, \omega_2, \omega_3$ according to the procedure

that has been described, theoretically justified and exemplified in [8, Section 4.4], [10, 11]. Eventually, the function $g_{\lambda,\omega}(t)$ with tuned parameters λ, ω has values $g_{\lambda,\omega}(t_i^\mu)$ that are close to $g_i^\mu,\ i = 1, 2, \ldots, n$.

As the result, each input data segment x_μ is associated with the values of the kernel $\omega^\mu = (\omega_1^\mu, \omega_2^\mu, \omega_3^\mu)$ and the regularization λ^μ parameters, which allow an accurate extrapolation of data y^μ from x^μ by means of $g_{\lambda^\mu,\omega^\mu}(t)$. On the other hand, each input data segment x_μ is also associated with the vector $u^\mu = (u_1^\mu, u_2^\mu)$ of coefficients of the linear least squares fit calculated according to (7) for $t_i = t_i^\mu,\ g_i = g_i^\mu$.

Then the functions $\omega_k(u),\ k = 1, 2, 3,\ \lambda(u)$, which provide the kernel and the regularization parameters are defined by the values $\omega_k^\mu,\ k = 1, 2, 3,\ \lambda^\mu$ at the point $u = u^\mu$. In particular, the functions we are looking for can be reconstructed from these values by means of a penalized regularization performed in RKHS. A procedure on the kernels choice for such regularization is provided in detail in [8, Chap. 5], [9, 12].

Figuratively speaking, the FARL algorithm chooses a suitable kernel for performing a prediction from the given input data in a similar fashion as a physician analyzes symptoms to make a diagnosis (see Fig. 1). In this analogy, the role of the symptoms is played by the feature vector $u = (u_1, u_2)$, while ω is seen as the diagnosis. Moreover, like a physician, the algorithm has been trained on the cases where the diagnoses ω^μ for the symptoms u^μ were known.

It is important to note that in the version of the FARL algorithm, which was tested in the extended clinical trials, the functions calculating the values of the kernel and the regularization parameters were found using the data segments (x_μ, y_μ) from just one patient. Then the obtained fully trained BG predictors were tested without any readjustment on data that were collected during a 3-day inpatient stay and 5 days outside the hospitals from 90 other patients.

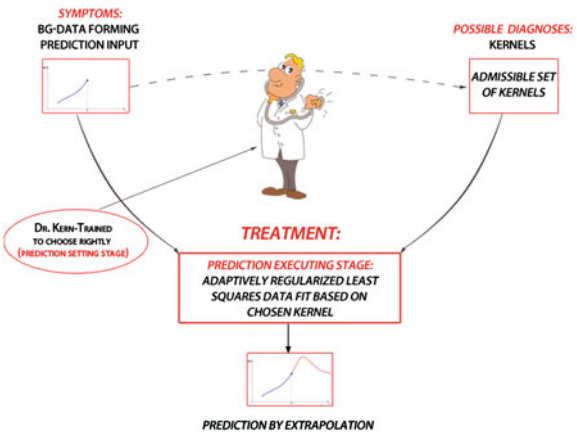

Fig. 1 A cartoon analog of the FARL algorithm for performing the blood glucose predictions

3 Android Version of the FARL Algorithm

3.1 Translation of the Algorithm from Matlab to Android System

In this section, we provide a short description of the translation process of the FARL algorithm presented in the previous section from Matlab environment to Android. Android mobile operating system was a system of choice due to several important and attractive features. According to the recent surveys, Android has the largest installed base of any mobile OS and as of 2013, its devices also sell more than Windows, iOS, and Mac OS devices all together. As of July 2013, the Google Play store has had over 1 million Android apps published, and over 50 billion apps downloaded [2]. Last but not least, Android's source code is released by Google under open source licenses.

Even though the Matlab version of the algorithm does not require any specific toolboxes, the operations as matrix inversion with high precision, which is, for instance, needed to find the vector of coefficients $\mathbf{c}_\lambda$ for the minimizer (4) cannot be performed using the standard Java libraries. However, the mathematical libraries such as Apache Commons Math [3] allow to achieve the same precision by mathematical operations as in Matlab. Therefore, the FARL algorithm was effectively and easily translated into Java compatible with the Android operating system.

The Android version of the predictor was built on a new Android development environment, Android Studio [1], and includes a set of classes which implement the algorithm and the data structures with some imported libraries for special operations: matrix calculus, time stamps, etc.

In the next section, we provide an assessment of the performance of the blood glucose predictors based on the FARL algorithm of the Matlab and Android versions using the clinically accepted metrics. However, as a short but important remark, we would like to mention that the results of the simple pointwise comparison between Matlab and Android versions are quite promising.

In particular, we consider the data of ten patients randomly chosen from the database, established in the course of the DIAdvisor project, with more than 300 data entries for more than 90 patients. The assessment period is up to 10 days for each patient and for all ten patients the predictor ran for more than 20,000 times. The values obtained at these points then were used to compare both versions of the FARL algorithm. Only 38 entries among all these collected points were different and the biggest difference was 9 mg/dl. This difference can be explained by the usage of the third-party libraries [3] for performing mathematical calculations. At the same time, the observed differences are below the minimum error of 15 mg/dL required by the international standard ISO 15197 for blood glucose monitoring systems performance.

3.2 Microprocessor and Power Consumption Analysis

Another important issue which has to be extensively tested is the consumption of memory access and, definitely, power consumption in a smartphone. In order to test the predictor's implementation on a real smartphone, we performed several tests to identify these parameters. Even in the worst case when the predictor is running nonstop (for each 4 s), the power consumption made by the application is acceptable (approx. 4 %). But in the real life, the predictor has to run once at 5 or even 10 min depending on the CGM system and, thus, the power consumption decreases accordingly.

The processor time is mainly used for performing prediction and their graphical representation. Other operations as CGM data receiving are rare in time and do not consume much processor time.

From these findings, we conclude that the predictor can run on Android systems with no big impact on the smartphone battery or performance.

4 Performance Assessment

The present section contains the summary results of the performance assessment of the Android version of the FARL predictor as well as provides the comparison with the state of the art. Since the aim of this paper is to illustrate a possibility to use the predictor as a smartphone app, we do not provide here an extensive comparison of the performance assessment but rather refer the reader to the recent papers [9, 11, 12], where such results are reported in detail. However, following the constant developments, we still compare the FARL algorithm with the predictors recently appeared in the literature as in [18], for instance.

4.1 Clinical Accuracy Metrics

The analysis of the algorithms for diabetes therapy management raises the natural problem of their assessment. The use of the proper assessment metrics helps to understand advantages of the considered algorithms. Of course, the key point is the choice of the assessment metrics that are suitable for diabetes technology.

There are several assessment metrics adopted in the literature specifically for the problem of the blood glucose readings and predictions; however, the vast majority of the prediction performance assessment is still performed using the general measures such as root mean square error, mean squared error, for example. We find the latter once, however, unillustrative and even irrelevant for a proper clinical assessment of the algorithm. We refer the reader to [15] for more details.

We consider two metrics for the assessment of the predictor's performance, which were validated and accepted by clinicians: the Clarke error grid analysis (EGA) [5] and the prediction error grid analysis (PRED-EGA) [15].

The Clarke error grid analysis [5] was developed in 1987 to quantify clinical accuracy of patient estimates of their current blood glucose as compared to the blood glucose value obtained in their meter. It was then used to quantify the clinical accuracy of blood glucose estimates generated by meters as compared to a reference value. Eventually, EGA became accepted as one of the "gold standards" for determining the accuracy of BG meters. Zone A of EGA corresponds to the clinically accurate BG-estimations. A schematic illustration of an EGA is given in Fig. 2.

The prediction error grid analysis [15] has been designed especially for the BG-prediction assessment. It calculates combined accuracy in three clinically relevant regions hypoglycaemia (low blood glucose concentration), euglycemia, and hyperglycaemia (high blood glucose concentration), and provides three estimates of the predictor's performance in each of the three regions: Accurate, Benign, and Erroneous (Error).

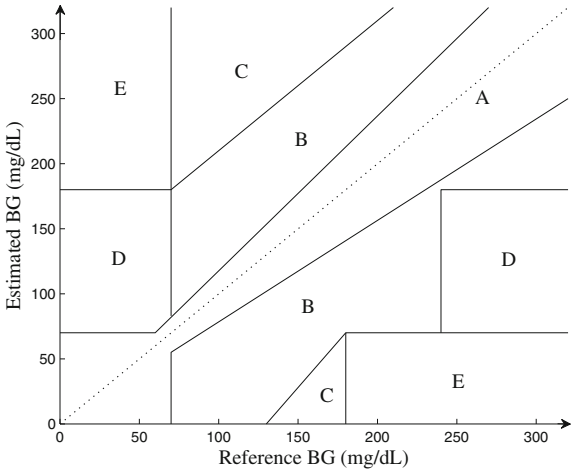

Fig. 2 The Clarke error grid analysis uses a Cartesian diagram, in which the estimated/predicted values are displayed on the y-axis, whereas the reference values are presented on the x-axis. This diagram is subdivided into 5 zones: A, B, C, D, and E. The points that fall within zones A and B represent, respectively, sufficiently accurate and acceptable glucose results, points in zone C may prompt unnecessary corrections, points in zones D and E represent erroneous and incorrect treatment

4.2 Performance Assessment

In Fig. 3, we present the comparison of the average percentage values of the clinically accurate predictions (zone A of EGA) with $PH = 15, 20, 25, 30, 35, 40, 45$ (min) for 10 patients, where the Android version of the FARL algorithm is compared against the predicted reported in the literature [18]. The assessment period for the predictors reported in the literature is 30 h; whereas for the FARL predictor it is 33 h on average.

4.3 Comparison of the Matlab and Android Versions

As mentioned above, we also compare the Matlab and Android versions of the predictor against the CGM measurements for different prediction horizons, i.e., $PH = 20, 30, 40, 60, 75, 120$ (min). In this case, we use the EGA and the PRED-EGA as assessment metrics. The results displayed in Tables 1, 2, and 3 clearly illustrate that both versions perform similarly and only very slight differences in terms of the assessment metrics might be observed. For the sake of brevity, we report

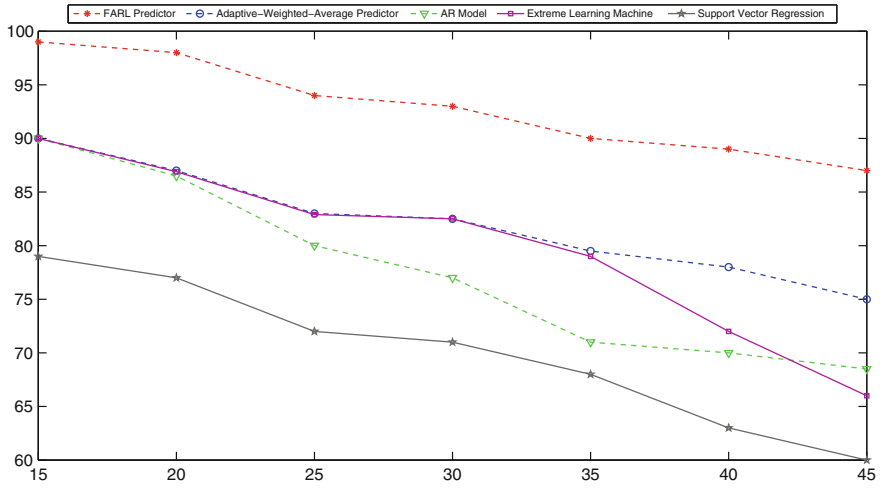

Fig. 3 The comparison of the average percentage values of the clinically accurate predictions (zone A of EGA) for different prediction horizons of the five different predictors

Table 1 Percentage of points in EGA zones for the FARL predictor, Android and Matlab versions, $PH = 20$ (min)

ID	Matlab Version					Android Version				
	A %	B %	C %	D %	E %	A %	B %	C %	D %	E %
101	97.69	1.96	0.00	0.20	0.00	97.69	1.96	0.00	0.20	0.00
102	97.43	2.47	0.00	0.10	0.00	97.49	2.41	0.00	0.10	0.00
103	95.25	4.58	0.00	0.17	0.00	95.22	4.61	0.00	0.17	0.00
108	96.09	3.88	0.00	0.03	0.00	96.09	3.88	0.00	0.03	0.00
120	99.27	0.70	0.03	0.00	0.00	99.27	0.70	0.03	0.00	0.00
121	96.79	3.04	0.00	0.17	0.00	96.89	2.94	0.00	0.17	0.00
128	98.52	1.48	0.00	0.00	0.00	98.56	1.44	0.00	0.00	0.00
206	99.00	0.85	0.00	0.15	0.00	99.04	0.81	0.00	0.15	0.00
212	98.20	1.80	0.00	0.00	0.00	98.27	1.73	0.00	0.00	0.00
213	99.26	0.59	0.15	0.00	0.00	99.26	0.59	0.15	0.00	0.00
Average	97.75	2.14	0.02	0.08	0.00	97.78	2.11	0.02	0.08	0.00

Table 2 PRED-EGA for 20-min predictions by the Matlab version of the FARL Algorithm

Subject ID	Hypoglycaemia			Euglycemia			Hyperglycaemia		
	Accurate %	Benign %	Error %	Accurate %	Benign %	Error %	Accurate %	Benign %	Error %
101	94.05	1.08	4.86	96.96	2.98	0.06	98.18	1.09	0.73
102	90.62	0.00	9.38	98.75	1.25	0.00	96.63	2.45	0.92
103	88.06	4.48	7.46	96.11	3.78	0.11	96.95	1.53	1.53
108	98.96	0.52	0.52	98.49	1.51	0.00	100.00	0.00	0.00
120	100.00	0.00	0.00	99.16	0.78	0.06	98.50	0.83	0.67
121	87.27	3.64	9.09	97.58	2.42	0.00	97.47	1.54	0.99
128	100.00	0.00	0.00	98.82	1.18	0.00	97.03	0.56	2.41
206	78.95	0.00	21.05	98.16	1.79	0.00	97.68	0.64	1.67
212	100.00	0.00	0.00	98.20	1.80	0.00	99.16	0.28	0.56
213	–	–	–	97.14	1.43	1.43	99.00	0.50	0.50
Average	93.10	1.08	5.82	97.94	1.89	0.17	98.06	0.94	1.00

here only the results of the performance assessment (EGA and PRED-EGA) for $PH = 20$ (min) for both versions of the predictor. We also mention that these results are representative for other PH as well.

Table 3 PRED-EGA for 20-min predictions by the Android version of the FARL Algorithm

Subject ID	Hypoglycaemia			Euglycemia			Hyperglycaemia		
	Accurate %	Benign %	Error %	Accurate %	Benign %	Error %	Accurate %	Benign %	Error %
101	94.05	1.08	4.86	97.09	2.85	0.06	98.18	0.97	0.85
102	90.62	0.00	9.38	98.75	1.25	0.00	96.93	2.15	0.92
103	88.06	4.48	7.46	96.22	3.67	0.11	96.95	1.53	1.53
108	98.96	0.52	0.52	98.53	1.47	0.00	100.00	0.00	0.00
120	100.00	0.00	0.00	99.11	0.84	0.06	98.50	0.83	0.67
121	87.27	3.64	9.09	97.58	2.37	0.05	97.36	1.76	0.88
128	100.00	0.00	0.00	98.82	1.18	0.00	97.03	0.37	2.60
206	78.95	0.00	21.05	98.10	1.84	0.06	97.81	0.51	1.67
212	100.00	0.00	0.00	98.29	1.71	0.00	98.88	0.56	0.56
213	–	–	–	97.14	1.43	1.43	99.00	0.50	0.50
Average	93.10	1.08	5.82	7.96	1.86	0.18	98.06	0.92	1.02

5 Conclusions and Discussion

In this paper, we presented a technical and algorithmical framework of the diabetes app for short-term predictions of the BG evolution of a diabetes patient. The developed solution demonstrates several attractive features, such as robustness, high clinical accuracy, patient and sensor independence, no need of manual input, which are of great importance for future potential users.

Despite these promising results, we definitely interested in continuing improvement and development of the solution for the integrated diabetes self-management. In particular, our future ambitions include identification and formalization of user preferences and needs for mHealth solutions capable of delivering efficient and realistic long-term diabetes management; establishing connectivity with external sensors and valuable external applications (e.g., for diet tracking); and extension of the algorithm to a longer prediction horizons. The latter one will require change of paradigms, e.g., we are interested in producing BG estimates with lowest possible uncertainty over the prediction horizon, and incorporation of additional inputs such as meal intake and insulin injection due to the intrapatient variability. As already mentioned, a short-term predictor is less concerned by this problem and it becomes essential once we want to increase the PH (e.g., from 1 to 3 h), because small variations of insulin action result in larger glucose deviations with time. Therefore, various possibilities and tools, which allow a reduction of the burden of tedious manual input, will have to be thoroughly studied.

Acknowledgments The first and the fourth authors acknowledge the support of the Austrian Fonds zur Förderung der Wissenschaftlichen Forschung (FWF), grant P25424 “Data-driven and problem-oriented choice of the regularization space.”

References

1. Android studio environment. http://developer.android.com
2. Android's google play beats app store with over 1 billion apps, now officially largest. http://phonearena.com. Accessed 28 Aug 2013
3. Apache commons. http://commons.apache.org/
4. Charpentier, G., et al.: The Diabeo software enabling individualised insulin dose adjustments combined with telemedicine support improves HbA1c in poorly controlled type 1 diabetic patients. Diabetes Care **34**, 533–539 (2011)
5. Clarke, W., Cox, D., Gonder-Frederick, L., Carter, W.R., Pohl, S.: Evaluating clinical accuracy of systems for self-monitoring of blood glucose. J. Diabetes Care **10**, 622–628 (1987)
6. DIAdvisor: personal glucose predictive diabetes advisor. http://www.diadvisor.eu
7. European commission: green paper on mobile health. SWD (2014) 135 final, pp. 1–20 (2014)
8. Lu, S., Pereverzyev, S.V.: Regularization theory for ill-posed problems, selected topics. DeGruyter, Berlin (2013)
9. McKennoch, S., Pereverzyev, S.V., Randlov, J., Sampath, S.: Patent WO2012143505A2: Glucose predictor based on regularization networks with adaptively chosen kernels and regularization parameters (2011). http://worldwide.espacenet.com
10. Naumova, V., Pereverzyev, S.V., Sampath, S.: Reading blood glucose from subcutaneous electric current by means of a regularization in variable reproducing Kernel Hilbert spaces. In: Proceedings of 50th IEEE CDC-ECC, pp. 5158–5163, USA (2011)
11. Naumova, V., Pereverzyev, S.V., Sampath, S.: Extrapolation in variable RKHSs with application to the blood glucose reading. Inverse Probl. **27**, 1–13 (2011)
12. Naumova, V., Pereverzyev, S.V., Sampath, S.: A meta-learning approach to the regularized learning—case study: blood glucose prediction. Neural Netw. **33**, 181–193 (2012)
13. Naumova, V., Pereverzyev, S.V.: Blood glucose predictors: an overview on how recent developments help to unlock the problem of glucose regulation. Recent Pat. Comput. Sci. **5**, 177–187 (2012)
14. Research2guidance: mHealth App Developer Economics. mHealthEconomics, pp. 1–42 (2014)
15. Sivananthan, S., Naumova, V., Dalla Man, C., Facchinetti, A., Renard, E., Cobelli, C., Pereverzyev, S.V.: Assessment of blood glucose predictors: the prediction-error grid analysis. Diabetes Technol. Ther. **13**, 787–796 (2011)
16. Vashist, S.K.: Continuous glucose monitoring systems: a review. Diagnostic **3**, 385–412 (2013)
17. Wahba, G.: Spline models for observational data (CBMS-NSF Regional Conference Series in Applied Mathematics vol. 59) SIAM Philadelphia
18. Wang, Y., Wu, X., Mo, X.: Adaptive-weighted-average framework for blood glucose prediction. Diabetes Technol. Ther. **15**, 792–801 (2013)

Predicting Glycemia in Type 1 Diabetes Mellitus with Subspace-Based Linear Multistep Predictors

Marzia Cescon, Rolf Johansson and Eric Renard

Abstract A major challenge for a person with diabetes is to adapt insulin dosage regimens and food intake to keep blood glucose within tolerable limits during daily life activities. The accurate prediction of blood glucose levels in response to inputs would support the patients with invaluable information for appropriate on-the-spot decision making concerning the management of the disease. Against this background, in this paper we propose multistep data-driven predictors to the purpose of predicting blood glucose multiple steps ahead in the future. We formulate the predictors based on the state-space construction step in subspace identification methods for linear systems. The clinical data of 14 type 1 diabetic patients collected during a 3-days long hospital visit were used. We exploited physiological models from the literature to filter the raw information on carbohydrate and insulin intakes in order to retrieve the inputs signals to the predictors. Predictions were based on the collected CGMS measurements, recalibrated against finger stick samples and smoothed through a regularization step. Performances were assessed with respect to YSI blood glucose samples and compared to those achieved with a Kalman filter identified from data. Results proved the competitiveness of the proposed approach.

1 Introduction

Diabetes mellitus is a chronic disease of disordered glucose metabolism due to a constant defect in insulin secretion by the pancreatic β-cells that can be combined with a defect in insulin action [36]. Type 1 diabetes (T1DM), previously called insulin-dependent diabetes mellitus (IDDM) is characterized by almost complete

M. Cescon · R. Johansson (✉)
Department of Automatic Control, Lund University, PO Box 118, 221 00 Lund, Sweden
e-mail: Rolf.Johansson@control.lth.se
M. Cescon
e-mail: marzia.cescon@control.lth.se

E. Renard
Department of Endocrinology, Montpellier University Hospital, University of Montpellier, Montpellier, France

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*,
Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_7

failing production of insulin, whereas type 2 diabetes is caused by reduced insulin production and decreased sensitivity of the tissues to the metabolic effect of insulin. The basic consequence of lack of insulin or insulin resistance is the prevention of the efficient uptake and utilization of glucose by insulin-dependent cells of the body, mainly the skeletal muscles, resulting in abnormally high blood sugar levels (hyperglycemia). Sustained hyperglycemia is associated with microvascular complications (nephropathy, retinopathy), neuropathy, and damages to the cardiovascular system, which are irreversible once they develop and can mean serious disability for the person who experiences them [36]. In order to avoid the long-term health complications, the patients need to regulate their blood glucose maintaining its levels as close to normal as possible. This is achieved by exogenous insulin replacement. Focusing on tight blood glucose targets, i.e., 70–140 (mg/dL) [32], the philosophy of insulin replacement is to mimic the physiological endogenous insulin secretion pattern of the nondiabetic person. In the nondiabetic subjects, insulin is secreted into the portal circulation at two rates: a slow basal secretion throughout the 24 h and an augmented rate at meal times. This pattern can be achieved to some extent with the so-called basal-bolus regime: a basal dose of long-acting insulin is sufficient to keep a constant glucose concentration during fasting conditions and a prandial bolus of rapid-acting insulin enhances an increased glucose uptake during and after meals. In current medical practice, the rough calculation of insulin doses and eventually extra carbohydrate intakes is based on empirical rules-of-thumbs. Many factors have to be considered in this decision process: health status, current blood glucose level, blood glucose target, foreseen activities, insulin sensitivity, expected future glycemia evolution, and approximation of the estimated meal carbohydrate content effects as well as insulin impact on the subject own blood glucose, taking into account medical advice and patients previous experience of his/her own metabolism. One common measure used in this regard is the carbohydrate-to-insulin ratio, which is an estimate of how many insulin units to administer to match the amount of digested carbohydrates. The task is nontrivial and demanding, although the standard tools in diabetes care improved significantly during the last decades.

Against this background, a personalized blood glucose predicting system to be used on the spot by the users in different daily situations, predicting glycemic excursions following meals, insulin intakes and exercise, would support the patients in optimizing their therapy. Actually, this was the focus of the European FP7-IST research project DIAdvisor™ [12] during the quadrennium 2008–2012. The predictor system needs user inputs concerning patient condition (e.g., fasting, meal time, rest, or physical activity), time, and size of meals and insulin doses, minimally invasive glucose sensors and produces short term, i.e., up to 120 min, blood glucose predictions to be graphically shown to the patients.

Although seemingly simple in concept, the problem of glucose prediction in an active individual has to date proved difficult. Currently, continuous glucose monitoring (CGM) devices are the available technology able to provide high/low glucose alarms when certain user specified preset threshold levels have been crossed and to deliver warnings of events that are likely to occur if the current trend continues. However, patients will benefit more from an early alarm that predicts the episodes before

they occur. To date many studies have investigated the possibility of predicting blood glucose concentration for the purpose of regulating glucose intervention, in order to enable individuals to take corrective actions and avoid low or high glucose values (see e.g., [2, 3] for a comprehensive review on T1DM modeling). Bremer and Gough in [5] originally developed the idea of T1DM CGM time series analysis using 10-min sampled data from ambulatory T1DM patients to identify autoregressive (AR) models. They explored 10, 20, and 30 min prediction horizons, and report that the 10-min predictions are accurate and the 20-min or 30-min predictions may also be acceptable for a limited set of data only. Sparacino et al. collected 48 h of continuously sampled (3-min) glucose data from 28 T1DM subjects in ambulatory conditions. In their retrospective analysis, presented in [30] they recursively identified simple polynomial and AR models from the CGM time series data, prefiltered to remove noise spikes. They investigated prediction horizons of 10 and 15 steps (i.e., 30 and 45 min) and concluded that hypoglycemia can be detected 25 min before the hypoglycemic threshold is passed. A tutorial overview of algorithms for CGM time series analysis to the purpose of alarm generation was provided in [31]. Eren-Oruklu et al. proposed in [13] a recursive second-order AR and ARMA model identification strategy with an adjustable forgetting factor for healthy and type II diabetics. Their models utilized only recent glucose history from a CGM device, achieving 3–5 % error for 30-min ahead prediction. In [28] a kernel-based regularization learning algorithm, in which the kernel and the regularization parameter are adaptively chosen on the basis of previous similar learning tasks, using past glucose concentration information, was presented. A short-coming of the methods listed above is the lack of exploitation of the dynamic interplay between previously injected insulin, meal intake, and eventually exercise to the purpose of improving glucose prediction. Patient-specific ARX models both batch-wise and recursively identified from nine patients data records by Finan et al. showed in [16] a mean 30-min prediction error RMSE of 26 mg/dL. An ARX model with a nonlinear forgetting factor scaled according to the glucose range was considered in [6, 7], and a 45-min prediction horizon showed good results. A feedforward neural network (NN) was exploited in [29] and tested on 10 real datasets, incorporating, in addition to CGM data, other inputs such as SMBG readings, information on insulin, meal, hypo- and hyperglycemia symptoms, lifestyle, activity and emotions and predict glucose values up to 75 min. In [37–39], 30-min ahead prediction was performed with a feedforward NN in cascade with the first-order polynomial model in [30]. The inputs to the linear predictor were the past CGM values weighted using a forgetting factor, while the inputs to the NN were current CGM and its trend, information on the past error committed by the polynomial model and information on meal, supplied as plasma glucose rate of appearance obtained from the physiological model of [11]. In [27] a recurrent NN (RNN) was shown to outperform a feedforward NN in predicting glycemia in 12 children on a short-term prediction horizon. Zhao et al. in [40] used a latent variable-based approach to predict future CGM values from past CGM and known carbohydrate and insulin boluses, transformed into time-smoothed inputs using second-order transfer functions. The method was applied to collected clinical data and simulated data generated by the model described in [24, 25]. They concluded that their LV-based method resulted in models whose prediction

accuracy was as least as good as the accuracies of standard AR/ARX models. In [14] a multisensor body monitor providing seven signals related to activity and emotional conditions was used in addition to a CGM monitor to improve glucose prediction. A multivariate ARMAX model with weighted recursive least-squares estimation of the unknown parameters using a variable forgetting factor was proposed. Results showed that the prediction error can be significantly reduced with the addition of the vital signs measurements, as compared to an ARMA model based only on CGM signals. Balakrishnan et al. developed personalized glucose prediction models in [4] taking into account physical exercise, meal, and insulin interventions as input and CGM sensor measurements as output and validated them using the clinical data of 34 children and adolescent. The input-output relationship was described by means of transfer functions with different number of poles (varying between 1 and 3) and with or without zeros for each input.

This contribution is concerned with subspace-based multistep linear predictors, previously introduced in [8, 9], to the purpose of short-term T1DM glycemia prediction. The approach is data-driven to overcome the limitations arising from the lack of the underlying physiological system model and takes into account meal and insulin informations as well as subcutaneously measured blood glucose concentration. To the best knowledge of the authors this is the first time such multistep predictors are applied to diabetes management.

The paper is organized as follows. Section 2 deals with the subspace-based predictors, Sect. 3 introduces the clinical data acquisition, while in Sect. 4 the proposed subspace predictors are used to forecast blood glucose concentrations multiple steps ahead. Results of such a procedure are given in Sect. 5, followed by some discussion and conclusions on the achievements in Sect. 6.

2 Subspace-Based Linear Multistep Predictors

Consider, a discrete-time linear time-invariant system $\mathscr{S}_n(A, B, C, K)$ in innovation form

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Ke_k \\ y_k &= Cx_k + e_k \end{aligned} \tag{1}$$

where $u_k \in \mathbb{R}^m$ is the input, $y_k \in \mathbb{R}^l$ the output, $x_k \in \mathbb{R}^n$ the state, $e_k \in \mathbb{R}^l$ the zero-mean white noise innovation process uncorrelated with u_k and $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{l\times n}$, $K \in \mathbb{R}^{l\times n}$ are constant matrices. In the following, assume that (A, B) is reachable and (C, A) is observable. Further, assume that no linear feedback from the states to the input is present, i.e., input–output data are obtained from an open-loop experiment and the sampling is uniform.

2.1 Notation

The available data sequences $\{u_k\}$, $\{y_k\}$, the state $\{x_k\}$ and the innovation process $\{e_k\}$ will be organized in Hankel matrices denoted by uppercase letters. Subscript indices $[\alpha, \beta]$ of a Hankel matrix will be used to indicate the argument of the upper-left and the lower-left element, respectively, e.g., $U_{[t_1,t_2]}$ will contain in the first column the input history between instants t_1 and t_2. Accordingly, data records of finite length N will be represented by the block rows of the Hankel data matrices and will be denoted by uppercase letters, the subscript indicating the first time instants of the time series:

$$U_{[t_1,t_2]} = \begin{bmatrix} U_{t_1} \\ U_{t_1+1} \\ \vdots \\ U_{t_2} \end{bmatrix} = \begin{bmatrix} u_{t_1} & u_{t_1+1} & \cdots & u_{t_1+N-1} \\ u_{t_1+1} & u_{t_1+2} & \cdots & u_{t_1+N} \\ \vdots & \vdots & \cdots & \vdots \\ u_{t_2} & u_{t_2+1} & \cdots & u_{t_2+N-1} \end{bmatrix}$$

The orthogonal projection of the rows of a given matrix A onto the row space of a given matrix B will be denoted by $\hat{E}\{A \mid B\}$, whereas the symbol $\hat{E}_{\|C}\{A \mid B\}$ will denote the oblique projection of the row space of A onto the row space of B along the row space of C, the projection operator being $\hat{E}\{\cdot\}$. Throughout the paper k will be the discrete-time index, t shall denote the current time instance in the identification problem, t_0 shall be the initial time from which the data are collected, so that $t - t_0 = p$ is the past horizon in the identification problem, T shall be such that $T - t + 1 = f$ represents the future horizon. The two integers p and f are such that $p \geq \max(f, n)$, n model order. Last, the number of steps in the look ahead that one wishes to investigate will be denoted by τ, where $\tau \leq f$.

2.2 Predictors Construction

Suppose ideally that we have observations of the processes $\{u_k\}$, $\{y_k\}$, $\{x_k\}$, $\{e_k\}$, $k \in [t - p, t + f - 1]$. In addition, assume the finite length N of the interval to be large. Since the finite length observed data sequences are a realizations of the underlying stochastic processes in Eq. (1), the following holds:

$$\begin{aligned} X_{k+1} &= AX_k + BU_k + KE_k \\ Y_k &= CX_k + E_k \end{aligned} \tag{2}$$

Furthermore, from the observed samples construct the following block Hankel matrices:

$$U^p = U_{[t-p,t-1]} \in \mathbb{R}^{p \cdot m \times N} \tag{3}$$

$$U^f = U_{[t,t+f-1]} \in \mathbb{R}^{f \cdot m \times N} \tag{4}$$

called the past and future input data matrices, respectively, and

$$Y^p = Y_{[t-p,t-1]} \in \mathbb{R}^{p\cdot l\times N} \tag{5}$$

$$Y^f = Y_{[t,t+f-1]} \in \mathbb{R}^{f\cdot l\times N} \tag{6}$$

called the past and future output data matrices, respectively.

Define the extended observability matrix

$$\mathscr{O}_f = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{f-1} \end{bmatrix} \tag{7}$$

and the Toeplitz matrices containing the impulse responses of the system due to the input u_k and the innovation process e_k, respectively

$$\mathscr{H}_f = \begin{bmatrix} 0 & 0 & \cdots 0 \\ CB & 0 & \cdots 0 \\ \vdots & \vdots & \ddots \vdots \\ CA^{f-2}B & CA^{f-3}B & \cdots 0 \end{bmatrix}, \mathscr{W}_f = \begin{bmatrix} I & 0 & \cdots 0 \\ CK & I & \cdots 0 \\ \vdots & \vdots & \ddots \vdots \\ CA^{f-2}K & CA^{f-3}K & \cdots 0 \end{bmatrix} \tag{8}$$

Then, by iteration of the system equations in Eq. (1), the following matrix input–output relations may be written to express the future output matrix:

$$Y^f = \mathscr{O}_f X_f + \mathscr{H}_f U^f + \mathscr{W}_f E^f \tag{9}$$

In the practical scenario the state sequence X_f is not known, so future outputs cannot be computed. An estimator of future output can be found, however, from the available data as a linear combination of the joint input–output past and the future input (see Appendix A.6 in [34] for a proof), provided that u_k and e_k are uncorrelated:

$$\hat{Y}^f = \hat{\Gamma} Z^p + \hat{\Lambda} U^f \tag{10}$$

where we have introduced the short-hand notation

$$Z^p = \begin{bmatrix} U^p \\ Y^p \end{bmatrix} = Z_{[t-p,t-1]} \in \mathbb{R}^{p\cdot(m+l)\times N} \tag{11}$$

to denote the past joint input–output data matrix. Now, the problem of finding an optimal output predictor can mathematically be formulated as a least-squares problem:

$$\hat{\Gamma}, \hat{\Lambda} = \arg \min_{\substack{\Gamma \in \mathbb{R}^{lf \times (l+m)p} \\ \Lambda \in \mathbb{R}^{lf \times mf}}} \left|\left| Y^f - \begin{bmatrix} \Gamma & \Lambda \end{bmatrix} \begin{bmatrix} Z^p \\ U^f \end{bmatrix} \right|\right|_F^2 \tag{12}$$

where $||\cdot||_F$ stands for the Frobenius norm of a matrix. Note that in this way f prediction problems are solved simultaneously row-wise. Each problem consists in estimating $\hat{Y}_{t+\tau}$, $\tau \in [0, f-1]$:

$$\hat{Y}^f = \begin{bmatrix} \hat{Y}_t \\ \hat{Y}_{t+1} \\ \vdots \\ \hat{Y}_{t+f-1} \end{bmatrix} = \begin{bmatrix} \hat{y}_t & \hat{y}_{t+1} & \cdots & \hat{y}_{t+N-1} \\ \hat{y}_{t+1} & \hat{y}_{t+2} & \cdots & \hat{y}_{t+N} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{y}_{t+f-1} & \hat{y}_{t+f} & \cdots & \hat{y}_{t+f+N-2} \end{bmatrix} \tag{13}$$

A geometric interpretation can be given to the least-squares problem as the orthogonal projection of Y^f onto $\begin{bmatrix} Z^p & U^f \end{bmatrix}^T$, i.e.,

$$\hat{Y}^f = \hat{E}\left\{ Y^f \mid \begin{bmatrix} Z^p \\ U^f \end{bmatrix} \right\} \tag{14}$$

Actually, the orthogonal projection (14) corresponds to the sum of two oblique projections ([21, Lemma 1]) under the assumption $\text{span}(Z^p) \cap \text{span}(U^f) = \{0\}$, with $\text{span}(\cdot)$ standing for the space spanned by the row vectors of a matrix:

$$\hat{Y}^f = \hat{E}_{\|U^f}\{Y^f \mid Z^p\} + \hat{E}_{\|Z^p}\{Y^f \mid U^f\} \tag{15}$$

$$= \hat{\Gamma} Z^p + \hat{\Lambda} U^f \tag{16}$$

Once the operator $\hat{\Gamma}$ and $\hat{\Lambda}$ have been estimated, they can be applied to new data generated by the same underlying mechanisms, to forecast the output τ steps ahead.

Numerically, the efficient implementation of such projection operations relies upon LQ decompositions (i.e., transpose of the QR decomposition) [18]. The following algorithm is outlined:

Algorithm 1 ([34, 35])

1. Consider the following

$$\begin{bmatrix} U^f \\ Z^p \\ Y^f \end{bmatrix} = \begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \begin{bmatrix} {Q_1}^T \\ {Q_2}^T \\ {Q_3}^T \end{bmatrix} \tag{17}$$

where $L_{11} \in \mathbb{R}^{fm \times fm}$, $L_{22} \in \mathbb{R}^{p(m+l) \times p(m+l)}$, $L_{33} \in \mathbb{R}^{fl \times fl}$ and Q_i, $i = 1, \ldots, 3$ are orthogonal.

2. Define

$$\hat{\Gamma} = L_{32} {L_{22}}^{\dagger} \tag{18}$$

$$\hat{\Lambda} = (L_{31} - L_{32} {L_{22}}^{\dagger} L_{21}) {L_{11}}^{-1} \tag{19}$$

3. Use Eq. (10) to calculate the future outputs as $Y^f = \hat{\Gamma} Z^p + \hat{\Lambda} U^f$.

3 Experimental Conditions and Clinical Data Acquisition

In the framework of the DIAdvisor™ project [12], clinical data acquisition from insulin-treated diabetic subjects was accomplished during a 3 days in-hospital trial. The investigations focused on a population of basal-bolus regimen treated subjects, either as combination of multiple daily insulin injections or as continuous insulin infusion from a pump. Prior to any study procedure the subjects participating in the study signed an informed and witnessed consent form. Standard meals for breakfast (8.00 am), lunch (1.00 pm), and dinner (7.00 pm) were served, the amount of administered carbohydrates being 42, 70, and 70 g, respectively. During the whole 3-days-long visit, the participants were permanently equipped with the Abbott Freestyle

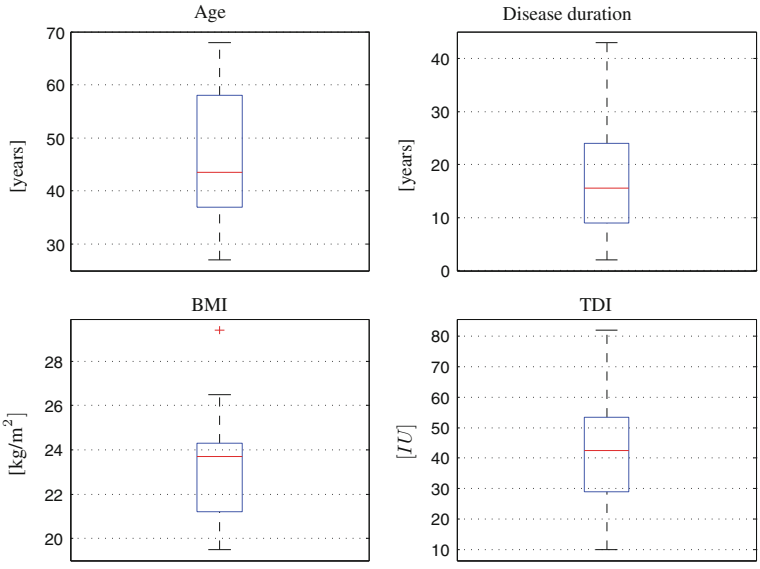

Fig. 1 Population statistics. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles. *Top left* Age (years); *Top right* Disease duration (years); *Bottom left* Body Mass Index (BMI) (kg/m^2); *Bottom right* Total Daily Insulin [IU]

Navigator™ continuous glucose monitoring sensor (CGMS) [1] which provided estimates of glycemia levels from interstitial glucose measurements every 10 min. No specific intervention on usual diabetes treatment was scheduled during the period. The patients decided their insulin needs according to the glucose measurements from the HemoCue® 201+ Glucose Analyzer [19] as usually in activities of daily life. As needed, particularly when hypo- or hyperglycemia occurred, support was provided by nurses and physicians. We present results for a population of 14 patients (9 males and 5 females, age 45.8 ± 12.7 (years), disease duration 18 ± 11.6 (years), BMI 23.4 ± 2.7 (kg/m^2), 10 MDI and 4 CSII, HbA1c 7.6 ± 1 (%), total daily insulin 41.9 ± 18.9 [IU]). The selection criteria were the quantity of data collected (>80 % of the expected), no sensor failures and patient diary correctly filled in. Figure 1 presents population statistics of patients clinically relevant parameters, whereas Fig. 2 gives the empirical distribution of the output signal for all the patients in the population. We compared the empirical distribution with a normal distribution having mean $\mu = 155.32$ (mg/dL) and standard deviation $\sigma = 61.20$ (mg/dL) and a lognormal

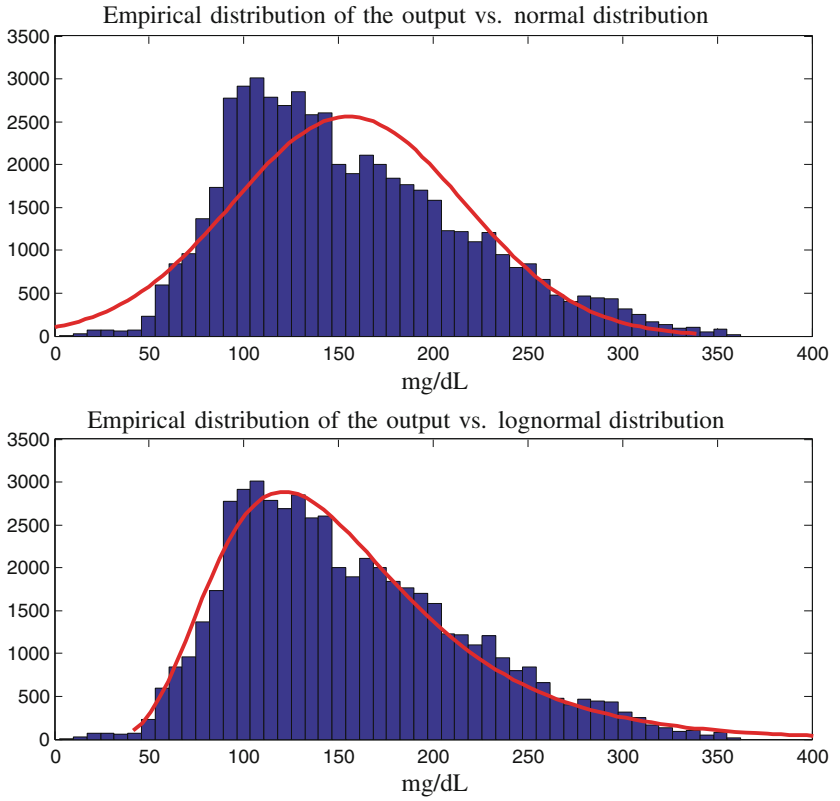

Fig. 2 Empirical distribution of the output for all the patients in the population. *Top* Histogram versus normal distribution; *Bottom* Histogram versus lognormal distribution

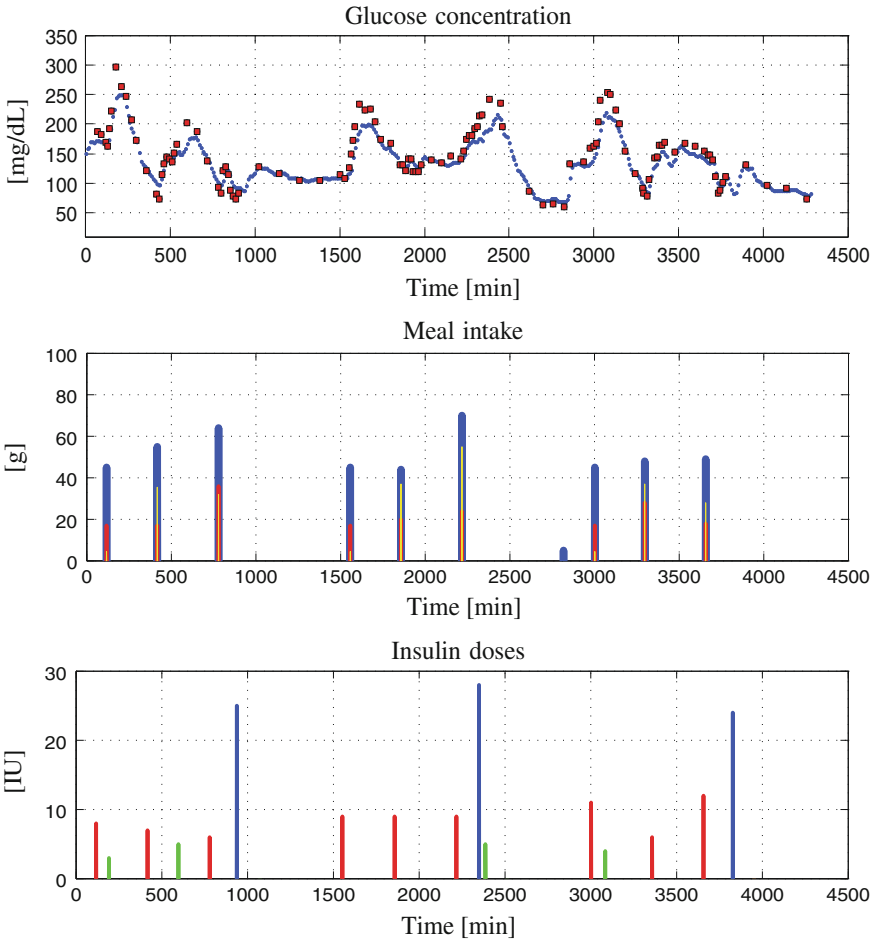

Fig. 3 Patient 102. DAQ trial, *Visit 2*. *Top* Glucose concentration (mg/dL): Navigator™ (*blue*), HemoCue® (*red*); *Upper Center* Meal intake (g): carbohydrates (*blue*), lipids (*red*), proteins (*yellow*); *Lower Center* Insulin doses [IU]: basal (*blue*), bolus (*red*), correction (*green*) versus time (min)

distribution having geometric mean $\mu^* = 144.51\,(\text{mg/dL})$ and geometric standard deviation $\sigma^* = 1.46$ (see Fig. 2). Last, Fig. 3 shows the data for a representative patient.

4 Predicting Diabetes Glycemia with the Multistep Predictors

The physiology of glucose metabolism in diabetes was considered having one output, i.e., glucose level in the bloodstream y_{BG}, and two main inputs, i.e., carbohydrate intake u_{carb} and administered insulin I_{ir} (Fig. 4). Physiological models from the literature were used to filter the raw informations on carbohydrate and insulin intakes collected by the study participants in their diaries. In particular, the glucose intestinal absorption model first presented in [10] was used to retrieve the glucose rate of appearance in plasma after a meal, i.e., $\hat{u}_g$, while the insulin pharmacokinetics model in [25] allowed us to compute the insulin appearing in plasma after subcutaneous injection, i.e., $\hat{u}_i$. Both models exploited mean population values for the parameters appearing therein which were provided us by Dalla Man and co-workers. Figure 5

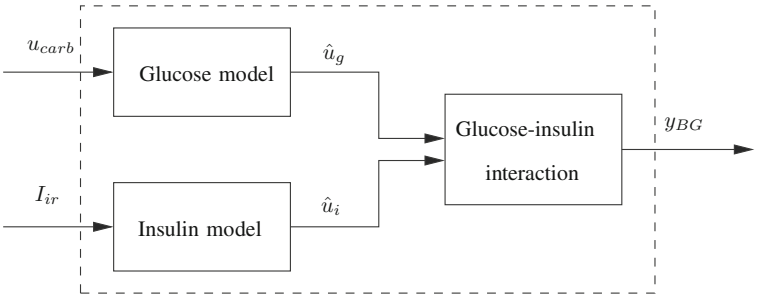

Fig. 4 Diagram of the physiological model describing diabetes glucose metabolism

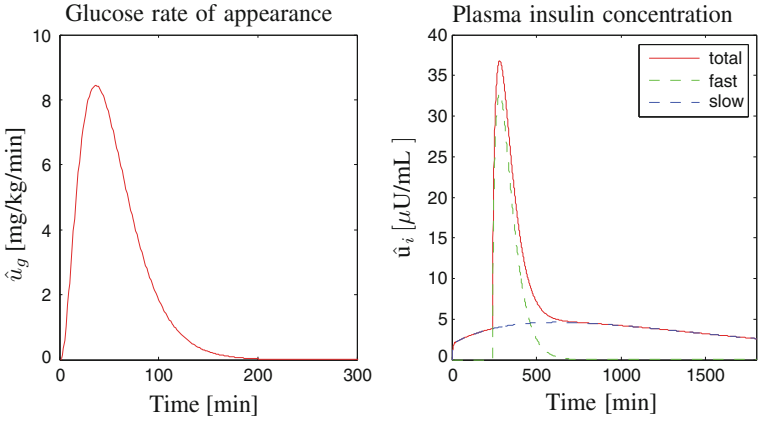

Fig. 5 *Left* Glucose rate of appearance $\hat{u}_g$ after ingestion of 40 (g) carbohydrate by a patient with $m_b = 65$ (kg) at time t = 0, simulated with the glucose intestinal absorption model in [10]; *Right* Plasma insulin concentration: total $\hat{u}_i$ (μU/mL) (*red*), slow-acting (*blue*) and fast-acting (*green*) resulting from a basal dose of 20 [IU] at t = 0 and a bolus of 5 [IU] at t = 240 (min), taken by a patient with $m_b = 65$ (kg), simulated with the insulin pharmacokinetics model in [25]

reproduces $\hat{u}_g$ and $\hat{u}_i$ simulated for a representative patient parameter set. The Abbott FreeStyle Navigator™ [1] records were taken as assessment of glycemia. The time series were interpolated and uniformly resampled at 1 min rate, under the assumption of zero-order hold intersample behavior. A retrospective recalibration against the collected HemoCue [19] samples was performed offline. The algorithm proposed by King et al. in [22], later evaluated by Facchinetti et al. in [15], was used. For each patient a recalibration parameter α^r was found by least-squares estimation:

$$\alpha^r = \arg\min_{\alpha} ||y_{BG} - \alpha y_{IG}||^2 \tag{20}$$

with y_{BG} and y_{IG} denoting the HemoCue and the CGM samples, respectively. The recalibrated traces $y^r_{IG} = \alpha^r y_{IG}$ were provided by Facchinetti and co-workers within the DIAdvisor™ project and were used as output signals. Further, in order to improve numerical precision when computing projections by means of LQ decompositions (see Sect. 2) smoothing of the recalibrated traces was performed. We used the Tikhonov regularization approach [33], following the reasoning presented in [17]. We minimized the functional $f(\dot{y}^r_{IG})$ given by:

$$f(\dot{y}^r_{IG}) = || \, y^r_{IG} - I\dot{y}^r_{IG} \, ||^2 + \lambda^2 \, || \, D\dot{y}^r_{IG} \, ||^2 \tag{21}$$

where $\dot{y}^r_{IG}$ denotes the glucose rate of change, I is a matrix denoting the integral operator, λ is the regularization parameter and D is a second derivative operator chosen to impose smoothness constraints on the derivative of the glucose signal. As a result of such minimization an estimate of the recalibrated traces first derivative is obtained, i.e., $\hat{\dot{y}}^r_{IG}$. Finally, the smoothed signals $\bar{y}^r_{IG}$ can be computed as:

$$\bar{y}^r_{IG} = I\hat{\dot{y}}^r_{IG} \tag{22}$$

We were interested in predictions up to 120 (min) ahead, i.e., denoting by τ_{max} (min) the maximum prediction horizon we take into consideration, $\tau_{max} = 120$. The scheme used for calculation of multistep predictions is reported in Algorithm 2.

Algorithm 2 (*Multistep prediction*)

1. Choose the maximum prediction horizon τ_{max}.
2. Set $p = f = \tau_{max}$.
3. Estimate $\hat{\Gamma}$ and $\hat{\Lambda}$ from Eqs. (18) and (19) applying Algorithm 1 to the first half of the data.

4. Let t denote the current time step. Form the predictions $\hat{y}$ of the second half of the data (validation) using the relation

$$\begin{bmatrix} \hat{y}_t \\ \hat{y}_{t+1} \\ \hat{y}_{t+2} \\ \vdots \\ \hat{y}_{t+f-1} \end{bmatrix} = \hat{\Gamma} \begin{bmatrix} z_{t-p} \\ z_{t-p+1} \\ z_{t-p+2} \\ \vdots \\ z_{t-1} \end{bmatrix} + \hat{\Lambda} \begin{bmatrix} \hat{u}_t \\ \hat{u}_{t+1} \\ \hat{u}_{t+2} \\ \vdots \\ \hat{u}_{t+f-1} \end{bmatrix} \tag{23}$$

where $\left[z_{t-p}\ z_{t-p+1}\ \cdots\ z_{t-1}\right]^\intercal$ are the joint input–output data up to time $t-1$ and $\left[\hat{u}_t\ \hat{u}_{t+1}\ \cdots\ \hat{u}_{t+f-1}\right]^\intercal$ are the future inputs. Notation: $\hat{u} = \left[\hat{u}_i\ \hat{u}_g\right]$ and $y = \bar{y}^r_{IG}$.

5 Results

Given the importance of ultimately predicting glucose levels in blood, the proposed multistep predictors were evaluated with respect to prediction performances on YSI validation data. Denoting with y_{BG} the YSI blood sample measurements and with $\hat{\bar{y}}^r_{IG}$ the predicted glycemia levels, accuracy was assessed on the basis of the following metrics:

- prediction error standard deviation (mg/dL):

$$\sqrt{\mathscr{E}[(y_{BG} - \hat{\bar{y}}^r_{IG})(y_{BG} - \hat{\bar{y}}^r_{IG})^\intercal]} \tag{24}$$

- absolute difference (AD) (mg/dL):

$$\mathrm{AD}(k) = \mid y_{BG}(k) - \hat{\bar{y}}^r_{IG}(k) \mid \tag{25}$$

- relative difference (RD):

$$\mathrm{RD}(k) = \frac{(y_{BG}(k) - \hat{\bar{y}}^r_{IG}(k))}{y_{BG}(k)} \tag{26}$$

- International Organization for Standardization (ISO) [20] criteria.

Clinical accuracy was evaluated with Clarke error grid analysis [23]. Performances were assessed for the prediction horizons

- $\tau = 30, 60, 90, 120\,(\mathrm{min})$

Table 1 presents assessment of prediction per se with respect to blood glucose for the entire population, while Tables 2, 3, 4, 5, 6, 7 and 8 numerical and clinical accuracy of prediction with respect to blood glucose. Figures 6, 7 and 8 show predictions for

Table 1 Multistep short-term predictors performance evaluation

Patient ID	30 (min)	60 (min)	90 (min)	120 (min)
102	13.18	14.75	25.19	45.10
103	10.04	14.74	29.72	44.00
104	14.55	20.00	57.10	113.91
105	9.75	17.76	43.03	62.80
106	19.99	19.98	50.89	80.43
107	30.80	33.80	52.28	77.79
108	10.81	16.97	48.17	62.16
115	17.81	21.97	33.00	39.24
118	16.20	16.49	27.59	60.40
119	16.82	29.60	95.19	132.49
120	27.14	31.38	57.90	97.94
121	15.81	20.79	43.02	74.37
128	17.66	21.95	41.61	55.62
130	14.68	20.75	51.51	78.14

Prediction Error Standard Deviation (mg/dL) versus Prediction Horizon (min) on validation data

Table 2 Numerical and clinical accuracy of predictions

Patient	AD†(mg/dL)		RD‡		ISO (%)	CG-pEGA (%)		CG-rEGA (%)	
	Mean	Median	Mean	Median	∗	A	B	A	B
102	9.65	7.54	−0.01	0.00	95.45	95.65	4.34	91.11	8.88
103	8.00	7.07	−0.01	−0.02	94.87	93.47	6.52	91.11	8.88
104	11.57	9.89	−0.01	0.00	95.45	93.61	4.25	86.95	13.04
105	7.59	6.05	−0.01	−0.01	100	97.77	0.00	95.45	2.27
106	15.23	11.24	−0.03	−0.03	81.39	76.08	17.39	82.22	15.55
107	22.95	17.71	−0.04	0.00	73.91	72.34	27.65	71.73	17.39
108	7.94	6.38	0.00	0.00	100	97.82	0	88.89	11.11
115	12.51	9.07	−0.01	0.00	93.47	93.47	6.52	82.22	6.67
118	11.97	8.92	−0.02	0.00	95.12	91.48	4.25	89.13	8.69
119	13.18	9.62	0.00	0.01	93.61	93.61	6.38	80.43	17.39
120	19.57	13.47	−0.02	0.01	86.95	86.95	13.04	86.66	11.11
121	11.86	8.33	−0.02	0.00	90.47	89.36	8.51	86.95	13.04
128	11.99	6.78	−0.02	0.00	90.62	90.62	9.37	90.32	9.67
130	11.84	11.37	−0.02	−0.03	100	100	0	93.54	6.45

30 (min) prediction

∗Prediction within (20 %) from reference y_{BG} when $y_{BG} \geq 75$(mg/dL); †Eq. (25); ‡Eq. (26)

Table 3 Numerical and clinical accuracy of model-based predictions (Kalman predictor) 30 (min) prediction

Patient	AD†(mg/dL)		RD‡		ISO (%)	CG-pEGA (%)		CG-rEGA (%)	
	Mean	Median	Mean	Median	∗	A	B	A	B
102	12.30	9.30	0.00	0.00	92.05	88.82	7.07	66.59	20.06
103	12.94	9.80	−0.01	−0.02	85.53	83.17	13.23	63.67	22.67
104	13.50	8.03	0.00	0.00	89.89	85.58	9.84	69.42	18.57
105	12.82	8.33	−0.01	−0.03	90.37	90.30	9.33	72.49	19.18
106	13.25	9.37	0.01	0.01	88.00	87.33	11.89	52.84	25.91
107	7.31	5.08	0.02	0.023	97.89	97.89	2.10	78.14	15.90
108	13.12	9.80	0.01	0.02	82.92	81.74	17.53	74.60	15.59
115	9.07	7.34	0.00	0.00	98.28	97.94	1.79	80.75	15.39
118	13.64	10.56	0.03	0.03	91.33	90.76	8.56	50.74	27.39
119	14.04	9.81	0.00	0.00	88.76	88.76	11.23	56.54	25.55
120	12.76	8.69	0.01	0.01	93.60	93.84	6.15	70.60	22.57
121	12.33	7.30	0.01	0.01	85.92	85.53	14.20	61.87	21.08
128	13.11	8.00	0.00	0.00	90.46	90.46	9.43	68.54	18.52
130	12.94	7.66	0.01	0.00	92.64	92.76	6.61	62.23	24.57

∗Prediction within (20 %) from reference y_{BG} when $y_{BG} \geq 75$(mg/dL); †Eq. (25); ‡Eq. (26)

Table 4 Numerical and clinical accuracy of predictions

Patient	AD†(mg/dL)		RD‡		ISO (%)	CG-pEGA (%)		CG-rEGA (%)	
	Mean	Median	Mean	Median	∗	A	B	A	B
102	10.89	10.02	−0.01	0.00	93.02	93.33	6.66	84.09	15.90
103	12.75	12.14	−0.04	−0.05	84.21	84.44	15.55	84.09	15.90
104	16.35	12.93	−0.01	0.00	90.69	89.13	8.69	71.11	20.00
105	14.31	13.26	−0.01	−0.01	95	93.18	4.54	79.06	11.62
106	17.16	16.61	−0.06	−0.05	83.33	77.77	15.55	75.00	20.45
107	25.30	19.53	−0.04	0.00	66.67	65.21	34.78	66.67	15.55
108	12.75	11.25	0.00	0.00	88.63	88.89	11.11	70.45	25.00
115	15.42	10.10	0.00	0.01	91.11	91.11	8.89	75.00	9.09
118	12.81	10.22	−0.02	0.00	95	89.13	4.34	84.44	13.33
119	23.24	20.07	−0.01	0.01	82.60	82.60	17.39	71.11	15.55
120	23.38	17.44	−0.04	0.00	73.33	73.33	26.66	79.54	18.18
121	15.68	13.47	0.00	0.01	82.92	82.60	15.21	71.11	26.66
128	16.10	11.08	−0.01	0.00	81.25	81.25	18.75	83.87	12.90
130	15.77	13.41	−0.03	−0.02	96.67	96.87	3.12	80.64	12.90

60 (min) prediction

∗Prediction within (20 %) from reference y_{BG} when $y_{BG} \geq 75$(mg/dL) ; †Eq. (25); ‡Eq. (26)

Table 5 Numerical and clinical accuracy of model-based predictions (Kalman predictor) 60 (min) prediction

Patient	AD†(mg/dL)		RD‡		ISO (%)	CG-pEGA (%)		CG-rEGA (%)	
	Mean	Median	Mean	Median	∗	A	B	A	B
102	23.95	19.83	−0.01	0.00	70.10	62.82	26.71	60.54	21.70
103	22.75	19.46	−0.04	−0.03	62.24	60.30	32.71	55.1052	31.40
104	28.49	19.29	0.00	0.01	69.07	64.25	26.05	57.61	23.55
105	23.62	17.60	−0.02	−0.06	71.65	69.69	26.61	64.80	21.90
106	26.82	21.84	−0.01	0.02	68.98	67.53	30.56	40.94	24.26
107	14.41	10.67	0.05	0.05	90.20	90.20	9.79	72.65	20.47
108	25.18	17.45	0.03	0.04	54.28	53.17	46.05	79.88	13.34
115	17.20	12.38	0.00	0.00	83.20	82.05	17.69	74.80	18.88
118	28.26	19.86	0.08	0.09	63.67	64.51	33.58	44.99	30.11
119	25.70	16.13	0.01	0.00	76.05	76.05	22.10	49.97	27.03
120	29.50	21.94	0.03	0.06	66.27	67.02	31.38	58.69	25.03
121	23.91	14.70	0.01	0.03	70.38	68.20	28.10	58.33	22.37
128	27.25	19.29	0.00	−0.01	64.00	64.00	33.38	62.13	20.06
130	29.30	17.25	0.03	0.02	73.25	73.53	21.58	49.51	29.09

∗Prediction within (20 %) from reference y_{BG} when $y_{BG} \geq 75$(mg/dL); †Eq. (25); ‡Eq. (26)

Table 6 Numerical and clinical accuracy of predictions

Patient	AD†(mg/dL)		RD‡		ISO (%)	CG-pEGA (%)		CG-rEGA (%)	
	Mean	Median	Mean	Median	∗	A	B	A	B
102	20.32	18.67	0.00	0.02	76.74	77.77	22.22	72.72	18.18
103	27.37	24.15	−0.13	−0.17	48.64	45.45	47.72	62.79	25.58
104	45.79	38.69	−0.03	0.01	46.51	45.65	47.82	48.89	31.11
105	34.36	26.66	−0.07	−0.09	52.50	50.00	40.90	46.51	25.58
106	42.45	37.17	−0.11	−0.12	39.02	38.63	56.81	48.83	30.23
107	39.24	30.19	−0.06	−0.04	45.45	44.44	53.33	47.72	25.00
108	36.45	26.40	0.01	0.06	40.90	40.00	55.55	56.81	18.18
115	24.34	18.34	0.00	0.04	79.54	79.54	20.45	60.46	20.93
118	21.84	18.94	0.01	0.00	75	73.91	23.91	71.11	24.44
119	64.51	42.41	0.01	0.09	42.22	42.22	42.44	43.18	29.54
120	45.43	33.89	−0.14	−0.06	44.44	44.44	55.55	61.36	25.00
121	33.52	23.71	0.00	0.07	57.50	57.77	40.00	52.27	15.90
128	30.70	24.09	0.01	0.02	58.06	58.06	38.70	50.00	26.66
130	38.04	26.20	−0.04	−0.05	60	61.29	35.48	56.66	16.66

90 (min) prediction

∗Prediction within (20 %) from reference y_{BG} when $y_{BG} \geq 75$(mg/dL); †Eq. (25); ‡Eq. (26)

Table 7 Numerical and clinical accuracy of model-based predictions (Kalman predictor) 90 (min) prediction

Patient	AD†(mg/dL)		RD‡		ISO (%)	CG-pEGA (%)		CG-rEGA (%)	
	Mean	Median	Mean	Median	∗	A	B	A	B
102	32.70	27.92	−0.02	0.01	56.04	50.05	36.71	63.98	21.80
103	28.49	27.23	−0.07	−0.02	46.93	47.79	45.38	60.54	26.57
104	40.43	28.13	0.01	0.03	54.80	50.87	35.02	55.20	23.70
105	30.56	24.32	−0.04	−0.10	53.95	52.10	42.82	65.67	20.88
106	36.22	31.41	−0.01	0.06	55.10	53.94	42.56	43.61	22.06
107	20.85	15.95	0.09	0.09	76.61	76.61	23.33	70.54	20.67
108	27.64	21.43	0.04	0.05	50.45	49.53	49.58	83.53	12.51
115	23.73	17.89	−0.01	0.00	65.00	64.15	35.58	74.44	17.90
118	40.13	30.09	0.13	0.15	43.25	44.25	49.69	45.86	29.29
119	35.05	26.03	0.01	0.01	65.38	65.38	29.48	48.17	29.29
120	44.47	33.15	0.06	0.09	48.16	48.20	44.41	54.74	26.73
121	32.31	19.83	0.02	0.04	58.31	55.89	38.51	60.54	19.24
128	40.19	29.95	−0.02	−0.02	46.97	46.97	48.46	61.57	19.90
130	45.23	26.60	0.05	0.03	56.01	55.89	33.23	45.30	28.73

∗Prediction within (20 %) from reference y_{BG} when $y_{BG} \geq 75$(mg/dL); †Eq. (25); ‡Eq. (26)

Table 8 Numerical and clinical accuracy of predictions

Patient	AD†(mg/dL)		RD‡		ISO (%)	CG-pEGA (%)		CG-rEGA (%)	
	Mean	Median	Mean	Median	∗	A	B	A	B
102	37.24	34.30	−0.01	0.00	43.90	44.18	51.16	47.61	28.57
103	40.59	37.24	−0.25	−0.21	33.33	30.23	60.46	61.90	23.80
104	85.97	62.63	−0.02	0.08	26.82	29.54	36.36	27.90	30.23
105	51.41	37.67	−0.17	−0.16	39.47	35.71	52.38	41.46	31.70
106	61.32	47.07	−0.16	−0.05	37.50	34.88	51.16	38.09	30.95
107	61.87	61.74	−0.18	−0.09	32.55	31.81	52.27	34.88	23.25
108	44.27	28.38	0.01	0.06	42.85	41.86	48.83	52.38	26.19
115	31.25	23.10	0.00	0.03	53.48	53.48	46.51	52.38	21.42
118	50.36	41.08	0.04	0.08	34.21	34.09	52.27	48.83	30.23
119	85.83	49.56	0.06	0.07	43.18	43.18	34.09	46.51	18.60
120	79.88	68.11	−0.27	−0.02	32.55	32.55	48.83	54.76	19.04
121	54.22	40.31	−0.08	0.10	43.58	45.45	45.45	34.88	27.90
128	42.71	33.94	0.06	0.09	46.66	46.66	46.66	48.27	37.93
130	56.82	34.51	0.00	0.06	46.66	46.66	50.00	51.72	13.79

120 (min) prediction

∗Prediction within [20 %] from reference y_{BG} when $y_{BG} \geq 75$(mg/dL); †Eq. (25); ‡Eq. (26)

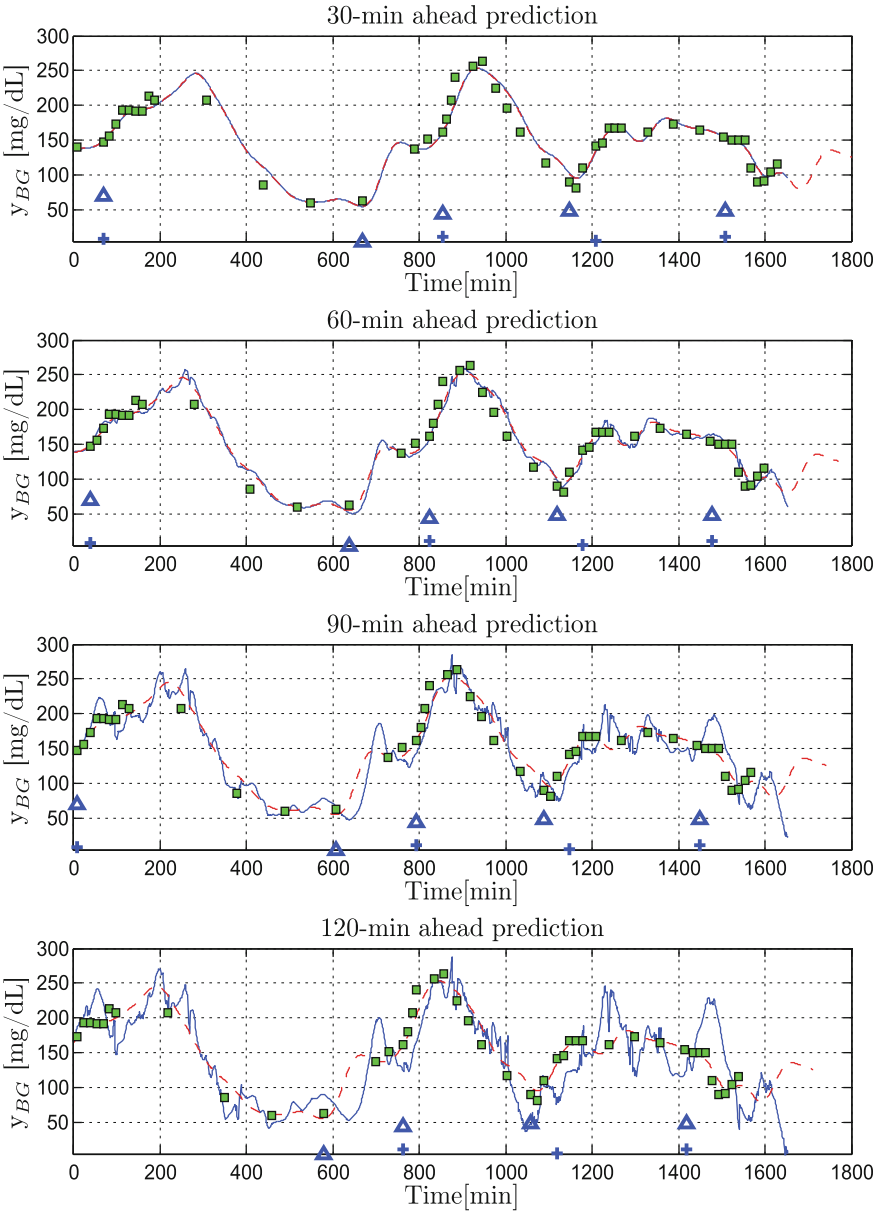

Fig. 6 Patient 102. Multistep predictions. $p = f = 120$. Evaluation on validation data. Predictor $\hat{\bar{y}}^r_{IG}$ (thin), recalibrated, and smoothed interstitial glucose $\bar{y}^r_{IG}$ (*red dashed*) and YSI measurements (*green squares*) (mg/dL) versus time (min). *Top* 30-min ahead; *Top Center* 60-min ahead; *Bottom Center* 90-min ahead; *Bottom* 120-min ahead prediction. Meals and injections are indicated with *triangles* and *plus signs*, respectively

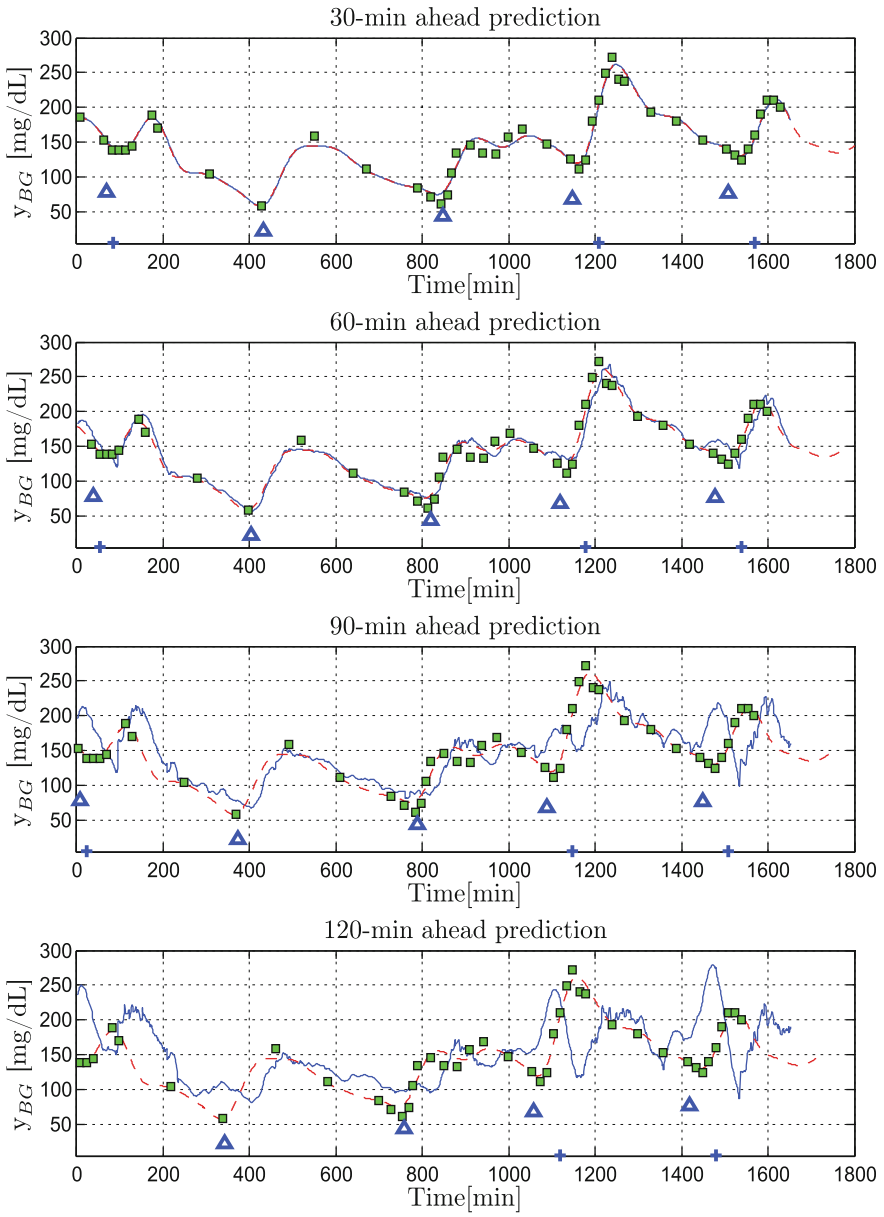

Fig. 7 Patient 105. Multistep predictions. $p = f = 120$. Evaluation on validation data. Predictor $\hat{\bar{y}}^r_{IG}$ (thin), recalibrated, and smoothed interstitial glucose $\bar{y}^r_{IG}$ (*red dashed*) and YSI measurements (*green squares*) (mg/dL) versus time (min). *Top* 30-min ahead; *Top Center* 60-min ahead; *Bottom Center* 90-min ahead; *Bottom* 120-min ahead prediction. Meals and injections are indicated with *triangles* and *plus signs*, respectively

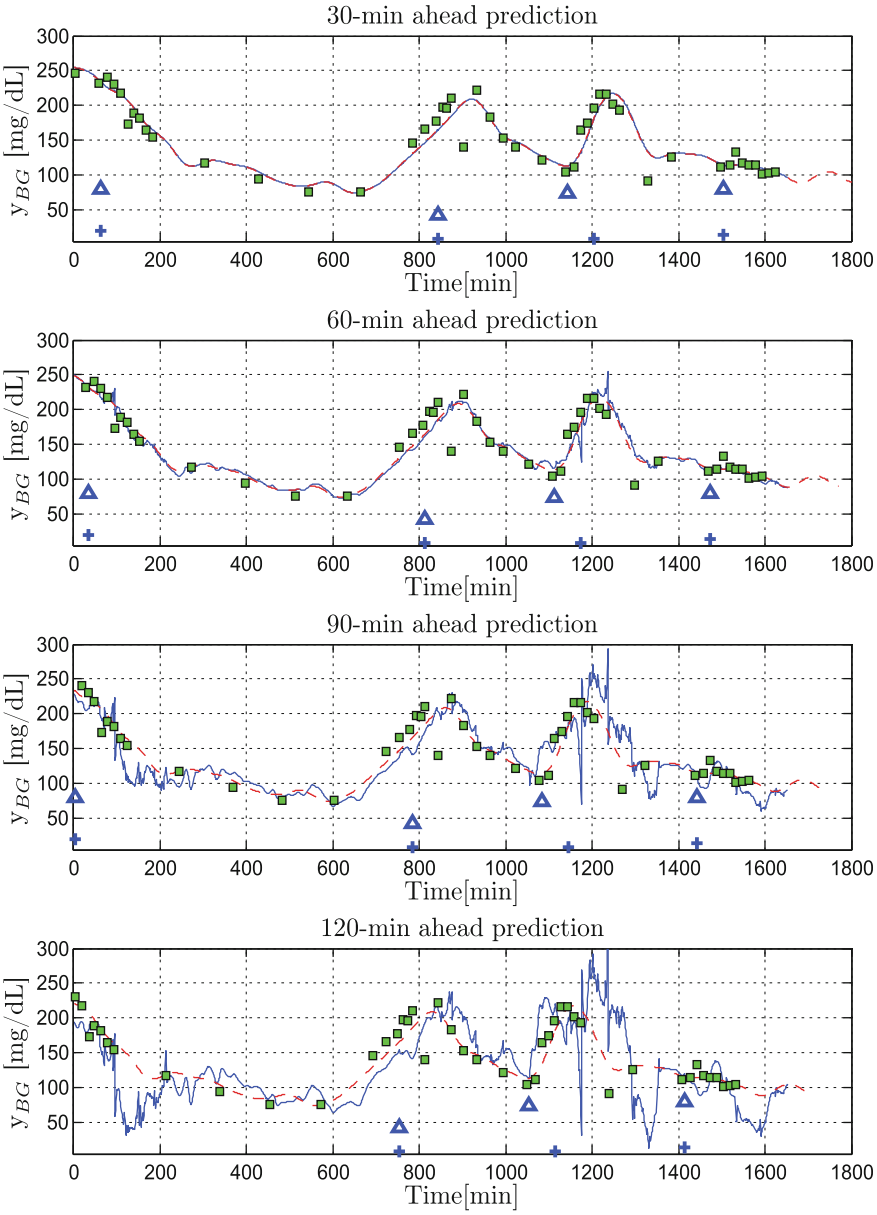

Fig. 8 Patient 115. Multistep predictions. $p = f = 120$. Evaluation on validation data. Predictor without logarithmic transformation of the data (thin), predictor with logarithmic transformation of the data (*dashed*) and recalibrated interstitial glucose $\hat{y}^r_{IG}$ (thick) (mg/dL) versus time (min). *Top* 30-min ahead; *Top Center* 60-min ahead; *Bottom Center* 90-min ahead; *Bottom* 120-min ahead prediction. Meals and injections are indicated with *triangles* and *plus signs*, respectively

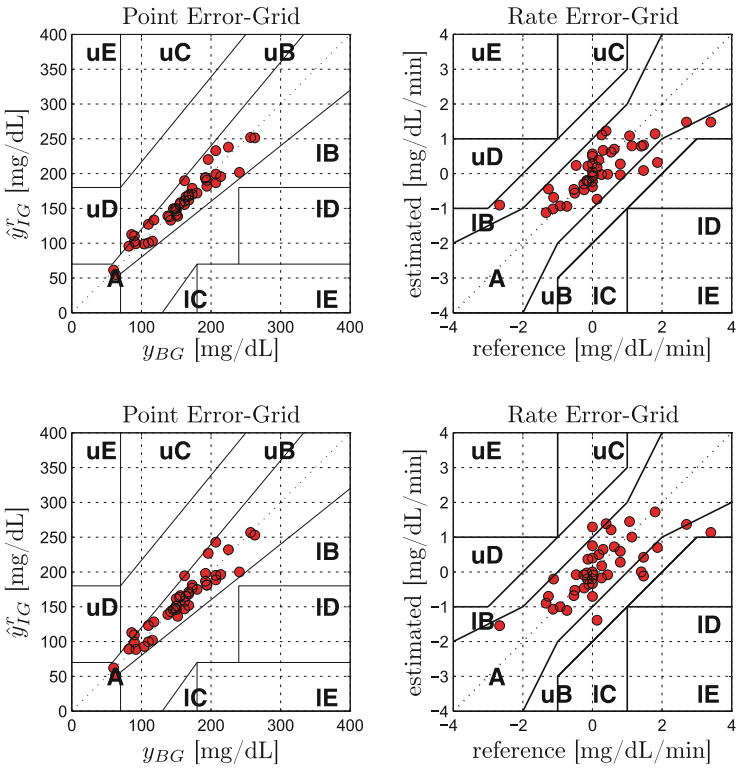

Fig. 9 Patient 102. Predictor performance evaluation. Clarke error grid analysis: *Left* pEGA; *Right* rEGA. *Top* 30 min ahead; *Bottom* 60 min ahead prediction

three representative subjects. Last, Figs. 9 and 10 display Clarke error grid analysis for the first representative subject relative to the various prediction horizons.

The performances achieved by the multistep predictors were compared to those obtained with a Kalman predictor based on individualized third-order ARMAX models identified from each patient data. The model parameters were estimated in a black-box fashion by means of the Matlab System Identification Toolbox function *armax.m* [26]. Results of such investigation are provided in Tables 3, 5, 7, 9 and 10.

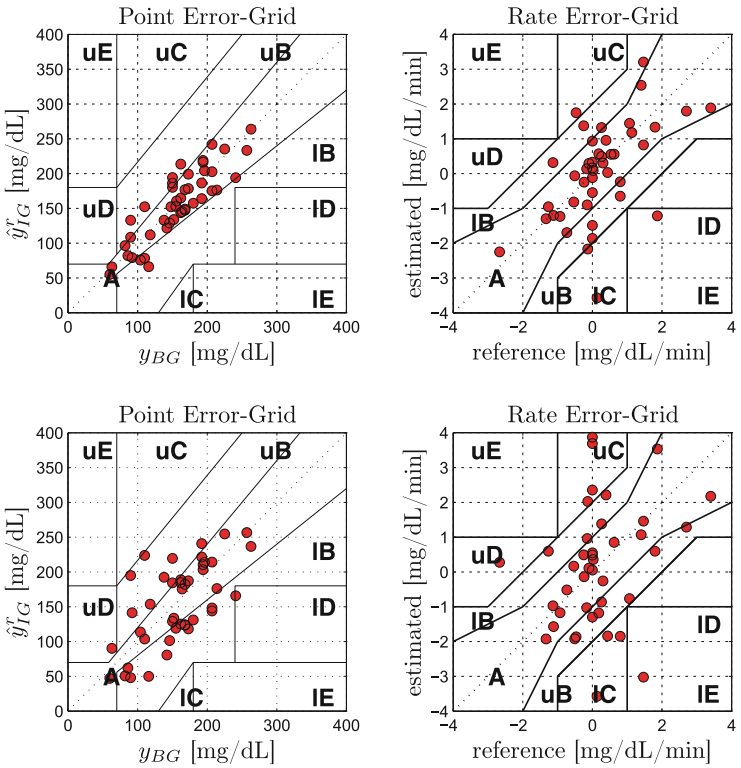

Fig. 10 Patient 102. Predictor performance evaluation. Clarke error grid analysis: *Left* pEGA; *Right* rEGA. *Top* 90 min ahead; *Bottom* 120 min ahead prediction

Table 9 Numerical and clinical accuracy of model-based predictions (Kalman predictor) 120 (min) prediction

Patient	AD†(mg/dL)		RD‡		ISO (%)	CG-pEGA (%)		CG-rEGA (%)	
	Mean	Median	Mean	Median	∗	A	B	A	B
102	39.99	36.41	−0.03	0.02	44.41	39.69	44.61	61.36	24.32
103	33.04	32.14	−0.11	0.00	42.91	40.76	49.53	62.03	25.19
104	49.79	33.59	0.03	0.04	47.03	44.41	37.79	55.56	24.32
105	34.11	29.32	−0.05	−0.11	43.58	41.84	53.02	67.93	22.31
106	41.79	37.16	−0.01	0.08	47.35	46.35	49.79	45.51	25.08
107	26.83	22.67	0.12	0.15	62.92	62.92	36.15	72.65	19.08
108	25.54	19.02	0.03	0.04	56.32	55.12	44.10	81.11	14.52
115	28.59	21.49	−0.02	−0.01	56.57	56.00	43.84	73.57	20.67
118	49.31	37.72	0.17	0.20	37.79	38.92	47.12	46.84	28.88
119	42.80	30.57	0.02	0.00	53.74	53.74	38.30	47.92	30.57
120	55.88	36.73	0.09	0.14	39.31	40.46	46.97	50.69	30.68
121	37.44	23.83	0.02	0.10	51.57	49.84	40.66	58.59	21.03
128	51.10	41.69	−0.03	−0.05	35.89	35.89	57.74	58.33	22.21
130	59.82	39.08	0.06	0.06	44.89	44.61	36.00	47.10	25.19

∗Prediction within (20 %) from reference y_{BG} when $y_{BG} \geq 75$(mg/dL); †Eq. (25); ‡Eq. (26)

Table 10 Kalman short-term predictors performance evaluation

Patient ID	30 (min)	60 (min)	90 (min)	120 (min)
102	21.96	26.71	38.64	45.15
103	15.09	16.47	26.71	37.52
104	16.90	21.99	41.86	71.39
105	11.88	18.82	25.79	37.54
106	12.00	23.92	48.85	67.51
107	20.24	25.48	40.11	49.53
108	19.66	26.41	37.24	43.26
115	16.46	21.96	29.56	39.82
118	14.27	18.18	29.15	48.03
119	14.24	15.02	38.33	53.78
120	13.02	28.68	45.59	53.66
121	10.99	17.43	33.09	45.14
128	19.88	31.20	46.37	61.83
130	26.73	36.58	53.30	75.82

Prediction Error Standard Deviation (mg/dL) versus Prediction Horizon (min) on validation data

6 Discussion and Conclusions

In this contribution, subspace-based data-driven linear multistep predictors have been applied to the problem of short-term prediction of blood glucose concentration in T1DM. Resorting to geometric operations on appropriate subspaces spanned by the measured input–output data sequences, predictor coefficients were estimated directly from the collected data. No model structure selection nor model order determination were, therefore, required.

Some conditions need to be fulfilled in order to have unbiased predictors. Specifically:

- input data should be "rich" enough, that is, the persistency of excitation of the data should be of sufficiently high order
- the space spanned by the row vectors of matrix Z_p and the space spanned by the row vectors of matrix U_f should have zero intersection: $\text{span}(Z_p) \cap \text{span}(U_f) = \{0\}$. This corresponds to no feedback interconnection

The first condition is related to the LQ decompositions required to solve the problem in Eq. (12). Matrix inversions in Eq. (18) is possible if and only if L_{11} is a full rank matrix. This is guaranteed when the input signal u is PE of order at least $f \cdot m$.

The second condition is necessary for the orthogonal projection to split uniquely into the sum of two oblique projections (Eq. (15)). When data are generated in closed loop this condition is no longer satisfied.

Throughout the paper, p was chosen so to meet the following condition:

$$p \geq \max\{n, \tau\}$$

in order for the system (1) to be observable (condition $p \geq n$) and to guarantee predictions up to the largest future horizon we wished to investigate (condition $p \geq \tau$). The parameter has strong connections with model order, and to optimize results it should be given by the user from knowledge or intuitions on the system dynamics. Indeed, requesting to predict with a look-ahead $\tau \gg n$ implies to choose also $p \gg n$, which leads to overparametrization of the model and suboptimal performances.

Figure 2 demonstrates the lognormality of the blood glucose data at hand. Motivated by this, a logarithmic transformation of the data was performed and a new set of predictors was obtained by applying Algorithm 2. Performances of such an approach were compared to those achieved without data transformation and lead to the conclusion that the log-transformed predictor does not perform significantly better.

A third-order ARMAX model was identified for each of the patients and a model-based Kalman predictor was constructed in order to allow comparison with the proposed algorithm. The competitiveness of our approach for the shorter prediction horizons is proven by Tables 3, 5, 7 and 9. However, for longer prediction horizons the Kalman filter outperforms our proposed method in all but four cases. Finally, we remind the reader in passing that the performances were assessed with respect to YSI samples pointwise. In this way, the error in prediction is made up by two components: the sensor error and the proper prediction error.

Acknowledgments This research was partly supported by the European project DIAdvisor ^{TM}, FP7 IST-216592 and the Swedish Research Council by the LCCC Linnaeus Center and the ELLIIT Excellence Center.

References

1. Abbott: FreeStyle Navigator ^{TM} (2013). www.abbottdiabetescare.co.uk. Accessed September 2013
2. Ajmera, I., Swat, M., Laibe, C., Le Novere, N., Chelliah, V.: The impact of mathematical modeling on the understanding of diabetes and realted complications. CPT Pharmacomet. Syst. Pharmacol. **2**(e54) (2013). doi:10.1038/psp.2013.30
3. Balakrishnan, N., Rangaiah, G., Samavedham, L.: Review and analysis of blood glucose (bg) models in type 1 diabetic patients. Ind. Eng. Chem. Res. **50**(21), 12041–12066 (2011)
4. Balakrishnan, N., Samavedham, L., Rangaiah, G.: Personalized hybrid models for exercise, meal and insulin interventions in type 1 diabetic children and adolescents. Ind. Eng. Chem. Res. **52**(36), 13020–13033 (2013)
5. Bremer, T., Gough, D.A.: Is blood glucose predictable from previous values? Diabetes **48**, 445–451 (1999)

6. Castillo-Estrada, G., Del Re, L., Renard, E.: Nonlinear gain in online prediction of blood glucose profile in type 1 diabetic patients. In: Proceedings of the 49th IEEE Conference on Decision and Control (CDC2009), Atlanta, GA, USA, pp. 1668–1673 (2010)
7. Castillo-Estrada, G., Kirchsteiger, H., Del Re, L., Renard, E.: Innovative approach for online prediction of blood glucose profile in type 1 diabetes patients. In: Proceedings of the American Control Conference ACC2010, Baltimore, MD, USA, pp. 2015–2020 (2010)
8. Cescon, M., Johansson, R.: Multi-step-ahead multivariate predictors: a comparative analysis. In: Proceedings of the 49th IEEE Conference on Decision and Control (CDC2010), Atlanta, USA, pp. 2837–2842 (2010)
9. Cescon, M., Johansson, R.: On data-driven multistep subspace-based linear predictors. In: Proceedings of the 18th IFAC World Congress (IFAC2011), Milano, Italy, pp. 11,447–11,452 (2011)
10. Dalla Man, C., Camilleri, M., Cobelli, C.: A system model of oral glucose absorption: validation on gold standard data. IEEE Trans. Biomed. Eng. **53**(12), 2472–2477 (2006)
11. Dalla Man, C., Rizza, R.R., Cobelli, C.: Meal simulation model of the glucose-insulin system. IEEE Trans. Biomed. Eng. **54**(10), 1740–1749 (2007)
12. DIAdvisor: The DIAdvisor ^{TM} (2012). http://www.diadvisor.eu. Accessed May 2013
13. Eren-Oruklu, M., Cinar, A., Quinn, L., Smith, D.: Estimation of future glucose concentrations with subject-specific recursive linear models. Diabetes Technol. Ther. **11**(4), 243–253 (2009)
14. Eren-Oruklu, M., Cinar, A., Rollins, D., Quinn, L.: Adaptive system identification for estimating future glucose concentrations and hypoglycemia alarms. Automatica **48**, 1892–1897 (2012)
15. Facchinetti, A., Sparacino, G., Cobelli, C.: Reconstruction of glucose in plasma from interstitial fluid continuous glucose monitoring data: role of sensor calibration. J Diabetes Sci. Technol. **1**(5), 617–623 (2007)
16. Finan, D., Palerm, C., Doyle, J., Seborg, D.: Effect of input excitation on the quality of empirical dynamic modes for type 1 diabetes. Process Syst. Eng. **55**(5), 1135–1146 (2009)
17. Gani, A., Andrei, G., Srinivasan, R., Ward, K., Jaques, R.: Predicting subcutaneous glucose concentration in humans: data-driven glucose modeling. IEEE Trans. Biomed. Eng. **56**(2), 246–254 (2009)
18. Golub, G., Van Loan, C.: Matrix Computations. The Johns Hopkins University Press, Baltimore (1996)
19. Hemocue: HemoCue ® (2013). www.hemocue.com. Accessed September 2013
20. International Organization for Standardization (ISO), Publication 15197: In vitro diagnostic test systems: requirements for blood glucose monitoring systems for self-testing in managing diabetes mellitus, Geneva (2003)
21. Katayama, T., Picci, G.: Realization of stochastic systems with exogenous inputs and subspace identification methods. Automatica **35**, 1635–1652 (1999)
22. King, C., Anderson, S., Breton, M., Clarke, W., Kovatchev, B.: Modeling and calibration effectiveness and blood-to-interstitial glucose dynamics as potential confounders of the accuracy of continuous glucose sensors during hyperinsulinemic clamp. J. Diabetes Sci. Technol. **1**(3), 317–322 (2007)
23. Kovatchev, B., Gonder-Frederik, L., Cox, D., Clarke, W.: Evaluating the accuracy of continuous glucose-monitoring sensors. Diabetes Care **27**(8), 1922–1928 (2004)
24. Kovatchev, B., Breton, M., Cobelli, C., Dalla Man, C.: Method, system and computer simulation environment for testing of monitoring and control strategies in diabetes (2008)
25. Kovatchev, B., Breton, M., Dalla Man, C., Cobelli, C.: In silico model and computer simulation environment approximating the human glucose/insulin utilization. Master File MAF-1521, Food and Drug Administration (FDA), Silver Spring, MA (2008)
26. MathWorks: MathWorks (2013). www.mathworks.com/products/matlab/
27. Mougiakakou, S., Prountzou, A., Iliopoulou, D., Nikita, K., Vazeou, A., Bartsocas, C.: Neural network based glucose-insulin metabolism models for children with type 1 diabetes. In: Proceedings of 28th IEEE EMBS, vol. 2, pp. 3545–3548. Piscataway, NJ (2006)
28. Naumova, V., Pereverzyev, S., Sivananthan, S.: A meta-learning approach to the regularized learning-case study: blood glucose prediction. Neural Netw. **33**(9), 181–193 (2012)

29. Pappada, S., Cameron, B., Rosman, P., Bourey, R., Papadimos, T., Olorunto, W., Borst, M.: Neural network-based real-time prediction of blood glucose in patients with insulin-dependent diabetes. Diabetes Technol. Ther. **13**(2), 135–141 (2011)
30. Sparacino, G., Zanderigo, F., Maran, A., Facchinetti, A., Cobelli, C.: Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series. IEEE Trans. Biomed. Eng. **54**(5), 931–937 (2007)
31. Sparacino, G., Facchinetti, A., Maran, A., Cobelli, C.: Continuous glucose monitoring time series and hypo/hyper glycemia prevention: requirements, methods, open problems. Current Diabetes Rev. **4**(3), 181–192 (2008)
32. The American Diabetes Association: Standards of medical care in diabetes 2013. Diabetes Care **36**(Supplement 1), S11–S66 (2013)
33. Tikhonov, A., Arsenin, V.: Solutions of Ill-Posed Problems. Winston, Washington (1977)
34. Van Overschee, P., De Moor, B.: Subspace Identification for Linear Systems : Theory-Implementation-Application. Kluwer Academic Publishers, Boston (1996)
35. Verhaegen, M.: Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. Automatica **30**(1), 61–74 (1994)
36. Williams, G., Pickup, J.: Handbook of Diabetes, 2nd edn. Blackwell Science, Oxford (1992)
37. Zecchin, C., Facchinetti, A., Sparacino, G., De Nicolao, G., Cobelli, C.: A new neural network approach for short-term glucose prediction using continuous glucose monitoring time-series and meal information. In: Proceedings of the of 33rd International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2011), Boston, Massachussets, USA (2011)
38. Zecchin, C., Facchinetti, A., Sparacino, G., De Nicolao, G., Cobelli, C.: Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration. IEEE Trans. Biomed. Eng. **59**(6), 1550–1560 (2012)
39. Zecchin, C., Facchinetti, A., Sparacino, G., Cobelli, C.: Reduction of number and duration of hypoglycemic events by glucose prediction methods: a proof-of-concept in silico study. Diabetes Technol. Ther. **15**(1), 66–77 (2013)
40. Zhao, C., Dassau, E., Jovanovic, L., Zisser, H., Doyle, F., Seborg, D.: Predicting subcutaneous glucose concentration using a latent-variable-based statistical method for type 1 diabetes mellitus. J. Diabetes Sci. Technol. **6**(3), 617–633 (2013)

Empirical Representation of Blood Glucose Variability in a Compartmental Model

Stephen D. Patek, Dayu Lv, Edward A. Ortiz, Colleen Hughes-Karvetski, Sandip Kulkarni, Qian Zhang and Marc D. Breton

Abstract Eating and exercise behaviors have an important effect on glycemic outcomes in type 1 diabetes, yet these influences are difficult to assess in real-life settings. While existing mathematical models faithfully represent the dynamic relationships of (i) oral carbohydrate ingestion and (ii) glucose and insulin transport/action in various compartments of the body, accurate models of meal and exercise behaviors are needed to realistically capture the wide excursions of blood glucose observed in the field, and this has been a bottleneck in preclinical *in silico* evaluation of advanced systems including the artificial pancreas. This work presents a method of using continuous glucose monitoring and insulin pump data to extract a BG variability signature represented by oral carbohydrate *net effect*, which can be "fed" back into the mathematical model to (i) reproduce the original BG time series from the original record of insulin delivery and (ii) be used to approximate the effect of a modified schedule of insulin delivery. We provide details of the basic method and illustrate the approach using both the Virginia/Padova Type 1 Simulator and human-subject data collected in a field study.

S.D. Patek (✉) · D. Lv · E.A. Ortiz · M.D. Breton
Center for Diabetes Technology, University of Virginia,
Charlottesville, Virginia, USA
e-mail: patek@virginia.edu

S.D. Patek
Systems and Information Engineering, University of Virginia,
Charlottesville, Virginia, USA

C. Hughes-Karvetski
Carolinas Healthcare System, Charlotte, North Carolina, USA

S. Kulkarni
Fischell Department of Bioengineering, University of Maryland,
College Park, Maryland, USA

Q. Zhang
BMW Brilliance Automotive Ltd., Shanyang, China

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*,
Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_8

1 Introduction

In maximizing the effectiveness of insulin therapy in Type 1 Diabetes (T1D) it is critical to adapt to the characteristics of the individual patient [4, 8]. Endocrinologists have long recognized in the importance of inter- and intra-patient variability with regard to insulin sensitivity and other physiological traits, and this is reflected in the methods used to titrate insulin therapy parameters including carbohydrate ratios, correction factors, correction thresholds and targets, and basal insulin doses [6, 15]. Behavioral characteristics (eating, exercise, and self-treatment behaviors) are equally important, particularly in separating the treatment of meals from the need to provide a basal insulin "baseline" that is adapted to circadian variability of insulin sensitivity throughout the day [7, 19].

Variability of physiological characteristics is well represented in existing mathematical models of T1D, which are increasingly important in the design and evaluation of advanced technologies and treatments [16, 18]. For example, inter-patient variability in endogenous glucose production is captured in the distinct *in silico* subjects of the Virginia/Padova Type 1 simulator [16, 18], as well as in other simulation platforms [9, 20]. *Intra*-patient variability is represented now in more recent works, cf. the updated Virginia/Padova Simulator [5] where physiological parameters that relate to insulin sensitivity are drawn from probability distributions that are specific to the time of day, capturing enhanced resistance overnight and exhibiting increasing insulin sensitivity throughout the day.

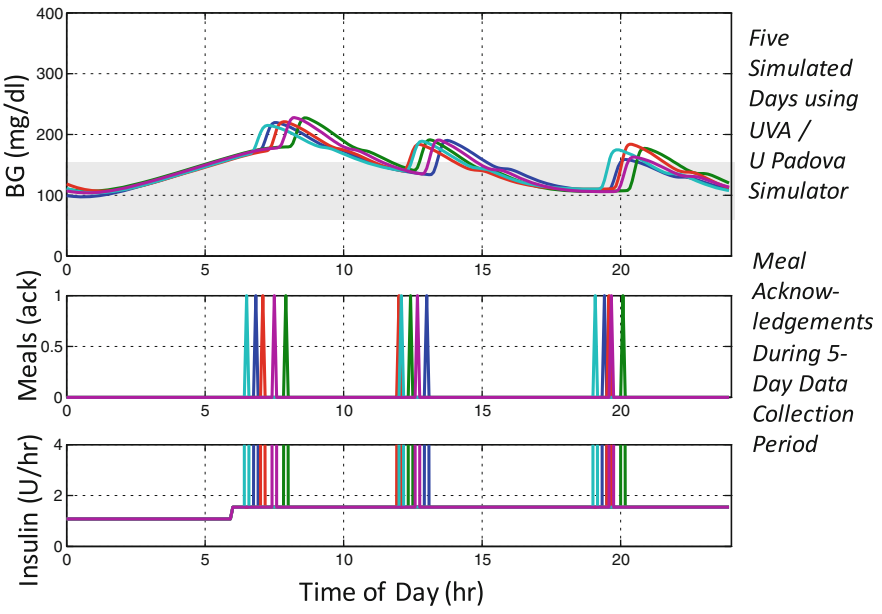

Fig. 1 An attempt to capture five days of BG variability in an *in silico* Type 1 patient

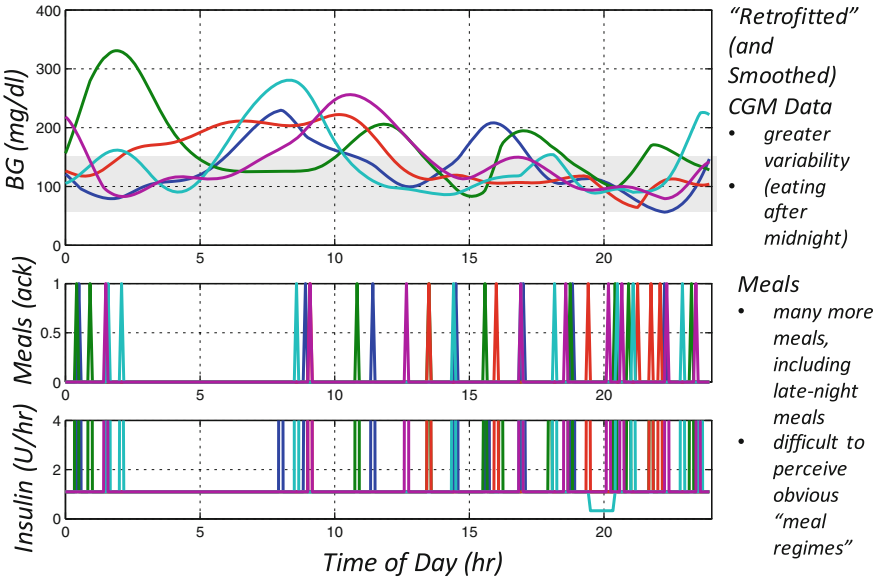

Fig. 2 Five-day snapshot of a real patient with Type 1 diabetes

While some attempts have been made to accurately represent behavioral influences in computer simulation models [11–14], eating and exercise behaviors are difficult to assess and represent. Figures 1 and 2 attempt to illustrate this point informally. Figure 1 was generated using the 2012 version of the FDA-accepted Virginia/Padova Type 1 simulator, showing BG traces for one of the *in silico* subjects for a five-day scenario in which meal times (three meals a day) and meal amounts (grams CHO) are randomized, with meal-related boluses computed using the patient's individualized carbohydrate ratio, and where the patient's "optimal" basal rate is intentionally reduced by 25 % to roughly capture a dawn effect. While the simulated traces do exhibit the morning hyperglycemia that one would expect, along with some variability in the timing and depth of postprandial glycemic excursions, the simulated traces look nothing like the five-day snapshot shown in Fig. 2 of CGM data from a real patient with Type 1 diabetes, exhibiting much greater BG variability (due in part to the fact that the patient, like most people, eats throughout the day, and often eats more than three meals a day).

In this chapter, we present our preliminary experience with the use of deconvolution methods [3, 10] in reconciling continuous glucose monitoring (CGM) data and insulin pump records leading to the estimation of an oral carbohydrate "net effect" signal. Here, given a linear time-invariant (LTI) model of the patient, the recovered net effect signal is such that if the modeled patient were (i) to absorb carbohydrates exactly as specified by the net effect function and (ii) to insulinize exactly as recorded, then the output of the model would accurately recover the CGM signal that was originally recorded. In a sense, the net effect signal itself is designed

to allow for a "simulation" of the patient data, and the advantage of this is that we can replay the patient's treatment day experimenting with different amounts of insulin at different times to see under what circumstances it is possible to achieve better glycemic outcomes [17]. Our goal here is to provide a snapshot of the method, providing details of the basic method, the compartmental model currently being used, the CGM preprocessing steps taken, and other implementation issues. We illustrate the approach with (i) data collected in a month-long field study of patients with Type 1 diabetes and (ii) simulated data using the Virginia/Padova Type 1 Simulator.

2 Oral Carbohydrate "Net Effect": Reconciling CGM and Pump Data via Regularized Deconvolution

Many sources of uncertainty confound the estimation of metabolic parameters of patients with Type 1 diabetes, including (i) unknown and time-variable insulin sensitivity (affected by physical activity, menstrual cycle, circadian rhythm, illness, stress), (ii) variations in insulin transport related to physical activity, (iii) the content of meals and the profile of meal ingestion, and (iv) variable gastric emptying, among many others. In analyzing historical data, meals, and physical activity pose significant challenges. Even with conscientious logging of meals and physical activity, errors stemming from inaccurate carb counting, meal content, and meal timing can significantly disrupt attempts to estimate insulin sensitivity from CGM and pump data. Failure to acknowledge meals or physical activity can lead to nonsensical estimates of physiological parameters. Ultimately, it is not always possible to attribute fluctuations in BG specifically to meals, physical activity, changes insulin sensitivity and/or counter regulation.

In this work, we take the view that, out of all of the factors that contribute to trends in BG over time, meals (the existence of meals, the content of meals, the absorption of meals) represent the main confounding factors in the retrospective analysis of CGM and insulin data. With this in mind we introduce the notion of oral carbohydrate "net effect" corresponding to the meal arrival process that best explains, through a mathematical model of the patient's physiology, the correlated time series of CGM and insulin delivery data. As illustrated in Fig. 3, we define a process for estimating the unknown meal signal via regularized deconvolution from records of CGM (subcutaneous glucose concentration) and insulin delivery data.

For computational efficiency (leading to a closed-form solution), our net effect signal is derived from a discrete-time, linear time-invariant model:

$$x(k+1) = Ax(k) + Bu(k) + G\omega(k) \tag{1}$$

$$y(k) = Cx(k), \tag{2}$$

where each discrete-time step corresponds to a fixed sampling period and where (i) $x(k)$ is a vector of state variables that represent glucose and insulin trans-

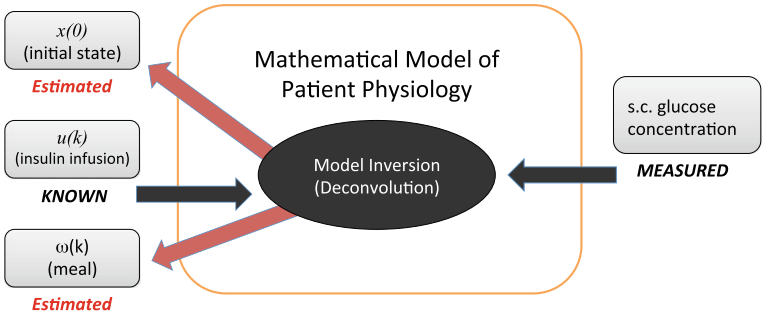

Fig. 3 Net effect: a model-based signature of patient behavior

port/interaction in various parts of the body, (ii) $u(k)$ is insulin delivery (mU/min) in the kth sampling interval relative to the patient's basal rate at that time, (iii) $\omega(k)$ is meal carbohydrates (mg-CHO/min) consumed in the kth sampling interval, and (iv) $y(k)$ is the kth sample of interstitial glucose concentration (mg/dl) relative to a reference BG value G_{ref}.

For the numerical results presented in the sequel the specific model that we use is derived from the minimal model of glucose-insulin kinetics [1] and is referred to as the subcutaneous oral glucose minimal model (SOGMM), which we have used in our prior work [11–13]. The structure and parameters of the SOGMM, along with the steps taken to linearize and discretize the model, are described in detail in the Appendix.

Before proceeding, it is useful to introduce notation for both the actual insulin delivered in any given stage, along with the preprogrammed basal profile. To this end, let $J_{del}(k)$ (U) denote the total insulin delivered in stage k. Similarly, let $basal(k)$ (U/hr) represent the patient's preprogrammed basal profile specification of basal insulin in stage k. Then, we have that

$$u(k) = 1000 \cdot J_{del}(k)/T_S - 1000 \cdot basal(k)/60, \tag{3}$$

representing the effect of insulin boluses, extended boluses, and temporary basal rates, where T_S (min) is the length of the sampling interval. ($T_S = 5$ for most CGM devices.)

Given the initial state $x(0)$, a vector of differential control inputs $\tilde{u} = (u(0), u(1), \ldots, u(T-1))'$, and a vector of meal inputs $\tilde{\omega} = (\omega(0), \omega(1), \ldots, \omega(T-1))'$, the LTI model predicts future differential glucose values

$$\tilde{y} = (y(0), y(1), \ldots, y(T))' = \mathscr{A}x(0) + \mathscr{B}\tilde{u} + \mathscr{G}\tilde{\omega}, \tag{4}$$

where $\mathscr{A}$, $\mathscr{B}$, and $\mathscr{G}$ are appropriately defined matrices:

$$\mathscr{A} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^T \end{bmatrix} \mathscr{B} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ CB & 0 & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{T-1}B & CA^{T-2}B & \cdots & CB \end{bmatrix} \mathscr{G} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ CG & 0 & \cdots & 0 \\ CAG & CG & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{T-1}G & CA^{T-2}G & \cdots & CG \end{bmatrix}.$$

2.1 Net Effect Core Algorithm

The goal of the net effect core algorithm is to compute estimates $\hat{x}(0)$ and $\hat{\tilde{\omega}}$ of the initial state $x(0)$ and the vector of meal inputs $\tilde{\omega}$ from knowledge of insulin delivery $\tilde{u}$ and (differential) CGM measurements:

$$\tilde{y}_{cgm} = (CGM(0) - G_{ref},\ CGM(1) - G_{ref},\ \ldots,\ CGM(T) - G_{ref})', \tag{5}$$

where $CGM(0)$, $CGM(1)$, $\ldots$, $CGM(T)$ are a vector of CGM values that may be the result of CGM preprocessing (cf. Sect. 2.2). The basic idea is to compute $\hat{x}(0)$ and $\hat{\tilde{\omega}}$ to satisfy Eq. (4) as closely as possible with given CGM data ($\tilde{y}_{cgm}$ instead of $\tilde{y}$) and known insulin inputs $\tilde{u}$. We refer to the estimate $\hat{\tilde{\omega}}$ of the meal input signal as *net effect* because it, along with the estimate of the initial state $\hat{x}(0)$, are the only means of "explaining" variability in measured BG, variability that is due in fact to many factors beyond meals, including dramatic variability of insulin sensitivity owing to physical activity, illness, circadian effectors, or, on a longer time-scale, menstrual cycle effects.

The approach taken here is essentially a regularized deconvolution method, with the additional feature of wrapping the estimation of the initial state into the quadratic optimization model for the deconvolution. We express the net effect estimate as a linear combination of basis vectors $e_i \in \Re^T$, $i = 1, \ldots, n$:

$$\hat{\tilde{\omega}} = \sum_{i=1}^{n} v_i e_i = Ev, \tag{6}$$

where the columns of E are the basis vectors e_i and $v = (v_1, \ldots, v_n)'$. Defining

$$\xi = \tilde{y}_{cgm} - \mathscr{B}\tilde{u} \tag{7}$$

we compute the net effect and initial state estimates to solve the optimization problem

$$\begin{aligned} \min_{x(0),v} \big[&(\xi - \mathscr{A}x(0) - \mathscr{G}Ev)'\Lambda_{fit}(\xi - \mathscr{A}x(0) - \mathscr{G}Ev) \\ &+ x(0)'\Lambda_{reg,1}x(0) \ + \ v'E'\Lambda_{reg,2}Ev \big], \end{aligned} \tag{8}$$

where in the first term of the objective function $\Lambda_{fit} \geq 0$ is a positive semindefinite weighting matrix that defines the penalty for failing to satisfy Eq. (4) and $\Lambda_{reg,1} \geq 0$ and $\Lambda_{reg,2} \geq 0$ define regularization terms that penalize large initial state and net effect estimates, respectively. The solution to this optimization model can be expressed in closed form, and for this it is convenient to rewrite Eq. (8) as

$$\min_{\theta}[(\xi - M_{fit}\theta)'\Lambda_{fit}\left(\xi - \theta M_{fit}\right) + \theta'\Lambda_{reg}\theta] \tag{9}$$

where

$$\theta = \begin{pmatrix} x(0) \\ \nu \end{pmatrix}, \quad M_{fit} = [\mathscr{A} \quad \mathscr{G}E], \quad \Lambda_{reg} = \begin{bmatrix} \Lambda_{reg,1} & 0 \\ 0 & E'\Lambda_{reg,2}E \end{bmatrix}. \tag{10}$$

The closed-form solution of Eq. (9) is then

$$\theta^* = \left(M'_{fit}\Lambda_{fit}M_{fit} + \Lambda_{reg}\right)^{-1} M'_{fit}\Lambda_{fit}\xi = \begin{pmatrix} \hat{x}(0) \\ \nu^* \end{pmatrix} \tag{11}$$

where the inverse is guaranteed to exist under certain circumstance, such as when $\Lambda_{reg} > 0$. We take $E\nu^*$ to be the estimate of net effect based on the CGM and insulin delivery data.

2.2 CGM Preprocessing

The net effect core algorithm above is essentially a nonstatistical procedure that ignores sensor noise and treats the meal arrival process as the only unknown input to the system. For this reason, it is necessary to filter the CGM measurements and/or apply smoothing. (Otherwise, the net effect core algorithm is prone to attempting to explain erratic signal artifacts as abrupt changes in the meal arrival process.) While filtering and smoothing of the CGM data can be achieved in different ways, we have found it convenient, as in Sect. 4 to make use of a "retrofitting" procedure in which (i) the CGM trace is smoothed via cubic spline interpolation, (ii) the smoothed CGM trace is warped via scaling and shifting, forcing the modified signal to pass through the SMBG samples, and (iii) gaps in the CGM data are interpolated.

Another critical preprocessing step is to ensure that the net effect estimate is free of "edge effects." While the core net effect algorithm is designed to produce an estimate of the initial condition of the dynamic system $\hat{x}(0)$, the interaction between $\hat{x}(0)$ and the estimate of net effect itself $\hat{\tilde{\omega}}$ is complicated. For this reason, it is generally helpful to prepend the CGM and insulin data streams with relevant data *before* the timeframe of interest. (We find that these "begin-game" effects usually last 30 min to one hour.) Thus, if it is desired to have an accurate assessment for the 24 h period from midnight to the following midnight, it is beneficial to associate the discrete stage $k = 0$ with a time at least 30 min *before* midnight, chosen to allow for a suitable "warm-up"

period. Similarly, there is an "end-game" effect that must be addressed. Intuitively, in order to estimate the meal arrival process through a particular time (e.g., midnight), then it is necessary to observe CGM data beyond that time (e.g., several hours past midnight) in order to perceived the impact of oral carbohydrates that arrive just prior to the end of the interval. Thus, it is beneficial to associate the end stage T with a time *after* the 24-h day, chosen to allow for the transient effect of any last-minute carbs to die out. (In the numerical results of the sequel, $k = 0$ corresponds to eight hours before midnight and $k = T$ corresponds to four hours after the following midnight.)

2.3 Discussion: "Net Effect" Versus "Meal Estimation"

We refer to the estimate $\hat{\tilde{\omega}}$ of as carbohydrate "net effect" since, in reality, there are many sources of BG variability beyond meals, e.g., dramatic variability of in insulin sensitivity owing to physical activity, illness, circadian effectors, or, on a longer time-scale, menstrual cycle effects. Glucose-insulin interactions in vivo are more complicated than natively expressed in the compartmental model described in the Appendix. Thus, other effects besides meals are "wrapped up" in the net effect computation. First, since the mathematical model assumes a fixed (patient-specific) insulin sensitivity, actual changes in insulin sensitivity appear as artifacts in the net effect trace. Enhanced insulin sensitivity shows up as a tendency toward negative values of net effect. Exercise can show up as negative "meals." Since the model assumes that there is a steady-state plasma insulin concentration associated with the patient's basal profile, large fluctuations in basal rate throughout the day (especially basal rates that are not matched to circadian variability in insulin sensitivity) can show up as artifacts in the net effect trace. Since the gut model is not adapted to the content of the meal, discrepancies in meal rate of appearance show up as transient artifacts in the net effect trace. Similarly, since the insulin transport model is homogenous, discrepancies in insulin pharmacokinetics show up as transient artifacts in the net effect trace. Finally, since endogenous hypoglycemic counter-regulation is not reflected in the model, counter-regulation shows up as an unacknowledged meal.

3 Net Effect Simulation

As discussed above, net effects are computed to an allow for an accurate "simulation" of the patient data, providing the opportunity (i) to replay the patient's treatment day and (ii) to experiment with alternative insulin strategies to see under what circumstances it is possible to achieve better glycemic outcomes.

3.1 "Replay" Simulation

Having computed the net effect estimates

$$\hat{x}(0) \text{ and } \hat{\tilde{\omega}} = (\hat{\omega}(0), \ldots, \hat{\omega}(T-1))$$

from differential CGM data and insulin delivery relative to the basal profile,

$$\tilde{y}_{hist} = (y_{hist}(0), \ldots, y_{hist}(T)) \quad \text{and} \quad \tilde{u}_{hist} = (u_{hist}(0), \ldots, u_{hist}(T-1)), \tag{12}$$

it is possible to simulate a "replay" of that data as follows:

$$x_{replay}(0) = \hat{x}(0) \tag{13}$$

$$x_{replay}(k+1) = Ax_{replay}(k) + Bu_{hist}(k) + G\hat{\omega}(k), \qquad k = 0, \ldots, T-1 \tag{14}$$

with

$$y_{replay}(k) = Cx_{replay}(k) \tag{15}$$

3.2 Simulating Modified Insulin Delivery

With the net effect estimates $\hat{x}(0)$ and $\hat{\tilde{\omega}}$ in hand, it is straightforward to simulate a modified schedule of insulin delivery

$$\tilde{u}_{mod} = (u_{mod}(0), \ldots, u_{mod}(T-1)) \tag{16}$$

as follows:

$$x_{mod}(0) = \hat{x}(0) \tag{17}$$

$$x_{mod}(k+1) = Ax_{mod}(k) + Bu_{mod}(k) + G\hat{\omega}(k), \qquad k = 0, \ldots, T-1 \tag{18}$$

with

$$y_{mod}(k) = Cx_{mod}(k). \tag{19}$$

In simulating a modified insulin regimen care must be taken to ensure that $\tilde{u}_{mod}$ accurately reflects the desired change. For example, if the goal is to simulate the effect of multiplying the patient's preprogrammed basal profile $basal(k)$ (U/hr) by a constant factor α, then (as a rough cut) one could use $\tilde{u}_{mod} =$

$$\max\{\ [-basal(k) \cdot 1000/60], \quad [u_{hist}(k) + (\alpha - 1) \cdot basal(k) \cdot 1000/60]\ \}. \tag{20}$$

Note this assumes that the patient would actually continue to deviate from his/her preprogrammed profile according to $u_{hist}(k)$. That is, if $u_{hist}(k) \neq 0$ represents a bolus in stage k, then the assumption is that the patient would issue the *same* bolus in the modified insulin regime. If the bolus actually includes a correction component, then it is necessary to separate the carb- and correction-related components of u_{hist}, retaining the meal-related component and modifying the correction component according the prevailing (simulated) BG at the time of the bolus. (The meal-related and correction components could be inferred, for example, from the patient's insulin pump record based on the carb-value provided by the patient to the bolus wizard.) Similarly, if $u_{hist}(k) \neq 0$ represents a temporary basal rate, then it may be necessary to adjust $u_{hist}(k)$ in simulating the modified insulin regime to reflect the fact that a temporary basal rate may, or may not, be warranted in stage k.

4 Results

In this section, we present results that (i) illustrate the ability to use net effect replay simulation to reproduce BG variability and (ii) give a preliminary view to the accuracy of net effect simulation in predicting the impact of changes to the patient's basal rate profile. For the former, we present net effects and net effect replays of data collected in a human-subject field study (phase 1 of NIH/NIDDK RO1 DK 085623), and we demonstrate that the net effect replay accurately captures the kinds of BG variability observed in practice. For the latter, we present results from an in silico preclinical trial using the FDA-accepted Virginia/Padova Type 1 simulator, showing that net effect simulation can capture, at least to first order, the impact of basal rate changes. Both studies involved the CGM preprocessing steps and the net effect parameterization described below.

CGM Preprocessing Step 1—Valid Extended Days: Gaps in CGM records from the field are common. While the occasional "missed" sample can easily be accommodated via interpolation, extended intervals of missing CGM data can obscure the picture. With a view to understanding daily BG variability, we introduce the notion of an *extended* CGM day, where the goal is to estimate net effect from midnight to the following midnight, and the last T^- hours of the preceding day are prepended to the record (to deal with initial condition errors) and the first T^+ hours of the following day are appended to the record (to deal with "end-game" effects), as discussed in Sect. 2.2. For the two studies below, we used $T^- = 8$ and $T^+ = 4$, and, for the field study, we considered an extended day to be *valid*, and included it in the study, if there are no gaps longer than three hours and no more than 60 of the normal 5-min samples are missing from the record.

CGM Preprocessing Step 2—Retrofitting: The data from each patient's CGM was retrofitted [2] to smooth the data and force agreement with the available fingerstick values in the historical record. In the retrofit cubic splines were used to smooth the

data and interpolate between gaps in the CGM record. A smoothing parameter of $1 - 10^{-6}$ was used to force the retrofit to give preference to matching second and third derivatives at anchor points (over minimizing the squared difference between the smoothed curve and anchor points). Next, the available fingerstick values were used as reference points, and a linear warping function was applied between samples to force agreement between the smoothed CGM trace and the fingerstick record. (If multiple fingerstick values were recorded within 10 minutes of each other, the sample that is numerically closest to the prevailing CGM value is the one that was used.) Note that the retrofitting procedure used here did not attempt to correct for transients in the CGM record due to calibration events. (Correcting for calibration transients would make the replay simulation more accurate, and this would be important for applications that focus on replaying specific glycemic events. Here, however, our interest is in characterizing BG variability over time timeframe of several weeks, where the effect of calibrations would "average out.")

Net Effect Parameterization: For both studies below we parameterized the net effect calculation as follows. First, given the $T^- = 8$ h of data prepended to the record for each extended day we directly imposed the initial condition estimate $\hat{x}(0) = 0$, relying on the observation that initial condition errors are "forgotten" within the first three to four hours of the extended day. Thus, the calculation deviated slightly from the methodology of Sect. 2.1 (cf. Eq. (6)) in that we took

$$\theta = v, \quad M_{fit} = \mathscr{G}E, \quad \Lambda_{reg} = E'\Lambda_{reg,2}E. \tag{21}$$

With each extended day consisting in $24 + 8 + 4$ h, corresponding to 2160 min, or 432 five-minute intervals, we used the following orthogonal basis:

$$\begin{aligned} e_1 &= (1,\ 1,\ 0,\ 0,\ \ldots 0,\ 0)' \\ e_2 &= (0,\ 0,\ 1,\ 1,\ \ldots 0,\ 0)' \\ &\vdots \\ e_{216} &= (0,\ 0,\ 0,\ 0,\ \ldots 1,\ 1)' \end{aligned} \tag{22}$$

so that each one of the $n = 216$ basis functions corresponds to a 10-min interval of the extended day. The weighting matrices used in the calculations are $\Lambda_{fit} = I_{432}$ and $\Lambda_{reg,2} = \lambda I_{216}$, where $\lambda = 5 \times 10^{-5}$ was chosen (empirically) to reflect a good tradeoff between replay accuracy and net effect smoothing.

4.1 Net Effects and Net Effect Simulation Replay from Field Data

Using de-identified field data collected during Phase 1 of the NIH/NIDDK project RO1 DK 085623, we illustrate (i) net effects derived from the data and (ii) the ability

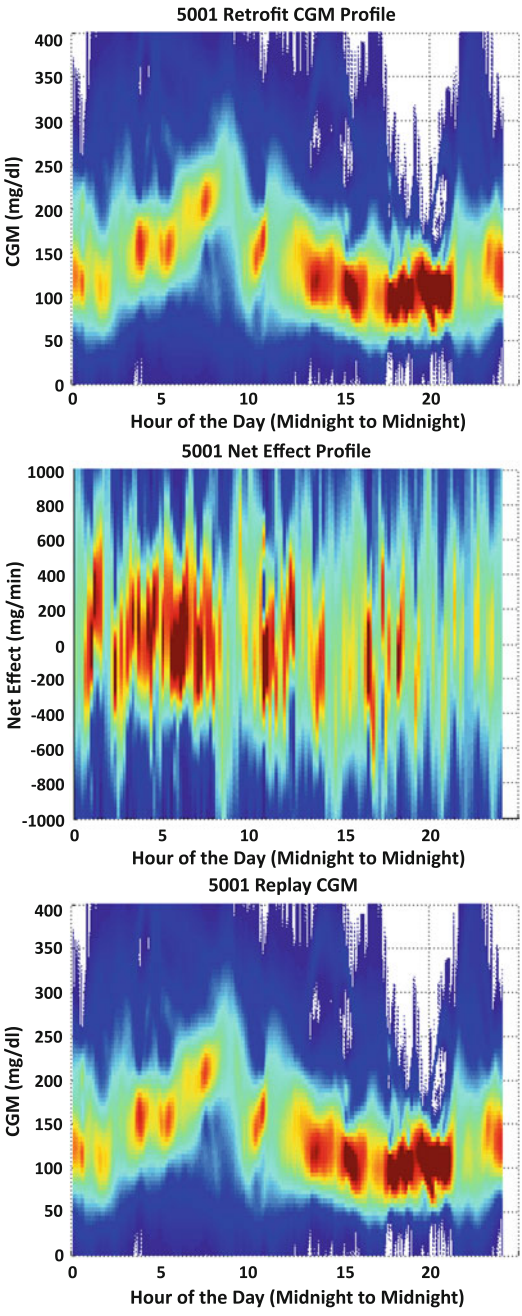

Fig. 4 "Density profile" plots of **a** retrofitted CGM data, **b** associated net effect, **c** replayed CGM data for the valid extended days from subject 5001

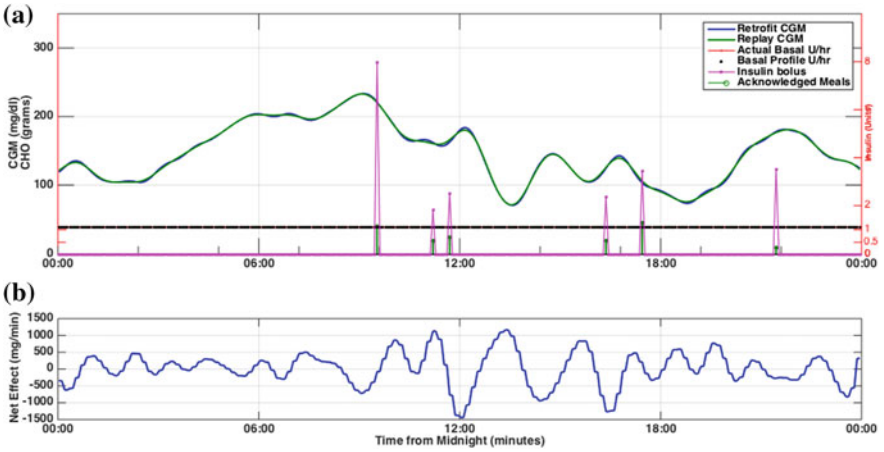

Fig. 5 Example 24-h period from subject 5001: **a** retrofitted CGM, actual basal, profile basal, boluses, acknowledged meals, net effect replay **b** net effect

to reproduce BG variability observed in the field via net effect simulation. The study involved insulin pump users with Type 1 diabetes. Study subjects were asked to wear CGM for a month, simultaneously recording SMBG, CGM and insulin pump data, as well as information about meals and physical activity. Using "valid" extended days from representative Subjects 5001, 5009, and 5041, respectively, Figs. 4, 6 and 8 show "density profile" plots of retrofitted CGM data (left panel), net effect (middle panel), and net effect replay simulation of the data (right panel).[1] Figs. 5, 7, and 9, respectively, show representative 24 h traces for each of the subjects.

4.1.1 Subject 5001

From the top panel of Fig. 4 we observe that Subject 5001 consistently experiences BG values in the range of 125–175 (mg/dl) between 3 and 6 AM, with a consistent trend toward BG above 200 (mg/dl) around 8 AM, and with tight control of BG around 100 (mg/dl) in the afternoon and early evening. The middle panel of Fig. 4 shows the corresponding oral carbohydrate net effect "density profile" plot with a fairly consistent trend to have sharply increasing net effect between 1 AM and 2 AM, corresponding to regular eating late at night (as confirmed by meal-tagging in the patient's insulin pump). Also, from the middle panel, it appears that net effect is generally positively biased between 3 and 7 AM (perhaps due to an insufficient basal rate during the early hours of the night), contributing to the observed elevated

[1] To explain the "density profile," for each five-minute interval of the day (x-axis), we plot a colorized strip representing a kernel density estimate of the empirical distribution function for the corresponding metric (y-axis) in that interval. The spectrum from white/blue to red corresponds to the range of density values from low to high likelihood, respectively.

BG between 3 and 6 AM. The "dispersion" of net effect kernel density at around corresponds to taking breakfast at that time (again confirmed in the meal record), leading to a fairly consistent post prandial peak between 200 and 250 (mg/dl). The "dispersed" nature of the density profile plot of net effect for the remainder of the day suggested that the patient has a flexible eating routine during waking hours. The bottom panel of Fig. 4 shows the BG density profile plot from the net effect replay simulation of the patient's data. The fact that the top and bottom panels of the Fig. 4 closely agree demonstrates that the net effect replay accurate recreates the patient's daily variability of blood glucose.

Figure 5 shows BG, insulin, meal, and net effect traces (lower panel) for a representative day of Subject 5001. The net effect traces has an oscillatory nature, perhaps due to a mismatch of time constants in the meal and insulin transport models relative to the patient's actual dynamics. Note that the net effect trace is generally positive between 01:00 and 08:00, "explaining" the strong trend toward hyperglycemia overnight, a combination of late night snacking (often acknowledged by the patient) and "dawn effect." Note that large positive swings in net effect appear in conjunction with acknowledged meals, though sometimes the net effect swing comes late, perhaps due to slower meal absorption than represented in the SOGMM. This may be the case with with the large negative-then-positive swings in net effect at 12:00 and 13:30. The carbohydrates acknowledged just prior to noon seem to have had little immediate effect on BG, and indeed BG subsequently drops below 80 mg/dl. Here, the rapid drop in BG has to be "explained" by the profoundly negative net effect signal at 12:00, and the subsequent large positive swing may be accounting for delayed absorption of the meal. Alternatively, the large positive swing at 1300 may be the result of counter-regulation (or unannounced hypoglycemia treatment) in response to the precipitous drop in BG.

4.1.2 Subject 5009

From the density profile plot of retrofitted CGM of Fig. 6 (top panel) we see that Subject 5009 consistently experiences hyperglycemia in excess of 200 mg/dl around 7AM and then consistently experiences a rapid drop in BG to below 100 mg/dl by 10:30AM. From the middle panel of Fig. 6, this patient has a generally positively biased net effect overnight, contributing to high blood sugar in the morning. In addition, from the shift in net effect density at around 6AM in the middle panel of Fig. 6, it appears that patient takes breakfast consistently at that time. (Interestingly, the patient's insulin pump shows insulin boluses with meal tags somewhat *later* in the morning, contributing possibly to post prandial BGs above 200 mg/dl.) With the net effect *not* showing an obvious trend to negative values in the morning, it seems that the patient may be compensating for high BG in the morning with exaggerated breakfast time boluses and temporary basal settings above normal. Again, the strong agreement between the top and bottom panels of the Fig. 6 demonstrates that the net effect replay accurately recreates the patient's daily variability of blood glucose.

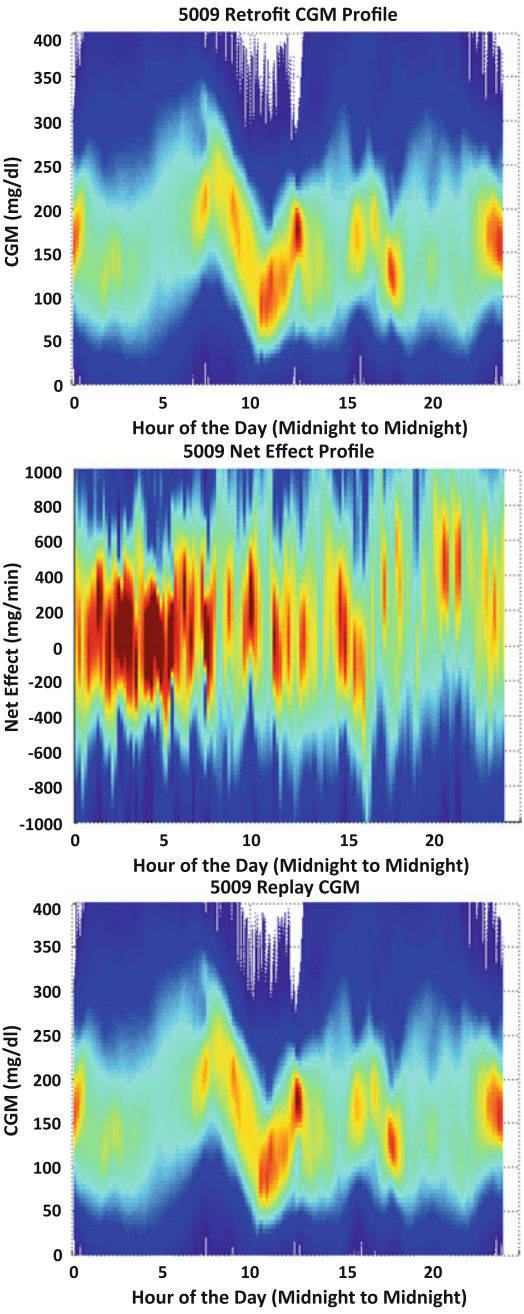

Fig. 6 "Density profile" plots of **a** retrofitted CGM data, **b** associated net effect, **c** replayed CGM data for the valid extended days from subject 5009

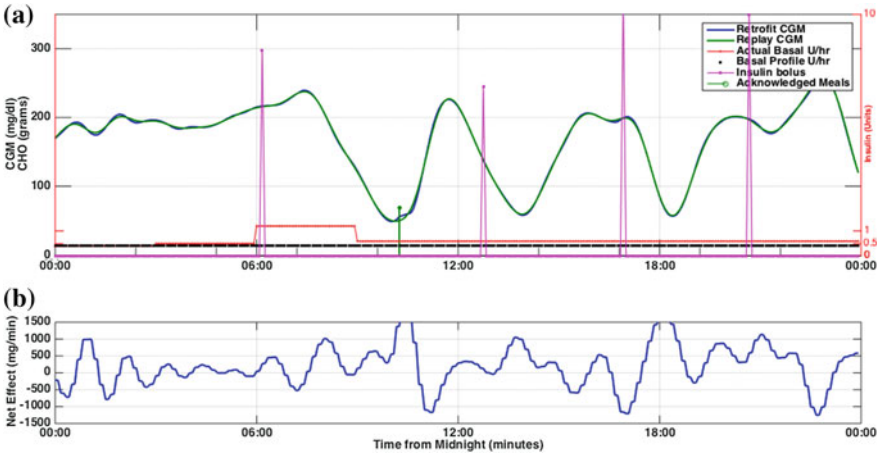

Fig. 7 Example 24-h period from subject 5009: **a** retrofitted CGM, actual basal, profile basal, boluses, acknowledged meals, net effect replay **b** net effect. (Note: only one *acknowledged* meal for the entire day.)

Figure 7 shows BG, insulin, meal, and net effect traces (lower panel) for a representative day of Subject 5009. As with Subject 5001, the net effect plot is oscillatory in nature, with a large positive swing corresponding to the only *acknowledged* meal of the day. The positive swings at around 14:00 and 18:30 probably correspond either to counter-regulation or unannounced meals/hypo treatments.

4.1.3 Subject 5041

The retrofitted CGM density profile plot of Fig. 8 shows that Subject 5041 is aggressively maintaining BG between 50 and 100 mg/dl throughout the day. Compared to Subjects 5001 and 5009, the density profile of net effect throughout the day is relatively "tight," indicating that the patient is engaging in consistent eating and physical activity behaviors, with meals showing up regularly at around 8AM and 7PM, and exercise (trend toward negative net effect) at around 3 PM. Figure 9 shows BG, insulin, meal, and net effect traces (lower panel) for a representative day of Subject 5041, where it is clear that (i) the patient is aggressively treating with insulin overall and (ii) is eating sparingly throughout the day according to the observed shifts in the net effect density profile of Fig. 8.

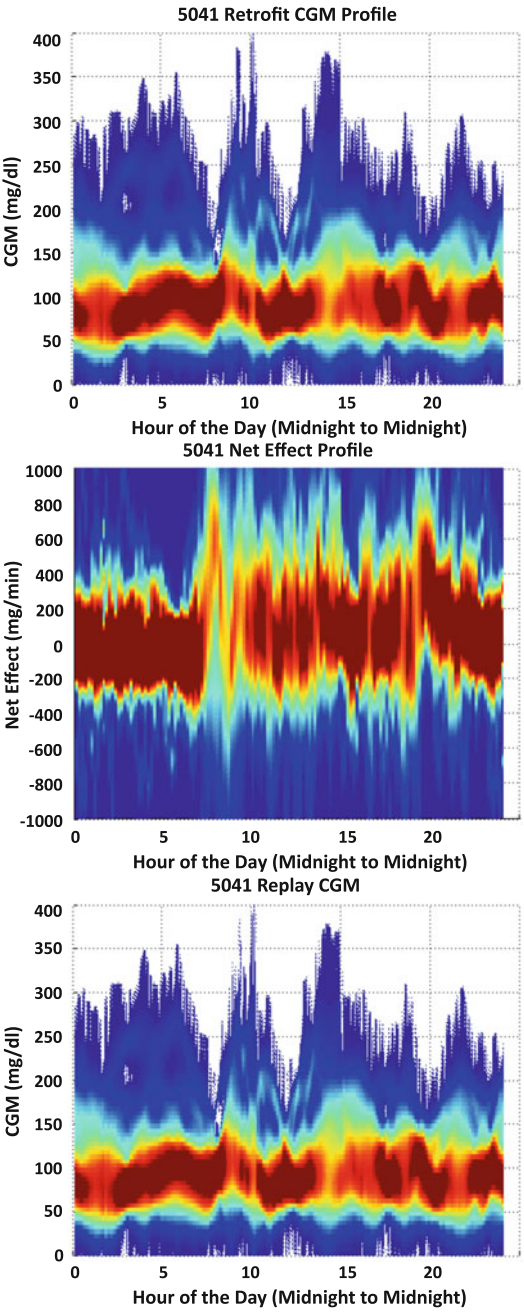

Fig. 8 "Density profile" plots of **a** retrofitted CGM data, **b** associated net effect, **c** replayed CGM data for the valid extended days from subject 5041

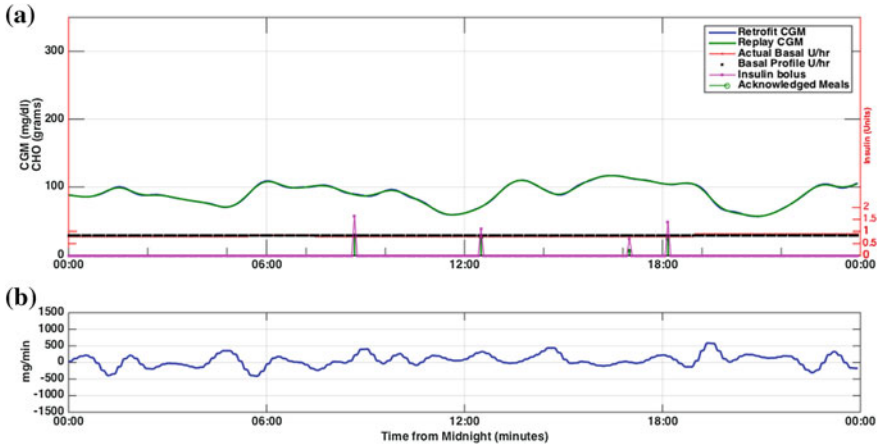

Fig. 9 Example 24-h period from subject 5041: **a** retrofitted CGM, actual basal, profile basal, boluses, acknowledged meals, net effect replay **b** net effect

4.2 In Silico Experiments: Using Net Effect to Design Basal Rate Adjustments

In the preceding section, we demonstrated the ability to accurately reproduce the CGM data of a patient using via net effect replay simulation, cf. Sect. 3.1. To gain a sense of the accuracy of net effect simulation in predicting changes to insulin therapy, we have a set of *in silico* preclinical trials using the adult population of patients in the Virginia/Padova Type 1 Simulator.

Simulated Collection of Field Data and Computation of Net Effect: Each *in silico* patient was simulated for 5 days with three meals a day (with varying meal times and amounts) using a self-treatment strategy that includes (i) detuned basal therapy (resulting in fasting BG between 140 and 180 mg/dl) and (ii) mealtime boluses based on accurate carb counts and patient-specific carbohydrate ratios (no corrections). Here, even though the simulator was set up with zero sensor noise, the same CGP preprocessing steps described above (retrofitting and segmentation into "extended days"). Net effects were calculated using (i) a reference BG computed from the time-average BG value from the 5-day simulation and (ii) patient-specific BW and insulin sensitivity (using TDI and total daily basal).

Computing Modified Basal Rates from the Net Effect Simulation: Based on the techniques of Sect. 3.2, we used the net effect traces above to numerically adjust each patient's basal rate to achieve targeted changes to time-average BG. (The targeted changes were $-20\,\%$, $-10\,\%$, $+10\,\%$, $+20\,\%$, respectively.) To be specific, we computed the constant value of $u_{mod}(k)$ (modulo the carb-related meals boluses that were originally given) needed to achieve the desired change to average BG. This constant offset corresponds to a specific change to the patient's actual basal profile.

Table 1 Accuracy of BG predictions (MARD) for basal rate adjustments designed to achieve given changes in average BG

	–20 % Δ Avg. BG	–10 % Δ Avg. BG	+10 % Δ Avg. BG	+20 % Δ Avg. BG
Regression SI	8.51 (3.93)	4.79 (1.62)	4.87 (2.22)	7.50 (4.82)
Population SI	13.51 (19.88)	5.54 (2.50)	6.13 (3.36)	10.06 (6.73)

Verifying the Basal Change in the Virginia/Padova Type 1 Simulator: Having designed a modified basal profile for each *in silico* subject, we used the Virginia/Padova Simulator to verify that the new basal setting actually achieves the desired effect. Table 1 presents Mean Average Relative Difference (MARD) between the net effect predicted five-day effect of the basal change to the actual effect produced in the Virginia/Padova simulator. The first row of the table shows that we achieve MARD values between 4.79 and 8.51 using the patient-specific insulin sensitivity parameters, with the largest MARD value associated with the attempt to achieve a –20 % change in average blood glucose. To see the importance of using an individualized model for the net effect calculation, we repeated the process using a common insulin sensitivity value (equal to the average individualized insulin sensitivity) for all of the *in silico* subjects. The second row of the table shows that the MARD values increase and become more variable with a population average model.

5 Conclusions

Adjustment of CSII therapy is complicated by variability in daily BG, including variability in eating and exercise behaviors and intra-patient, day-to-day variability in physiological responses. Separating behavioral and physiological factors is difficult. In addition, BG variability observed in field data is far more significant than what is often seen in simulation models. In this work, we have shown that field data can be used to extract BG variability signatures in the form of oral carbohydrate *net effect* as a data-driven extension to the underlying compartmental model. This approach creates the ability to replay real-life scenarios and approximate answers to "what if" questions regarding modified insulin delivery. In Sect. 4 we present numerical results that (i) illustrate the ability to use net effect replay simulation to reproduce BG variability and (ii) give a preliminary view to the accuracy net effect simulation in predicting the impact of changes to the patient's basal rate profile. The fact that we achieve MARD values less than 10 % comparing predicted effect of therapy changes to actual changes in the Virginia/Padova simulator suggests that net effect simulation can be a viable tool for characterizing and addressing BG variability in patients with Type 1 diabetes.

In ongoing work we are continuing to develop the method, including (i) bypassing the gut model and estimating plasma glucose rate of appearance as a net effect, (ii)

meal-informed net effect, where the computation is informed by the patient's estimate of carbohydrates in meals, (iii) using other norms for regularized deconvolution, including mixed-norm methods, (iv) using BG-dependent weights (emphasizing close fit at peaks and valleys), and (v) using nonlinear model inversion. In addition, we are exploring a number of different applications of the approach, including retrospective detection of meals, multiple-hypothesis model predictive control (e.g., [17]), and individualized patient simulation. From our preliminary work we have found that it is sometimes necessary to adapt parameters of SOGMM model (including SI) in an application-specific way, e.g., estimating SOGMM parameters from clinical BG data with precise meal knowledge in order to obtain net effect traces that look more like the impulse train that one would expect from discrete meal events. Generally, we have found that the quality of the net effect replay is dependent upon the accuracy of the underlying compartmental model, and that the "net effect" approach should be seen as *complimentary* to the estimation and use of high fidelity compartmental models.

Acknowledgments This material is based upon work supported by the National Science Foundation under Grant No. CNS0931633, and by the National Institutes of Health under Grant No. NIH/NIDDK RO1 DK 085623. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the sponsors.

APPENDIX A: Mathematical Model

A.1 Subcutaneous Oral Glucose Minimal Model (SOGMM)

For the net effect calculations presented here, we use an extended version of the minimal model [1] known as the Subcutaneous Oral Glucose Minimal Model (SOGMM). As outlined below, the SOGMM includes compartments for oral consumption of carbohydrates, subcutaneous insulin infusion, and subcutaneous measurement of glucose concentration.

Core (Minimal) Submodel—As in the minimal model, glucose-insulin interactions are captured in a two-state compartmental model:

$$\dot{G}(t) = -(S_g + X(t)) \cdot G(t) + S_g \cdot G_b + (R_a(t)/V_g) \tag{23}$$

$$\dot{X}(t) = -p_2 \cdot X(t) + p_2 \cdot S_I(I(t) - I_b) \tag{24}$$

where

1. $G(t)$ (mg/dl) represents the plasma glucose concentration and $R_a(t)$ (mg/kg/min) represents the glucose rate of appearance. G_b (mg/dl) represents "basal" glucose concentration associated with the patient's basal rate of insulin delivery. S_g (1/min) represents the fractional glucose effectiveness measuring glucose abil-

ity to promote glucose disposal and inhibit glucose production, and V_g (kg/dl) represents the distribution volume of glucose.

2. $X(t)$ (1/min) is the proportion of insulin in the remote compartment, $I(t)$ (mU/L) is the plasma insulin concentration, while I_b (mU/L) represents the reference value for $I(t)$ associated with the fasting plasma glucose concentration of the patient. p_2 (1/min) is the rate constant of the remote insulin compartment from which insulin action is emanated, and S_I, (1/min per mU/L), often referred to as insulin sensitivity, represents the ability of insulin to control glucose production and utilization.
3. Glucose rate of appearance $R_a(t)$ can be further related to the second state of the gastrointestinal submodel (see below) by:

$$R_a(t) = (Q_2(t) \cdot k_{abs} \cdot f)/BW, \tag{25}$$

where $Q_2(t)$ is defined below, k_{abs} (1/min) represents a rate constant associated with oral glucose absorption, f (dimensionless) is the fraction of intestinal absorption which actually appears in the plasma, and BW (kg) represents the body weight of the subject.

4. Insulin concentration $I(t)$ can be expressed as a function of $I_p(t)$ (mU), the plasma insulin, as below:

$$I(t) = (I_p(t))/(V_I \cdot BW), \tag{26}$$

where V_I (L/kg) represents the distribution volume of insulin.

Subcutaneous Glucose Measurement Submodel—Interstitial glucose concentration $G_{cgm}(t)$ is modeled as a first order delay from the plasma glucose concentration $G(t)$:

$$\dot{G}_{cgm}(t) = -k_{sc}(G_{cgm}(t) - G(t)), \tag{27}$$

where k_{sc} (1/min) is the time constant that encompasses both the physiological lag and the sensor lag.

Gastrointestinal Submodel—Transport of oral carbohydrates are expressed with a two-compartment "gut" model:

$$\dot{Q}_1(t) = -k_\tau \cdot Q_1(t) + \omega(t) \tag{28}$$

$$\dot{Q}_2(t) = -k_{abs}Q_2(t) + k_\tau \cdot Q_1(t), \tag{29}$$

where $Q_1(t)$ (mg) and $Q_2(t)$ (mg) are the two compartments represent the oral glucose transport, k_τ (1/min) is a rate constant associated with oral glucose absorption, and $\omega(t)$ (mg/min) is the rate of mixed-meal carbohydrate absorption at time t.

Subcutaneous Insulin Kinetic Submodel—Transport of insulin from subcutaneous infusion (via CSII) to plasma is expressed with three compartments:

$$\dot{I}_{sc1}(t) = -k_d \cdot I_{sc1}(t) + J_{ctrl}(t) \tag{30}$$

$$\dot{I}_{sc2}(t) = -k_d \cdot I_{sc2}(t) + k_d I_{sc1}(t) \tag{31}$$

$$\dot{I}_p(t) = -k_{cl} \cdot I_p(t) + k_d I_{sc2}(t), \tag{32}$$

where $I_{sc1}(t)$ (mU) and I_{sc2} (t) (mU) represent compartments related to the interstitial insulin transport, k_d (1/min) and k_{cl} (1/min) represent rate constants of subcutaneous insulin transport, J_{ctrl} (t) (mU/min) represents the insulin input signal, and $I_p(t)$ (mU) represents the plasma insulin.

Linearization and Discretization

For computational efficiency, our computation of net effect is computed using a discrete-time linearized approximation of the SOGMM, which ultimately allows net effect to be computed in closed form in a relatively computationally tractable manner.

First the model is linearized about basal plasma insulin I_b and glucose concentration G_b to get the continuous-time LTI model:

$$\dot{x}_c(t) = A_c x_c(t) + B_c u_c(t) + G_c \omega(t) \tag{33}$$

$$y_c(t) = C_c x_c(t) \tag{34}$$

where $x_c(t)$ represents the patient's metabolic state vector:

$$x_c(t) = \left(\partial G(t),\ \partial X(t),\ \partial I_{sc1}(t),\ \partial I_{sc2}(t),\ \partial I_p(t),\ \partial G_{sc}(t),\ \partial Q_1(t),\ \partial Q_2(t)\right)' \tag{35}$$

whose elements are all relative to the operating point described by G_b and I_b, and where

1. *differential insulin input:* $u_c(t)$ (mU/min) is the rate of subcutaneous insulin delivery relative to the patient's preprogrammed basal rate profile, i.e., $u_c(t) = J_{ctrl}(t) - J_{basal}$. (That is, $J_{ctrl}(t)$ (mU/min) is the rate of insulin infusion at time t, and J_{basal} is the rate of insulin delivery (mU/min) needed to achieve basal plasma insulin I_b.)
2. *differential subcutaneous glucose:* $y_c(t)$ is derived from $x_c(t)$ as the subcutaneous glucose concentration relative to G_b.

The elements of the state vector $x_c(t)$ are:

1. *differential blood glucose concentration:* $\partial G(t) = G(t) - G_b$ (mg/dl), where basal plasma glucose concentration G_b serves as the reference value G_{ref} for glucose concentration,
2. *differential remote compartment insulin action:* $\partial X(t) = X(t) - X_{ref}$ (1/min), where $X_{ref} = 0$ (we have "insulin action" ultimately when plasma insulin deviates from I_b, i.e., $X(t) = 0$ is the basal condition of no insulin action),

3. *differential interstitial insulin, first compartment:* $\partial I_{sc1}(t) = I_{sc1}(t) - I_{sc1,ref}$ (mU), where $I_{sc1,ref}$ is the steady-state interstitial insulin associated with I_b,
4. *differential interstitial insulin, second compartment:* $\partial I_{sc2}(t) = I_{sc2}(t) - I_{sc2,ref}$ (mU), where $I_{sc2,ref}$ is the steady-state interstitial insulin associated with I_b,
5. *differential plasma insulin:* $\partial I_p(t) = I_p(t) - I_b$ (mU), where basal plasma insulin I_b serves as the reference value for plasma insulin,
6. *differential interstitial glucose concentration:* $\partial G_{sc}(t) = G_{sc}(t) - G_{sc,ref}$ (mg/dl), where $G_{sc,ref} = G_b$ is the reference value of interstitial glucose concentration,
7. *differential gut, first compartment*: $\partial Q_1(t) = Q_1(t) - Q_{1,ref}$ (mg), where $Q_{1,ref} = 0$ is the reference value of glucose stored in the first of the two gut compartments,
8. *gut, second compartment*: $\partial Q_2(t) = Q_2(t) - Q_{2,ref}$ (mg), where $Q_{2,ref} = 0$ is the reference value of glucose stored in the second of the two gut compartments.

The coefficient matrices of the linearized model are given as

$$A_c = \begin{bmatrix} -S_G & -G_b & 0 & 0 & 0 & 0 & 0 & \frac{k_{abs} \cdot f}{BW \cdot V_G} \\ 0 & -p_2 & 0 & 0 & \frac{p_2 \cdot S_I}{V_I \cdot BW} & 0 & 0 & 0 \\ 0 & 0 & -k_d & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_d & -k_d & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & k_d & -k_{cl} & 0 & 0 & 0 \\ k_{sc} & 0 & 0 & 0 & 0 & -k_{sc} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -k_\tau & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & k_\tau & -k_{abs} \end{bmatrix} \tag{36}$$

$$B_c = [0\,0\,1\,0\,0\,0\,0\,0]', \quad G_c = [0\,0\,0\,0\,0\,0\,1\,0]', \quad C_c = [0\,0\,0\,0\,0\,1\,0\,0], \tag{37}$$

where it is important to note that plasma basal insulin I_b does not appear.

As a practical matter, we assume that the patient's basal rate profile $J_{basal}(t)$ is perfectly adapted to the patient's need for insulin throughout the day, so that $u_c(t) = J_{ctrl}(t) - J_{basal}(t) = 0$ leads to G_b as a steady-state value. The numerical value of G_b is taken to be the approximate glucose concentration associated with the patient's most recent HbA1c (glycated hemoglobin) reading:

$$G_b = \text{HbA1c} \cdot 28.7 - 46.7. \tag{38}$$

Thus, given the discussion of Sect. 2, if the patient's insulin sensitivity is such that basal insulin alone leads to a steady-state glucose concentration not equal to G_b, then this would be explained by a net effect estimate with an offset away from zero.

The LTI model is discretized using the standard zero-order hold method to obtain the discrete-time model of Eqs. (1) and (2).

Table 2 Fixed parameters of the SOGMM

$S_G = 0.01000$ (1/min)	$V_G = 1.6000$ (kg/dl)	$p_2 = 0.02000$ (1/min)
$k_{abs} = 0.01193$ (1/min)	$f = 0.90000$ (dimensionless)	$V_I = 0.06005$ (L/kg)
$k_{sc} = 0.09088$ (1/min)	$k_\tau = 0.08930$ (1/min)	$k_d = 0.02000$ (1/min)
$k_{cl} = 0.16000$ (1/min)		

Parameters

All but two of the parameters of the SOGMM are fixed at population values, as listed in Table 2. Body weight (BW) and insulin sensitivity S_I (1/min per mU/L) are the only individualized parameters.

Here S_I is computed according to the regression formula below using (i) the patient's average Total Daily Insulin (TDI_{whole}) computed from overall insulin utilization over the timeframe of one week of open-loop data collection and (ii) the patient's average Total Daily Basal Insulin (TDI_{basal}):

$$S_I = \exp(-6.4417 - 0.063546 * TDI_{whole} + 0.057944 * TDI_{basal}) \quad (39)$$

The regression formula above was derived from a set of experiments with the adult *in silico* population in the FDA-accepted Virginia/Padova Type 1 Simulator. Using the parameters of each individual *in silico* patient, we first computed the two steady-state BG values associated with (i) a nominal basal profile and then (ii) a modified basal profile. Next, we computed the value of the SI parameter of the linearized and discretized SOGMM model so that the effect of the change in basal would be reproduced as a constant insulin input relative to the baseline operating point. After computing individual SI values in this way, we performed a stepwise log-linear regression using TDI_{whole}, TDI_{basal} and other statistics from one simulated week of nominal therapy, ultimately resulting in Eq. (39).

References

1. Bergman, R.N., Ider, Y.Z., Bowden, C., Cobelli, C.: Quantitative estimation of insulin sensitivity. Am. J. Physiol. **236**, E667–E677 (1979)
2. Breton, M.D., Kovatchev, B.P.: Impact of blood glucose self-monitoring errors on glucose variability, risk for hypoglycemia, and average glucose control in type 1 diabetes: an in silico study. J. Diab. Sci. Tech. **4**, 562–570 (2010)
3. Carson, E., Cobelli, C.: Modeling Methodology for Physiology and Medicine, 2nd Ed. Elsevier (2013)
4. Chiang, J.L., Kirkman, M.S., Laffel, L.M.B., Peters, A.L.: Type 1 diabetes through the life span: a position statement of the american diabetes association. Diabet. Care 37(7), 2034-2054 (2014)

5. Dalla Man, C., Micheletto, F., Lv, D., Breton, M., Kovatchev, B., Cobelli, C.: The UVA/PADOVA type 1 diabetes simulator: new features. J. Diabet. Sci. Tech. 8(1), 26-34 (2014)
6. Davidson, P.C., Hebblewhite, H.R., Steed, R.D., Bode, B.W.: Analysis of guidelines for basal-bolus insulin dosing: basal insulin, correction factor, and carbohydrate-to-insulin ratio. Endocr. Pract. **14**(9), 1095–1101 (2008)
7. Diabetes Research In Children Network (DirecNet) study group: use of the direcnet applied treatment algorithm (DATA) for diabetes management with a real-time continuous glucose monitor (the FreeStyle Navigator. Pediat. Diabet. 9, 142-147 (2008)
8. Duron, F.: Intensive insulin therapy in insulin-dependent diabetes mellitus, the results of the diabetes control and complications trial. Biomed. Pharmacother. **49**(6), 278–282 (1995)
9. Haidar, A., Wilinska, M.E., Graveston, J., Hovorka, R.: Stochastic virtual population of subjects with type 1 diabetes for the assessment of closed-loop glucose controllers. IEEE Trans. Biomed. Eng. **60**(12), 3524–3533 (2013)
10. Hovorka, R., Chappell, M.J., Godfrey, K.R., Madden, F.N., Rouse, M.K., Soons, P.A.: Code: a deconvolution program implementing a regularisation method of deconvolution constrained to non-negative values. description and pilot evaluation. Biopharm. Drug. Dispos. **19**(1), 39–53 (1998)
11. Hughes, C.S., Patek, S.D., Breton, M., Kovatchev, B.P.: Anticipating the next meal using meal behavioral profiles: A hybrid model-based stochastic predictive control algorithm for T1DM. Comput. Methods Progr. Biomed. **102**(2), 138–148 (2011)
12. Hughes, C.S., Patek, S.D., Taylor, A., Park, R., Laubscher, K., Pillutla, S., Breton, M., Kovatchev, B.P.: In-silico experiments of human eating patterns: Making a case for behavior-informed control of T1DM. In: Proceedings of the UKACC International Conference of Control (2010)
13. Hughes-Karvetski, C., Patek, S.D., Breton, M.D., Kovatchev, B.P.: Historical data enhances safety supervision system performance in t1dm insulin therapy risk management. Comput. Methods Progr. Biomed. **109**(2), 220–225 (2012)
14. Hughes-Karvetski, C., Patek, S.D., Gonder-Frederick, L., Kovatchev, B.: Interactive internet intervention providing individually-tailored glycemic feedback based on cgm, insuin and meal data, and computer simulation: Algorithmic description. Abstract and poster presented at the Advanced Technologies & Treatments for Diabetes, Paris (2013)
15. King, A.B.: Continuous glucose monitoring-guided insulin dosing in pump-treated patients with type 1 diabetes: a clinical guide. J. Diabet. Sci. Tech. **6**(1), 191–203 (2012)
16. Kovatchev, B., Breton, M., Dalla Man, C., Cobelli, C.: In Silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes. J. Diabet. Sci. Tech. 3, 44-55 (2009)
17. Patek, S.D.: Open-loop feedback control under multiple disturbance function hypotheses. In: IEEE Conference on Decision and Control, pp. 4165-4170 (2010)
18. Patek, S.D., Bequette, B.W., Breton, M., Buckingham, B.A., Dassau, E., Doyle III, F.J., Lum, J., Magni, L., Zisser, H.: In Silico preclinical trials: methodology and engineering guide to closed-loop control in type 1 diabetes mellitus. J. Diabet. Sci. Tech. 3, 269-282 (2009)
19. Scheiner, G., Boyer, B.A.: Characterstics of basal insulin requirements by age and gender in type-1 diabetes patients using insulin pump therapy. Diabet. Res. Clin. Pract. **69**, 14–21 (2005)
20. Sorensen, J.T.: A physiologic model of glucose metabolism in man and its use to design and assess improved insulin therapies for diabetes. Ph.D. thesis, Department of Chemical Engineering, MIT (1985)

Physiology-Based Interval Models: A Framework for Glucose Prediction Under Intra-patient Variability

Jorge Bondia and Josep Vehi

Abstract In recent years interval models, i.e., models with interval parameters, have been proposed in literature introducing new methodological approaches for glucose prediction. In this paper, a review of physiological interval models applied to the prediction of blood glucose in type 1 diabetes glycemic control is presented. Predicting blood glucose for diabetic patients is a difficult challenge mainly due to intra-patient variability and uncertainty which may jeopardize model individualization, both for data-based and physiological models. Interval models provide a theoretical framework to express the imprecision and the uncertainty related to complex systems. In the context of physiological systems, interval models can represent intra-patient variability by means of interval parameters. In this paper, interval models are introduced as well as methods for simulation. The result of an interval simulation is an envelope, which can be computed efficiently and accurately with modal interval analysis and monotone systems theories. Finally, as model predictions are as good as the individual model itself, interval model identification methods are also introduced for model individualization.

1 Introduction

Type 1 diabetes is a metabolic disease characterized by an absolute deficiency of insulin. This leads to high blood glucose levels (hyperglycemia) and has deleterious effects such as ketoacidosis and chronic complications like microangiopathy leading to blindness and renal failure. Currently, patients follow a substitutive insulin therapy, either with multiple daily injections (MDI) or continuous subcutaneous infusion with an insulin pump (CSII). However, insulin dosing is still an empirical process

J. Bondia (✉)
Instituto Universitario de Automática e Informática Industrial, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain
e-mail: jbondia@isa.upv.es

J. Vehi
Institut D'Informàtica i Aplicacions, Universitat de Girona, Campus Montilivi, Edifici P4, 17071 Girona, Spain

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*, Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_9

based on the patient's auto-control from 3 to 4 capillary glucose measurements per day from a glucometer (SMBG). This frequently results in under- or overdosing of insulin, causing, respectively, hyper- and hypoglycemia. Hypoglycemia may lead to confusion, falls, seizures, coma, or even death. Nowadays glycemic control objectives are still unmet with an average exposure to hypoglycemia (<70 mg/dL) above 1 h/day and to hyperglycemia (>180 mg/dL) above 9 h/day, besides the prevalence of large glycemic variability [29].

Continuous glucose monitors (CGM) provide glucose readings every 1–5 min. Recent technological advances in CGM have fostered new approaches in type 1 diabetes management [24]. Examples are the recent launch of the first combo CGM–insulin pump with a hypoglycemia prediction system driving automatic pump shutdown [39] and the very active research field of closed-loop glucose control, coined as "the artificial pancreas" (AP) [11], with widespread use of models for glucose prediction both for controllers design and validation and as a controller component itself (e.g., Model Predictive Control-based systems).

However, effectiveness of any model-based approach depends on the accuracy of the model predictions for a given individual. Model individualization has been proved difficult for both data-based [12, 44, 52] and physiology-based models [43]. Proposed strategies for optimal experiment design [13, 33] yield to complex protocols of difficult practical application in an ambulatory context where patient's compliance is at stake. Two main barriers have been identified to model individualization:

1. CGM devices still have suboptimal accuracy, especially in hypoglycemia [38], with significant measurement error/noise that may jeopardize ambulatory model identification.
2. High variability in the patient's behavior is observed in the clinical practice [48] due to circadian rhythms, altered metabolic states, stress, illness, and other unknown sources.

CGM accuracy is a technological barrier that will be overcome with new generations of devices. Indeed, significant advances are being done in this sense with new increased accuracy devices [9, 14, 53]. However, patient's variability is a physiological barrier that cannot be eliminated. Significant variability has been reported in subcutaneous insulin absorption of insulin aspart, with an intra-subject coefficient of variation for the time-to-peak plasma insulin concentration of 27 % in subjects with type 1 diabetes [20]. Nearly 40 % of this variance was attributed to inter-occasion variability (variations in depth of cannula, infusion site age, local tissue perfusion). In [48], unexpected responses are reported as higher postprandial glucose area under the curve for significantly higher insulin prandial dosing in a same patient, which apparently would be classified as 'non-physiological." This illustrates the big challenge of model individualization.

Intra-subject variability can be mathematically represented with model parametric uncertainty. However, it has not generally been integrated into the identification process, leading to average models with poor predictive capabilities. In recent years interval models (i.e., models with interval parameters) have been proposed in literature introducing new methodological approaches for glucose prediction, both for

physiological models [34, 35] and data-driven models [28]. In this case predictions are given in terms of a glucose *envelope* describing the family of possible glycemic responses according to the patient's observed variability from training data. Clinical validations of such methods, although limited, are promising with robust predictions 4–5 h after a meal, as compared to the 30–90 min used in current approaches. The area of interval models in control is not new (see for instance [26]). However, demonstration of its feasibility in the diabetes field opens an exciting research on robust glycemic control, hypoglycemia risk prediction and fault tolerance to name a few applications, which are important components for a safe and efficient artificial pancreas.

In this work a review of the fundamentals of interval models is presented. General concepts will be presented in Sect. 2. At the core of interval methods is interval simulation, i.e., how to generate efficiently and accurately, with mathematical guarantee, the output envelope for a given interval model. Different approaches will be introduced for this purpose from the areas of interval analysis and monotone systems in Sect. 3 and illustrated with the Bergman glucose–insulin model [4] in Sect. 4. Finally, methods for interval model identification will be introduced in Sect. 5.

2 Interval Models

Complex systems are often subjected to uncertainties that make such a model difficult if not impossible to obtain. These uncertainties can be unstructured (the equations of the model are not entirely known) or structured (the equations are known but not the values of their parameters) and they arise from the following reasons:

- the knowledge of the system is not complete, e.g., because the real system cannot be observed as it frequently happens in biological systems;
- the model of the system is known but it is too complex and a simplified one is more appropriate for the task to be undertaken;
- the parameters of the model can change across time due to unknown, unpredictable, or difficult to model phenomena.

When the uncertainties are structured, i.e., the model structure is known and only the parameters undergo imprecisions, they can be handled with *interval models* in which model parameter values are allowed to vary within numeric intervals. Thus, interval models represent a *family* of quantitative (standard) models. Henceforth the following notation will be used: $X = [\underline{x}, \overline{x}]'$ is a real interval with $\underline{x}$ and $\overline{x}$ as left and right endpoints respectively; $\mathbf{X} = (X_1, \ldots, X_n)$ is an n-dimensional interval vector (or box); $I(\mathbb{R}^n)$ denotes the set of all n-dimensional interval vectors and $I(\mathbb{R})$ the set of all real intervals.

Such an interval model could be given by a transfer function with interval coefficients which represents a linear differential relation between an input $u(t)$ and an output $y(t)$:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{B_m s^m + \cdots + B_1 s + B_0}{A_n s^n + \cdots + A_1 s + A_0} \tag{1}$$

in which s is the Laplace variable, $U(s)$ and $Y(s)$ are the input and output Laplace transforms, respectively, and $A_i, B_j \in I(\mathbb{R}), i = 0, \ldots, n, j = 0, \ldots, m$. In its general form transfer function coefficients may be a function of physical interval parameters $\mathbf{P} \in I(\mathbb{R}^{n_p})$, with n_p the number of parameters

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_m(\mathbf{P}) s^m + \cdots + b_1(\mathbf{P}) s + b_0(\mathbf{P})}{a_n(\mathbf{P}) s^n + \cdots + a_1(\mathbf{P}) s + a_0(\mathbf{P})}. \tag{2}$$

This will be the case when the transfer function originates from a description of the system in differential equations with uncertain physical parameters. However, methods become more involved. Coefficient functions $a_i(\cdot), b_j(\cdot), i = 1, \ldots, n, j = 1, \ldots, m$, may depend linearly, multi-linearly or nonlinearly on $\mathbf{P}$. This distinction will be important when evaluating for instance the complex plane image of the interval model (value set) or in stability analysis [5].

Uncertain state space models can be represented in the same way

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t); \mathbf{P}), \quad \mathbf{x}(0) = \mathbf{X}_0 \\ \mathbf{y}(t) &= \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t); \mathbf{P}) \end{aligned} \tag{3}$$

where $\mathbf{P} \in I(\mathbb{R}^{n_p})$ is a vector of n_p interval parameters and $\mathbf{X}_0 \in I(\mathbb{R}^n)$ the vector of interval initial states, $\mathbf{x} \in \mathbb{R}^n$ is the state vector, $\mathbf{u} \in \mathbb{R}^{n_u}$ is the input vector and $\mathbf{y} \in \mathbb{R}^{n_y}$ is the output vector. As compared to interval transfer functions, interval state space models will generally retain the physical meaning of the uncertain parameters. They will be preferred in this work, leading to interval physiological models for glucose prediction.

In some cases an overbounding of the family may be carried out to reduce complexity, for instance in linear state space models with interval matrices computed from the interval evaluation of each of the entries

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \quad \mathbf{x}(0) = \mathbf{X}_0 \tag{4}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u} \tag{5}$$

where $\mathbf{A} := (A_{ij}(\mathbf{P})), A_{ij} : I(\mathbb{R}^{n_p}) \to I(\mathbb{R})$ and equivalently for $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$. However, this will lead to conservativeness.

3 Simulating Interval Models

An interval model represents a whole set or family of models. The behavior of interval models are represented by an envelope trajectory (*envelope* for short) that contains the set of temporal trajectories for each element in the family. These element-wise temporal trajectories are indiscernible, so that the qualitative features may not be captured, particularly if more than one element contributes to the envelope bounding curves as in Fig. 1.

The ultimate goal is to generate a complete and sound envelope, i.e., the *exact envelope*. However, more realistic goals are to produce either a minimally overbounding envelope or a minimally underbounding envelope, depending on the use that one wants to make of it. In the case of glucose prediction minimally overbounding envelopes (or exact when possible) are generally preferred for the sake of robustness and patient's safety in therapeutical actions derived from the prediction.

Another desirable property of an envelope is *stability*. If the envelope width grows with time, as the input signal remains the same, it is unstable. This means that the imprecision accumulates along the propagation through time. This is obviously an undesirable property, because an unstable envelope becomes useless at some time point.

Methods for the generation of envelopes include the superposition of a threshold (fixed or adaptive) to a nominal system simulation [27, 46], the simulation of a discrete subset of elements in the family (taken randomly as in Monte Carlo or from a parameter space grid) [31] or the solution of a global optimization problem in the parameter space [21]. However they generally lead to underbounding due to the

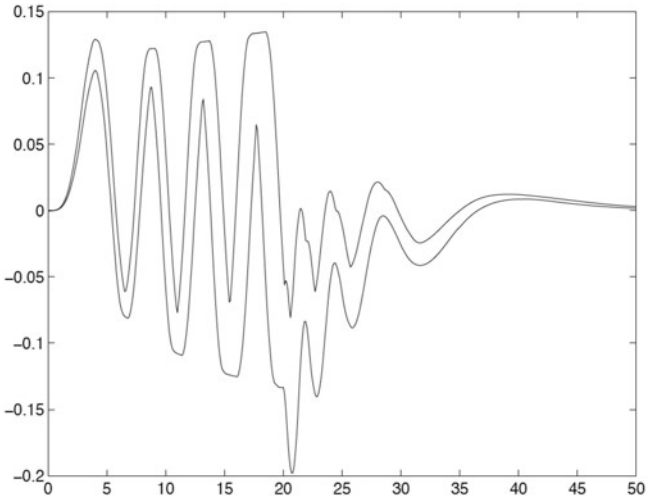

Fig. 1 The envelope is bounded by two curves

lack of guarantee of completeness. The reader is referred to [2, 47] for a review of simulators.

Methods with mathematical guarantee of completeness leading to exact or minimally overbounding envelopes are *interval analysis* and *monotone systems* theories. These methods are reviewed next. They will be applied in Sect. 4 to build guaranteed glucose predictors of a literature glucose–insulin model as illustration.

3.1 Interval Analysis

3.1.1 Classic Interval Analysis

Interval analysis [40, 41] allows to consider the whole continuous range of possible instances represented by an interval model; it hence guarantees that the envelopes are not underbounded. Interval analysis identifies the intervals with the set of real numbers they contain:

$$X = \left[\underline{x}, \overline{x}\right]' = \left\{x \in \mathbb{R} \mid \underline{x} \leq x \leq \overline{x}\right\} \tag{6}$$

Then, the set of closed intervals of $\mathbb{R}$ is:

$$I(\mathbb{R}) = \{[a, b]' \mid a, b \in \mathbb{R}, \quad a \leq b\} \tag{7}$$

In the classic set theory interval analysis, given a $\mathbb{R}^n$ to $\mathbb{R}$ continuous function $z = f(x_1, \ldots, x_n)$, the *interval united extension* R_f of f corresponds to the range of f-values on its interval argument $X = (X_1, \ldots, X_n)$ in $I(\mathbb{R}^n)$:

$$\begin{aligned} R_f(X_1, \ldots, X_n) &= \{f(x_1, \ldots, x_n) \mid x_1 \in X_1, \ldots, x_n \in X_n\} \\ &= [\min\{f(x_1, \ldots, x_n) \mid x_i \in X_i\}, \max\{f(x_1, \ldots, x_n) \mid x_i \in X_i\}] \quad i = 1, \ldots, n, \end{aligned} \tag{8}$$

which can be considered as a semantic extension of f, since it admits the logical interpretation:

$$(\forall x_1 \in X_1) \cdots (\forall x_n \in X_n)(\exists z \in R_f(X_1, \ldots, X_n))\, z = f(x_1, \ldots, x_n). \tag{9}$$

This logical interpretation represents the set of all trajectories that verify the model equation.

Since the interval united extension of a general rational function is generally not computable, set theory *interval rational extensions* $fR(X_1, \ldots, X_n)$, also denoted as natural extensions, are defined in a similar manner to their corresponding real rational

functions $f(x_1, \ldots, x_n)$ replacing their numerical arguments $x_1, \ldots, x_n$ by the interval arguments $X_1, \ldots, X_n$ and their "real" arithmetic operators by their corresponding interval operations, so that

$$R_f(X_1, \ldots, X_n) \subseteq fR(X_1, \ldots, X_n), \tag{10}$$

where $fR(X_1, \ldots, X_n)$ is computable from the bounds of the intervals $X_1, \ldots, X_n$, and usually represents an overestimation of $R_f(X_1, \ldots, X_n)$.

As a consequence, rational extension is very useful for computing the range of a function because it guarantees the result. None of the functions in the class can take values outside the range computed.

Unfortunately, interval arithmetic does not provide the exact result (complete and sound) in the general case. This comes essentially from the two following problems [41]:

- *The multi-incidence problem*. Interval arithmetic considers that each incidence of a variable in a function is independent of each other. Therefore, when there are multi-incident variables, i.e., variables that appear in several leaves of the syntactical tree of the function, the compiled range is overbounded. Similarly, interval arithmetic is unable to take into account dependencies or relations between variables and obtains overbounded results when they exist. This problem makes for instance that in classical interval arithmetic $X - X \neq [0, 0]'$ and $X/X \neq [1, 1]'$ (for instance, $[1, 2]' - [1, 2]' = [-1, 1]'$ and $[1, 2]'/[1, 2]' = [1/2, 2]'$).
- *The wrapping problem*. This is a problem that appears when interval arithmetic is used for simulation in the state space. The state, at some time point, of a model with interval parameters may be represented by a hypercube. However, it is possible that this state of the system does not evolve into another hypercube at the next time point. In Fig. 2 an example with two state variables is shown: the hypercube is a rectangle. It transforms into a rhombus at the following time step (it could actually evolve into any two dimensional shape). The projection of this rhombus on the

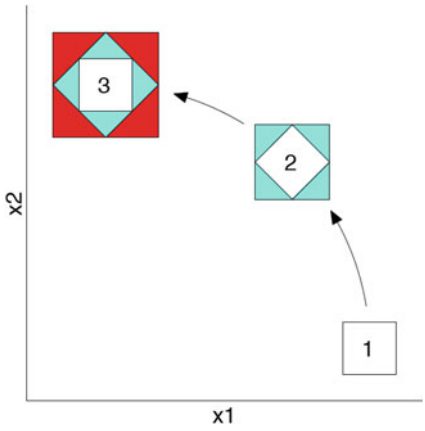

Fig. 2 The wrapping problem

variable axis imposed by the interval representation leads to a new rectangle which obviously includes spurious states, as shown shadowed in the figure. Therefore, the obtained envelopes are overbounded.

Some simulators use classical one-step-ahead numeric integration algorithms revised for intervals [47]. The Moore's interval simulator [41] obtains very overbounded envelopes because it simply ignores multi-incidences. Though, the result can be tightened by splitting the parameter space into smaller subspaces [41] to the detriment of computational burden. Several methods have been proposed to avoid the wrapping effect, or at least to reduce it, since it was first observed in the early 1960s [40]. These methods include a change of coordinates [40], the use of Taylor models [37], or a QR factorization [36] to rotate the state space of the interval system, as well as the use of new computer set representations like ellipsoids [42] or zonotopes [30].

3.1.2 Modal Interval Analysis

Modal interval analysis (MIA) [19, 49] allows the computation of a tight (sometimes exact) enclosure of the envelope that includes all possible behaviors of the system when given in a discretized form [3]. MIA is an algebraic and logical completion of classic intervals. A modal interval X is defined as a couple $X = (X', Q)$, where X' is its classic interval domain, $X' \in I(\mathbb{R})$, and Q is one of the classical quantifiers $\forall$ or $\exists$. The set of the modal intervals is represented by $I^*(\mathbb{R})$. Modal intervals of the type $X = (X', \exists)$ are called *proper intervals* and modal intervals of the type $X = (X', \forall)$ are called *improper intervals*. A modal interval can be represented using canonical coordinates in the form:

$$X = [a, b] = \begin{cases} ([a, b]', \exists) & \text{if } a \leq b \\ ([b, a]', \forall) & \text{if } a \geq b. \end{cases} \tag{11}$$

For example, the interval $[1, 3]$ is equal to $([1, 3]', \exists)$ and the interval $[3, 1]$ to $([1, 3]', \forall)$.

A convenient symmetry between proper and improper intervals is established by the duality operator, which changes the modality of a modal interval. Thus, $\text{Dual}([a, b]) = [b, a]$.

The lattice operations "meet" and "join" on $I^*(\mathbb{R})$ for a bounded family of modal intervals $A(I) := \{A(i) = [a_1(i), a_2(i)] \in I^*(\mathbb{R}) \mid i \in I\}$ (I is the index domain) are defined as a function of the interval bounds:

$$(meet) \quad \wedge_{i \in I} A(i) = \left[\max_{i \in I} a_1(i), \min_{i \in I} a_2(i) \right] \tag{12}$$

$$(join) \quad \vee_{i \in I} A(i) = \left[\min_{i \in I} a_1(i), \max_{i \in I} a_2(i) \right]. \tag{13}$$

In the context of modal intervals, the optimal modal interval extensions of a function f are the semantic $*$ and $**$-functions, denoted by f^* and f^{**}, and defined as:

$$f^*(X) := \bigvee_{x_p \in X'_p} \bigwedge_{x_i \in X'_i} [f(x_p, x_i), f(x_p, x_i)] = \left[\min_{x_p \in X'_p} \max_{x_i \in X'_i} f(x_p, x_i), \max_{x_p \in X'_p} \min_{x_i \in X'_i} f(x_p, x_i) \right] \tag{14}$$

and

$$f^{**}(X) := \bigwedge_{x_i \in X'_i} \bigvee_{x_p \in X'_p} [f(x_p, x_i), f(x_p, x_i)] = \left[\max_{x_i \in X'_i} \min_{x_p \in X'_p} f(x_p, x_i), \min_{x_i \in X'_i} \max_{x_p \in X'_p} f(x_p, x_i) \right], \tag{15}$$

where $x = (x_p, x_i)$ is component splitting corresponding to the proper and improper components of $X = (X_p, X_i)$.

The following two key results, named *semantic theorems*, give logical interpretation to these semantic extensions [49].

Theorem 1 ($*$-semantic theorem) *Let be* $X \in I^*(\mathbb{R}^n)$ *and* $Z \in I^*(\mathbb{R})$. *Then,* $f^*(X) \subseteq Z \Leftrightarrow (\forall x_p \in X'_p)(Qz \in Z')(\exists x_i \in X'_i)\ z = f(x_p, x_i)$.

Theorem 2 ($**$-semantic theorem) *Let be* $X \in I^*(\mathbb{R}^n)$ *and* $Z \in I^*(\mathbb{R})$. *Then,* $f^{**}(X) \supseteq Z \Leftrightarrow (\forall x_i \in X'_i)(Qz \in Dual(Z))(\exists x_p \in X'_p)\ z = f(x_p, x_i)$.

Both semantic theorems equate a logical formula, with intervals and functional predicates for which the universal quantifiers precede the existing ones, to an interval inclusion.

Semantic extensions f^* and f^{**} can be equal or not, but they are out of reach for any direct computation except for simple real functions. When the continuous function f is a rational function, there exist *modal rational extensions* that are obtained using a computing program defined by the syntax tree of the expression of the function in which the real arguments are transformed into interval arguments and the real operators are transformed into their *-semantic extensions. If f is a $\mathbb{R}^n$ to $\mathbb{R}$ rational function, its rational extension to the modal intervals $X_1, \ldots, X_n$, represented by $fR(X_1, \ldots, X_n)$, is the function fR from $I^*(\mathbb{R}^n)$ to $I^*(\mathbb{R})$ defined by the computational program indicated by the syntax tree of f when the real operators are transformed to their semantic extensions. Modal rational interval functions are not generally interpretable.

Thus, the problem is that semantic extensions f^* and f^{**} are interpretable but not generally computable and rational extensions are computable but not generally interpretable. The interpretation problem for modal rational functions, which are the core of numerical computing, consists of relating them by means of *inclusion relations to the corresponding semantic * and **-semantic extensions*, which have a standard meaning (defined by the semantic theorems) referring to their original

real continuous functions. Computations with $fR(X)$ must be carried out with external truncation of each operator to obtain inclusions $f^*(X) \subseteq fR(X)$, and with inner truncation to obtain inclusions $fR(X) \subseteq f^{**}(X)$. In many cases the rational extension $fR(X)$ is *optimal*, i.e.,

$$f^*(X) = fR(X) = f^{**}(X),$$

and, except for rounding, both semantic theorems are applicable to the computed interval $fR(X)$, providing a logical meaning for it.

MIA provides a group of results about these inclusions or equalities, which solves part of the double problem of interpretability of modal rational extensions and computability of semantic extensions. The following coercion theorem provides the conditions and a method to obtain optimal extensions in the case of total monotonicity [49].

Definition 1 (*Total monotonicity*) A continuous real function f is x-totally monotonic for a multi-incident variable $x \in \mathbb{R}$ if it is uniformly monotonic for this variable and for each one of its incidences (considering each leaf of the syntactic tree as an independent variable). Any uni-incident uniformly monotonic variable is totally monotonic too.

Theorem 3 (Coercion to optimality) *Let X be an interval vector, $fR(X)$ defined and tree optimal on the domain X' and f totally monotonic for all its multi-incident components. Let XD be defined as an enlarged vector of X, such that each incidence of every multi-incident component is included in XD as an independent component, but transformed into its dual if the corresponding incidence point has a monotony-sense contrary to the global one of the corresponding X component. Then, $f^*(X) = fR(XD) = f^{**}(X)$.*

If the function is not totally monotonic for each multi-incident component, theorems can be partially applied to reduce the complexity of the problem. All these computations are carried out by controlling the rounding of the operations and taking into account the multi-incident variables in the functions [49].

Using modal intervals, each interval function to be evaluated is automatically analyzed and put, if possible, in its optimal form. Then the exact range is computed. The C++ modal interval library IvalDB can be used for this purpose. If optimality cannot be reached, the f^* algorithm is available [49]. This algorithm uses many optimality and coercion theorems from modal interval theory to compute tight approximations of the range by using branch-and-bound techniques. However, the f^* algorithm has exponential complexity.

3.2 Monotone Input–Output Systems

The computation of envelopes for continuous-time systems like (3) can also be addressed by building upper and lower bounding *deterministic* models whose solutions are the envelope bounds (exact or minimally overbounding). Then, these ones

can be simulated with interval numerical methods with infinitesimal interval width describing rounding errors due to the finite computer number system, if rigorous solutions are needed. In this case, phenomena like the wrapping effect will have a minimal impact.

Uncertain parameters will be treated as stationary extended states. Thus, the system (3) will be rewritten as

$$\frac{d\tilde{\mathbf{x}}(t)}{dt} = \tilde{\mathbf{f}}(\tilde{\mathbf{x}}(t), \mathbf{u}(t)), \quad \tilde{\mathbf{x}}(0) = \tilde{\mathbf{X}}_0 \tag{16}$$

$$\mathbf{y}(t) = \tilde{\mathbf{h}}(\tilde{\mathbf{x}}(t), \mathbf{u}(t)) \tag{17}$$

where $\tilde{\mathbf{x}}(t) := (x_1(t), \ldots, x_n(t), p_1(t), \ldots, p_{n_p}(t))^T$ is the extended state vector including the original state vector $\mathbf{x}$ and the uncertain interval parameters $p_i \in P_i$, $i = 1, \ldots, n_p$, so that $\frac{dp_i(t)}{dt} = 0$, $i = 1, \ldots, n_p$. The uncertain initial conditions are then given by the interval vector $\tilde{\mathbf{X}}_0 := (X_1, \ldots, X_n, P_1, \ldots, P_{n_p})$.

Exact envelopes can be obtained for the class of *monotone systems* [50]. Although autonomous systems have traditionally been studied in the field of monotone systems, the notion of monotonicity was extended to input–output systems like (16)–(17) in [1]. Denoting by $\tilde{\mathbf{x}}(t; \tilde{\mathbf{x}}_\mathbf{0}, \mathbf{u})$ and $\mathbf{y}(t; \tilde{\mathbf{x}}_0, \mathbf{u})$ the solutions of (16) and (17), respectively, for the real initial condition $\tilde{\mathbf{x}}_\mathbf{0} \in \mathbb{R}^{n+n_p}$ and input $\mathbf{u}(t)$, the input–output system (16)–(17) is monotone for all initial states $\tilde{\mathbf{x}}_\mathbf{0}^1, \tilde{\mathbf{x}}_\mathbf{0}^2$ and inputs $\mathbf{u}^1(t), \mathbf{u}^2(t)$ if and only if $\tilde{\mathbf{x}}_\mathbf{0}^1 \preceq \tilde{\mathbf{x}}_\mathbf{0}^2$, $\mathbf{u}^1(t) \preceq \mathbf{u}^2(t) \Rightarrow (\tilde{\mathbf{x}}(t; \tilde{\mathbf{x}}_\mathbf{0}^1, \mathbf{u}^1) \preceq \tilde{\mathbf{x}}(t; \tilde{\mathbf{x}}_\mathbf{0}^2, \mathbf{u}^2)) \wedge (\mathbf{y}(t; \tilde{\mathbf{x}}_\mathbf{0}^1, \mathbf{u}^1) \preceq \mathbf{y}(t; \tilde{\mathbf{x}}_\mathbf{0}^2, \mathbf{u}^2))$ for all $t \geq 0$, where $\preceq$ is the relation order induced by a cone K ($\mathbf{a} \preceq \mathbf{b}$ if $\mathbf{b} - \mathbf{a} \in K$), which may be different for the state, input and output (in this case they will be denoted by K_x, K_u and K_y respectively). A special class of monotone input–output systems are *cooperative input–output systems* in which for all $t \geq 0$:

$$\frac{\partial \tilde{f}_i}{\partial \tilde{x}_j}(\tilde{\mathbf{x}}, \mathbf{u}) \geq 0, \quad \forall\, i \neq j, \qquad \frac{\partial \tilde{h}_s}{\partial \tilde{x}_j}(\tilde{\mathbf{x}}, \mathbf{u}) \geq 0, \quad \forall\, s, j, \tag{18}$$

$$\frac{\partial \tilde{f}_i}{\partial u_r}(\tilde{\mathbf{x}}, \mathbf{u}) \geq 0, \quad \forall\, i, r, \qquad \frac{\partial \tilde{h}_s}{\partial u_r}(\tilde{\mathbf{x}}, \mathbf{u}) \geq 0, \quad \forall\, s, r, \tag{19}$$

i.e., the Jacobian $\partial \tilde{\mathbf{f}}/\partial \tilde{\mathbf{x}}$ is a Metzler matrix (nonnegative off-diagonal elements) and $\partial \tilde{\mathbf{f}}/\partial \mathbf{u}$, $\partial \tilde{\mathbf{h}}/\partial \tilde{\mathbf{x}}$ and $\partial \tilde{\mathbf{h}}/\partial \mathbf{u}$ are nonnegative. For cooperative input–output systems the cones K_x, K_u and K_y are the positive orthants $\mathbb{R}^{n+n_p}_{\geq 0}$, $\mathbb{R}^{n_u}_{\geq 0}$ and $\mathbb{R}^{n_y}_{\geq 0}$, respectively, and the relation order $\mathbf{a} \preceq \mathbf{b}$ means that each coordinate of $\mathbf{a}$ is less than or equal than the corresponding coordinate in $\mathbf{b}$. In this case the exact envelope can be computed

as the solutions $\tilde{\mathbf{x}}_{\mathbf{L}}(t), \tilde{\mathbf{x}}_{\mathbf{U}}(t) : \mathbb{R} \to \mathbb{R}^n$ and $\mathbf{y}_{\mathbf{L}}(t), \mathbf{y}_{\mathbf{U}}(t) : \mathbb{R} \to \mathbb{R}^{n_y}$ of the following lower and upper bounding systems, respectively:

$$\text{(lower)} \qquad \begin{aligned} \frac{d\tilde{\mathbf{x}}_{\mathbf{L}}(t)}{dt} &= \tilde{\mathbf{f}}(\tilde{\mathbf{x}}_{\mathbf{L}}(t), \underline{\mathbf{u}(t)}), \quad \tilde{\mathbf{x}}_{\mathbf{L}}(0) = \underline{\tilde{\mathbf{X}}_{\mathbf{0}}} \\ \mathbf{y}_{\mathbf{L}}(t) &= \tilde{\mathbf{h}}(\tilde{\mathbf{x}}_{\mathbf{L}}(t), \underline{\mathbf{u}(t)}) \end{aligned} \tag{20}$$

$$\text{(upper)} \qquad \begin{aligned} \frac{d\tilde{\mathbf{x}}_{\mathbf{U}}(t)}{dt} &= \tilde{\mathbf{f}}(\tilde{\mathbf{x}}_{\mathbf{U}}(t), \overline{\mathbf{u}(t)}), \quad \tilde{\mathbf{x}}_{\mathbf{U}}(0) = \overline{\tilde{\mathbf{X}}_{\mathbf{0}}} \\ \mathbf{y}_{\mathbf{U}}(t) &= \tilde{\mathbf{h}}(\tilde{\mathbf{x}}_{\mathbf{U}}(t), \overline{\mathbf{u}(t)}) \end{aligned} \tag{21}$$

where $\underline{\tilde{\mathbf{X}}_{\mathbf{0}}}$ and $\overline{\tilde{\mathbf{X}}_{\mathbf{0}}}$ denote the vector of left and right endpoints, respectively, of the interval initial extended state and equivalently $\underline{\mathbf{u}(t)}$ and $\overline{\mathbf{u}(t)}$ for the input vector at time t.

Exact envelopes can also be obtained for cooperative systems with respect to general orthants K_x^{ε}, K_u^{δ}, K_y^{ν}, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{n+n_p})$, $\delta = (\delta_1, \ldots, \delta_{n_u})$, $\nu = (\nu_1, \ldots, \nu_{n_y})$, $\varepsilon_i, \delta_r, \nu_s \in \{0, 1\}$ so that $(-1)^{\varepsilon_i}\tilde{x}_i \geq 0$, $i = 1, \ldots, n + n_p$, $(-1)^{\delta_r}u_r \geq 0$, $r = 1, \ldots, n_u$ and $(-1)^{\nu_s}y_s \geq 0$, $s = 1, \ldots, n_y$. In this case, for all $t \geq 0$

$$(-1)^{\varepsilon_i+\varepsilon_j}\frac{\partial \tilde{f}_i}{\partial \tilde{x}_j}(\tilde{\mathbf{x}}, \mathbf{u}) \geq 0, \ \forall\, i \neq j, \qquad (-1)^{\nu_s+\varepsilon_j}\frac{\partial \tilde{h}_s}{\partial \tilde{x}_j}(\tilde{\mathbf{x}}, \mathbf{u}) \geq 0, \ \forall\, s, j, \tag{22}$$

$$(-1)^{\varepsilon_i+\delta_r}\frac{\partial \tilde{f}_i}{\partial u_r}(\tilde{\mathbf{x}}, \mathbf{u}) \geq 0, \ \forall\, i, r, \qquad (-1)^{\nu_s+\delta_r}\frac{\partial \tilde{h}_s}{\partial u_r}(\tilde{\mathbf{x}}, \mathbf{u}) \geq 0, \ \forall\, s, r \tag{23}$$

holds [1]. The state, after a proper permutation, can be expressed as $\tilde{\mathbf{x}}(t) = (\tilde{\mathbf{x}}^{+^T}\ \tilde{\mathbf{x}}^{-^T})^T$ so that $\tilde{\mathbf{x}}^+$ contains the "positive states" ($\varepsilon_i = 0$) and $\tilde{\mathbf{x}}^-$ contains the "negative states" ($\varepsilon_i = 1$) according to the orthant K_x^{ε}. This leads to a Morishima-type Jacobian block matrix $\partial\tilde{\mathbf{f}}/\partial\tilde{\mathbf{x}}$ with Metzler diagonal submatrices and non-positive off-diagonal submatrices. The same applies for the input and output vectors, which can be decomposed as $\mathbf{u}(t) = (\mathbf{u}^{+^T}\ \mathbf{u}^{-^T})^T$ and $\mathbf{y}(t) = (\mathbf{y}^{+^T}\ \mathbf{y}^{-^T})^T$ according to the orthants K_u^{δ} and K_y^{ν}, respectively. This yields block matrices with nonnegative diagonal submatrices and non-positive off-diagonal submatrices for $\partial\tilde{\mathbf{f}}/\partial\mathbf{u}$, $\partial\tilde{\mathbf{h}}/\partial\tilde{\mathbf{x}}$ and $\partial\tilde{\mathbf{h}}/\partial\mathbf{u}$. In this case the exact envelope can be computed as the solutions of the following lower and upper bounding systems, respectively:

$$\begin{aligned} &\text{(lower/upper)} \\ &\frac{d}{dt}\begin{pmatrix} \tilde{\mathbf{x}}_{\mathbf{L}}^+(t) \\ \tilde{\mathbf{x}}_{\mathbf{U}}^-(t) \end{pmatrix} = \tilde{\mathbf{f}}\left(\begin{pmatrix} \tilde{\mathbf{x}}_{\mathbf{L}}^+(t) \\ \tilde{\mathbf{x}}_{\mathbf{U}}^-(t) \end{pmatrix}, \begin{pmatrix} \mathbf{u}^+(t) \\ \overline{\mathbf{u}^-(t)} \end{pmatrix}\right), \quad \begin{pmatrix} \tilde{\mathbf{x}}_{\mathbf{L}}^+(t) \\ \tilde{\mathbf{x}}_{\mathbf{U}}^-(t) \end{pmatrix}(0) = \begin{pmatrix} \underline{\tilde{\mathbf{X}}_{\mathbf{0}}}^+ \\ \overline{\tilde{\mathbf{X}}_{\mathbf{0}}}^- \end{pmatrix} \\ &\begin{pmatrix} \tilde{\mathbf{y}}_L^+(t) \\ \tilde{\mathbf{y}}_{\mathbf{U}}^-(t) \end{pmatrix} = \tilde{\mathbf{h}}\left(\begin{pmatrix} \tilde{\mathbf{x}}_{\mathbf{L}}^+(t) \\ \tilde{\mathbf{x}}_{\mathbf{U}}^-(t) \end{pmatrix}, \begin{pmatrix} \mathbf{u}^+(t) \\ \overline{\mathbf{u}^-(t)} \end{pmatrix}\right), \end{aligned} \tag{24}$$

(upper/lower)

$$\frac{d}{dt}\begin{pmatrix}\tilde{\mathbf{x}}_{\mathbf{U}}^{+}(t)\\ \tilde{\mathbf{x}}_{\mathbf{L}}^{-}(t)\end{pmatrix}=\tilde{\mathbf{f}}\left(\begin{pmatrix}\tilde{\mathbf{x}}_{\mathbf{U}}^{+}(t)\\ \tilde{\mathbf{x}}_{\mathbf{L}}^{-}(t)\end{pmatrix},\begin{pmatrix}\overline{\mathbf{u}^{+}(t)}\\ \underline{\mathbf{u}^{-}(t)}\end{pmatrix}\right),\quad \begin{pmatrix}\tilde{\mathbf{x}}_{\mathbf{U}}^{+}(t)\\ \tilde{\mathbf{x}}_{\mathbf{L}}^{-}(t)\end{pmatrix}(0)=\begin{pmatrix}\overline{\tilde{\mathbf{X}}_{\mathbf{0}}}^{+}\\ \underline{\tilde{\mathbf{X}}_{\mathbf{0}}}^{-}\end{pmatrix}$$

$$\begin{pmatrix}\tilde{\mathbf{y}}_{\mathbf{U}}^{+}(t)\\ \tilde{\mathbf{y}}_{\mathbf{L}}^{-}(t)\end{pmatrix}=\tilde{\mathbf{h}}\left(\begin{pmatrix}\tilde{\mathbf{x}}_{\mathbf{U}}^{+}(t)\\ \tilde{\mathbf{x}}_{\mathbf{L}}^{-}(t)\end{pmatrix},\begin{pmatrix}\overline{\mathbf{u}^{+}(t)}\\ \underline{\mathbf{u}^{-}(t)}\end{pmatrix}\right), \tag{25}$$

Besides by inspection of the Jacobian matrix structure, cooperativeness with respect to orthants K_x^{ε}, K_u^{δ} and K_y^{ν} can be easily characterized by using a *species graph* [51], in which a node is assigned to each state, input and output. An activation ($\rightarrow$) or inhibition ($\dashv$) arrow is drawn from node $\tilde{x}_i$ to node $\tilde{x}_j$ if $\partial\tilde{f}_j/\partial\tilde{x}_i > 0$ or $\partial\tilde{f}_j/\partial\tilde{x}_i < 0$, respectively. In case of no definite sign both arrows are drawn. The same rule applies for connecting input and state nodes, state and output nodes and input and output nodes (according to the sign definiteness of the corresponding partial derivative). A *spin assignment* is then carried out in which each node is assigned a sign or spin ($\uparrow$, $\downarrow$), such that nodes connected by an activation arrow $\rightarrow$ have the same spin, while nodes connected by an inhibition arrow $\dashv$ have opposed spins. If at least one consistent assignment exists, the dynamical input–output system is monotone and the type of spin will classify the state, input and output into the groups $\tilde{\mathbf{x}}^{+}$, $\tilde{\mathbf{x}}^{-}$, $\mathbf{u}^{+}$, $\mathbf{u}^{-}$, $\mathbf{y}^{+}$ and $\mathbf{y}^{-}$ respectively. Figure 3 shows an example of a monotone and a non-monotone input–output system.

In case of non-monotonicity overbounding will appear. The application of change of variables to overcome or reduce the non-monotonicity can help in some cases, for instance by reducing multi-incidences. Also useful is the bounding of $\tilde{\mathbf{f}}(\tilde{\mathbf{x}}(t),\mathbf{u}(t))$ by two new vector fields $\tilde{\mathbf{f}}_{\mathbf{L}}(\tilde{\mathbf{x}}(t),\mathbf{u}(t))$ and $\tilde{\mathbf{f}}_{\mathbf{U}}(\tilde{\mathbf{x}}(t),\mathbf{u}(t))$ leading to cooperative

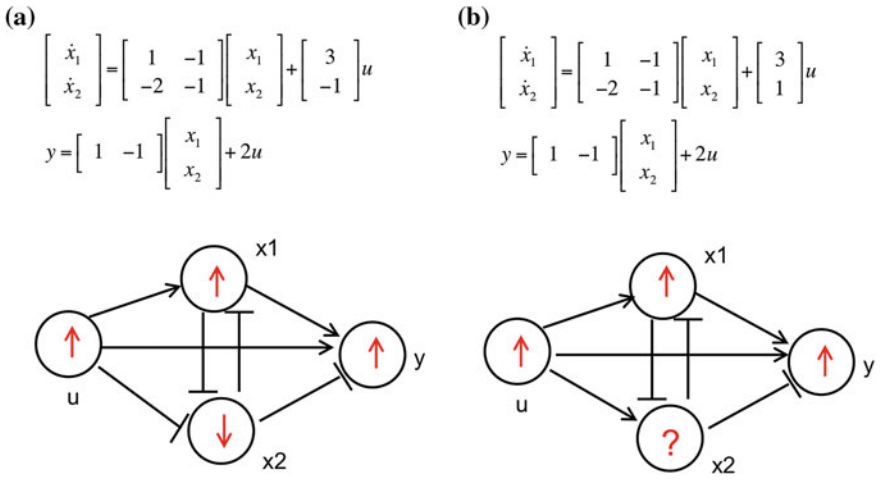

Fig. 3 Characterization of monotone and non-monotone input–output systems using species graphs: **a** monotone system; **b** non-monotone system

lower and upper bounding systems [50]. This is applied in [45] to build a minimally overbounding interval predictor for the Hovorka model.

4 Interval Glucose Predictors

Modal interval analysis [7, 8, 18] and monotone systems [45] have been successfully applied to the most common physiological glucose–insulin models (Bergman [4], Hovorka [25], Dalla Man [10]) yielding to computationally efficient and exact or minimally overbounding interval glucose predictors. These ones have been used in a variety of applications such as model-based open-loop insulin therapy [6, 17], hypoglycemia risk prediction [16] and fault detection [22].

In this section, the application of the techniques developed in Sect. 3 to build such predictors is presented. Bergman model will be used as illustration for the sake of simplicity:

$$\dot{x}(t) = -p_2 x(t) + p_3(i(t) - i_b), \qquad x(0) = x_0 \tag{26}$$

$$\dot{g}(t) = -(p_1 + x(t))g(t) + p_1 g_b + \frac{r_a(t)}{v_g}, \quad g(0) = g_0 \tag{27}$$

where $x(t)$ (1/min) represents remote insulin, $i(t)$ (mU/mL) is the plasma insulin concentration, with basal value i_b, p_2 (1/min) is the rate of appearance of insulin in the remote insulin compartment, p_3 (mL/mU min^2) is the disappearance of insulin from this compartment, $g(t)$ (mg/dL) is the plasma glucose concentration, g_b (mg/dL) is the basal plasma glucose concentration, p_1 (1/min) is the rate at which glucose is removed from the plasma space independent of the influence of insulin, $r_a(t)$ is the glucose rate of appearance from the gut from a preceding meal ingestion model (or intravenous glucose), and v_g (dL) is the glucose distribution volume.

The reader is referred to [15, 45] for a complete description of different interval glucose predictors.

4.1 Bergman Model Predictor Based on Modal Interval Analysis

The Euler discrete form of Bergman model is given by:

$$X(t+1) = X(t)\ (1 - \Delta t\ p_2) + \Delta t p_3(I(t) - I_b) \tag{28}$$

$$G(t+1) = \Delta t p_1(G_b - G(t)) + G(t)(1 - \Delta t X(t)) + \Delta t \frac{R_a(t)}{V_g} \tag{29}$$

where X (1/min) represents remote insulin, I (mU/mL) is the plasma insulin concentration, with basal value I_b, p_2 (1/min) is the rate of appearance of insulin in the

remote insulin compartment, p_3 (mL/mU min^2) is the disappearance of insulin from this compartment, G (mg/dL) is the plasma glucose concentration, G_b (mg/dL) is the basal plasma glucose concentration, p_1 (1/min) is the rate at which glucose is removed from the plasma space independent of the influence of insulin and V_g (dL) is the glucose distribution space.

Equation (28) does not have multi-incident variables, i.e., multiple occurrences of a variable, and therefore the rational computation can directly be applied to get $X(t+1)$. In Eq. (29), $G(t)$ is multi-incident with $\partial G(t+1)/\partial G(t) = 1 - \Delta t(p_1 + X(t))$. Global positive monotonicity with respect to $G(t)$ holds for a time step $\Delta t < 1/(p_1 + X(t))$. Denoting as $G^1(t)$ and $G^2(t)$ the first and second instances of $G(t)$ in (29), then $\partial G(t+1)/\partial G^1(t) = -\Delta t p_1 < 0$ and $\partial G(t+1)/\partial G^2(t) = (1 - \Delta t X(t)) > 0$ for the chosen time step. Thus, applying the coercion to optimality theorem (see Theorem 3) to plasma glucose concentration the rational computation is:

$$G(t+1) = \Delta t p_1 (G_b - Dual(G(t))) + G(t)(1 - \Delta t X(t)) + \Delta t \frac{R_a(t)}{V_g} \tag{30}$$

Equations (28) and (30) can now be computed using the modal interval library IvalDb [49] or the MATLAB class for modal interval arithmetic (without rounding) incorporated in VSIVIA software [23] (downloadable from http://www3.imperial.ac.uk/people/p.herrero-vinias/research).

4.2 *Bergman Model Predictor Based on Monotone Systems Theory*

Figure 4 shows the species graph associated to the model (26)–(27). A consistent spin assignment can be done, although the resulting monotonicity cone will depend

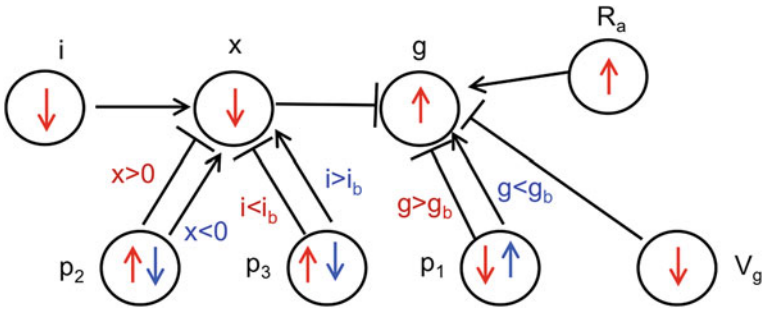

Fig. 4 Species graph for the Bergman model

on the sign of $x(t)$, $i(t) - i_b$ and $g(t) - g_b$. In particular:

$$\begin{aligned}
\tilde{\mathbf{x}}^+ &:= \{g\} \\
&\quad \cup \{p_1\} \quad (\text{if} \quad g(t) \leq g_b) \\
&\quad \cup \{p_2\} \quad (\text{if} \quad x(t) \leq 0) \\
&\quad \cup \{p_3\} \quad (\text{if} \quad i(t) \leq i_b) \\
\tilde{\mathbf{x}}^- &:= \{x, v_g\} \\
&\quad \cup \{p_1\} \quad (\text{if} \quad g(t) > g_b) \\
&\quad \cup \{p_2\} \quad (\text{if} \quad x(t) > 0) \\
&\quad \cup \{p_3\} \quad (\text{if} \quad i(t) > i_b) \\
\mathbf{u}^+ &:= \{r_a\} \\
\mathbf{u}^- &:= \{i\}
\end{aligned} \tag{31}$$

In this case the output corresponds to the state $g(t)$. The following time-varying bounding systems are obtained:

(lower/upper)

$$\dot{x_U}(t) = -p_2^* x_U(t) + p_3^*(\overline{i(t)} - i_b), \qquad x_U(0) = \overline{x_0} \tag{32}$$

$$\dot{g_L}(t) = -(p_1^* + x_U(t))g_L(t) + p_1^* g_b + \frac{r_a(t)}{\overline{v_g}}, \quad g_L(0) = \underline{g_0} \tag{33}$$

$$p_1^* = \begin{cases} \underline{p_1} & \text{if } g_L(t) \leq g_b \\ \overline{p_1} & \text{if } g_L(t) > g_b \end{cases}$$

$$p_2^* = \begin{cases} \underline{p_2} & \text{if } x_U(t) \leq 0 \\ \overline{p_2} & \text{if } x_U(t) > 0 \end{cases}$$

$$p_3^* = \begin{cases} \underline{p_3} & \text{if } \overline{i(t)} \leq i_b \\ \overline{p_3} & \text{if } \overline{i(t)} > i_b \end{cases}$$

(upper/lower)

$$\dot{x_L}(t) = -p_2^* x_L(t) + p_3^*(\underline{i(t)} - i_b), \qquad x_L(0) = \underline{x_0} \tag{34}$$

$$\dot{g_U}(t) = -(p_1^* + x_L(t))g_U(t) + p_1^* g_b + \frac{\overline{r_a(t)}}{\underline{v_g}}, \quad g_U(0) = \overline{g_0} \tag{35}$$

$$p_1^* = \begin{cases} \overline{p_1} & \text{if } g_U(t) \leq g_b \\ \underline{p_1} & \text{if } g_U(t) > g_b \end{cases}$$

$$p_2^* = \begin{cases} \overline{p_2} & \text{if } x_L(t) \leq 0 \\ \underline{p_2} & \text{if } x_L(t) > 0 \end{cases}$$

$$p_3^* = \begin{cases} \overline{p_3} & \text{if } \underline{i(t)} \leq i_b \\ \underline{p_3} & \text{if } \underline{i(t)} > i_b \end{cases}$$

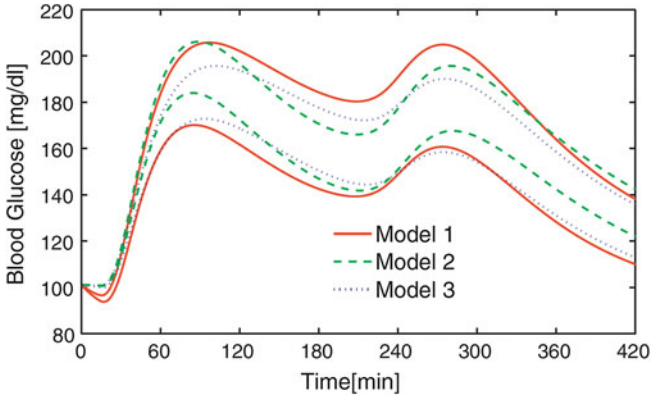

Fig. 5 Example of postprandial response for three interval models. *Model 1* Bergman; *Model 2* Hovorka; *Model 3* Dalla Man

4.3 Postprandial Glucose Prediction Using Interval Models

In the case of the postprandial response, in addition to intra-patient variability, there are other sources of uncertainty such as the estimation of ingested carbohydrates or preprandial glycemia. In [18] the postprandial response of three different models (Bergman [4], Hovorka [25], Dalla Man [10]) is compared for a given uncertainty variations in the estimation of carbohydrates, preprandial glycemia and insulin sensitivity. One of the main conclusions of this work is that, under variability, structurally simple glucose–insulin models (models with few interval parameters) may be sufficient to describe patient dynamics in most situations. These simple interval models are much more easily identifiable. Figure 5 shows the postprandial response of the three models for a representative virtual patient.

5 Interval Model Identification

Once the mathematical formulation of an interval glucose predictor is available, the problem of model identification can be posed. Robust predictions for therapeutic 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 available time instant t_i, $i \in I := \{0, \ldots, N\}$ (see Fig. 6). 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 (especially in a domiciliary context with the use of continuous glucose monitors) and compensation for non-modeled dynamics.

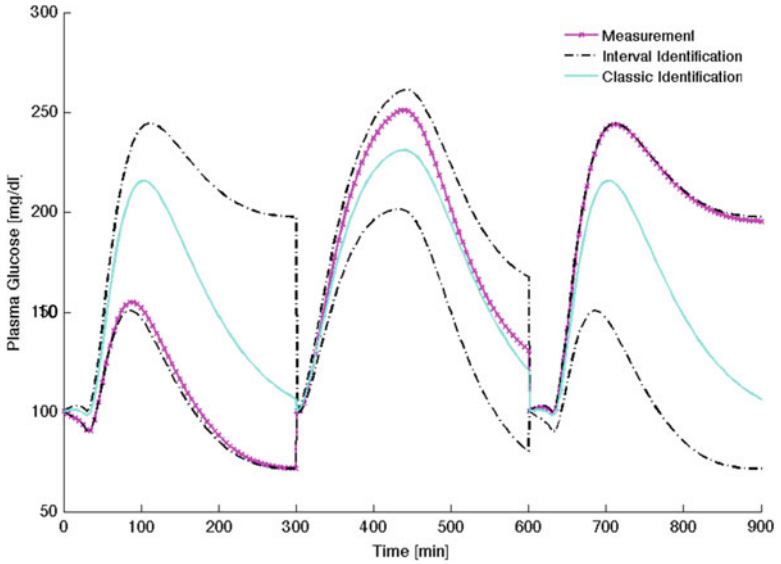

Fig. 6 Concept of interval identification versus classic identification for three-day 5 h postprandial periods in an in silico patient with time-varying dynamics (day 1: 0–300 min; day 2: 300–600 min; day 3: 600–900 min). Adapted from [35]

In [34] this is carried out as a standard global optimization problem with the interval model parameters/inputs/initial conditions endpoints $\mathbb{P} := \{\underline{p_1}, \overline{p_1}, \ldots, \underline{p_{n_p}}, \overline{p_{n_p}}\}$ as decision variables. Denoting as $G(t;\mathbb{P}) = [\underline{g(t;\mathbb{P})}, \overline{g(t;\mathbb{P})}]$ the interval model glucose prediction for the parameters $\mathbb{P}$ and $g^*(t_i)$ the measurement at time t_i, the following composite cost index is defined:

$$J_{we}(\mathbb{P}) := J_w(\mathbb{P}) + \gamma \cdot J_e(\mathbb{P}) \tag{36}$$

$$J_w(\mathbb{P}) := \max_{i \in I} \left(\overline{g(t_i;\mathbb{P})} - \underline{g(t_i;\mathbb{P})} \right) \tag{37}$$

$$J_e(\mathbb{P}) := \sum_{i=1}^{N} d_H^2 \left(g^*(t_i), G(t_i;\mathbb{P}) \right) \tag{38}$$

where d_H is the Hausdorff distance from the measurement to the glucose envelope (distance from a point to an interval)

$$d_H \left(g^*(t_i), G(t_i;\mathbb{P}) \right) := \begin{cases} 0 & \text{if } g^*(t_i) \in G(t_i;\mathbb{P}) \\ \min(|g^*(t_i) - \overline{G(t_i;\mathbb{P})}|, & \\ \quad |g^*(t_i) - \underline{G(t_i;\mathbb{P})}|) & \text{if } g^*(t_i) \notin G(t_i;\mathbb{P}) \end{cases} \tag{39}$$

J_w represents the envelope width computed as the maximum width along time. J_e represents the fitting error extended to set computations. Parameter γ is a weighting

factor among them which has to be tuned a priori, depending on the available data characteristics. It defines the degree of relaxation given to the optimization problem.

The above method was successfully applied to the Hovorka model identification in [32, 34] from 5 h postprandial clinical data for a cohort of 12 patients. Patients were monitored in their postprandial state on four occasions. 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 100 g of CHO. Preprandial plasma glucose was set around 100 mg/dL by means of a manual feedback intravenous insulin infusion.

In [34], plasma insulin and glucose reference measurements were considered (in-clinic setting). Interval uncertainty was included in insulin sensitivity, plasma-to-tissues glucose transport rate and meal rate of appearance. A value of $\gamma = 100$ was considered for the error term weight in the cost index. The first conclusion drawn from this study was a trivial one: identification data must be representative of patient's variability when aiming at its characterization. Four postprandial periods may fall short for this purpose since a highly variable patient behavior was observed among the different days. Following a leave-one-day-out validation strategy, a specific combination of identification–validation days was always found out to achieve good prediction ability, corresponding to the one with more representative identification days. Despite the limitation that this implied for this study, it is expected that data representativeness will hold in a real-time setting with long enough time series. A median glucose envelope width of 94 mg/dL enclosing 97.3 % of glucose reference data in the validation day (YSI 2300 STAT Plus Glucose analyzer, Yellow Springs Instruments, Ohio, USA) was achieved for the best identification–validation days combination. Median MARD (Mean Absolute Relative Deviation) of data outside the predicted envelope was 1.5 % and overall a MARD of 0.03 % was achieved considering zero error for data inside the envelope. Despite the large width identified, median envelope overbounding was estimated in 15.55 mg/dL, i.e., at each time instant at least a measurement as close as 15.55 mg/dL to the upper or lower envelope bound for some of the identification days existed (in median). This value is comparable to current continuous glucose monitoring accuracy and reveals that the large envelope width was due to actual patient variability.

Results under conditions closer to domiciliary data are presented in [32], where insulin pump and continuous glucose monitoring data were considered. Interval uncertainty in subcutaneous insulin absorption was additionally considered in this case. For the same tuning of $\gamma = 100$, continuous glucose monitoring induced larger envelope widths, as expected, with a median value of 115.4 mg/dL and an envelope overbounding estimation of 22.1 mg/dL due to the noisy measurements. An overall median MARD of 0.14 % was achieved with a median of 97.8 % of data enclosed by the predicted glucose envelope. Noise compensation was addressed by decreasing the weight γ, giving more importance to envelope width minimization. For a value of $\gamma = 25$, a median envelope width of 67.6 mg/dL was obtained, with an envelope overbounding estimation of 16 mg/dL, comparable to the in-clinic setting. In this case, a median value of 96.2 % of data were enclosed yielding an overall median MARD

of 0.27 %. These studies, although with limitations, demonstrate the feasibility of interval model identification in a practical setting.

6 Conclusions

Prediction of blood glucose in patients with diabetes is a very difficult challenge, mostly due to the intra-patient variability. Interval models have been proven to be a very useful tool to represent uncertainty and variability. In this paper we have presented the basics of the interval models as well as methods to simulate them.

Modal interval analysis and monotone systems principles have been introduced for the prediction of glucose excursions in type 1 diabetes under uncertainty and intra-patient variability. These methods allow to efficiently tackle important problems such as the study of the effect of a given insulin therapy on glycemia and its optimization given uncertain parameters such as hepatic and peripheral insulin sensitivities and food intake. As result, an envelope containing all possible glucose excursions suffered by the patient for the given uncertainty is obtained. In contrast to probabilistic methods such as Monte Carlo simulation, sharp envelopes containing all possible patient responses can be computed using the methods introduced in Sect. 3 (guaranteed simulation). Thus, worst-case analysis can be performed efficiently, which is extremely important in the context of diabetes. Prediction of possible hyper- and hypoglycemic episodes considering patient's variability may yield to safer and more robust insulin infusion algorithms.

Glucose prediction is a complex process owing to that simple models cannot represent all glucose dynamics and complex models are difficult or impossible to identify with simple experiments. Additionally, the presence of intra-patient variability (insulin sensitivity, s.c. insulin absorption, etc.) should be taken into account. However, as illustrated in Fig. 5, in the presence of intra-patient variability and different sources of uncertainty a reduction of model complexity may be feasible: a simple model can be easily identified for a single patient and represent all possible behaviors included in the simulation.

To this end, classic identification procedures may be jeopardized by high intra-patient variability, resulting in poor predictive capabilities of average models. The use interval models for identification, as introduced in Sect. 5, allow for the characterization of individual variability from experimental data. Interval models offer thus a powerful framework for the design of more robust therapies, control algorithms and fault detection, which are necessary components of an artificial pancreas for home use.

Acknowledgments This work was funded by the Spanish Ministry of Science and Innovation under grants DPI2010-20764-C02-01, DPI2010-20764-C02-02, DPI2013-46982-C2-1-R and DPI2013-46982-C2-2-R. The authors acknowledge Dr. Alejandro J. Laguna, Dr. Mei Calm, and Dr. Maira A. García-Jaramillo for providing some of the figures presented in this paper.

References

1. Angeli, D.E., Sontag, E.D.: Monotone control systems. IEEE Trans. Autom. Control **48**(10), 1684–1698 (2003)
2. Armengol, J., Travé-Massuyès, L., Vehí, J., de la Rosa, J.L.: A survey on interval model simulators and their properties related to fault detection. Annu. Rev. Control **24**, 31–39 (2000)
3. Armengol, J., Vehí, J., Travé-Massuyès, L., Sainz, M.A.: Application of modal intervals to the generation of error-bounded envelopes. Reliab. Comput. **7**(2), 171–185 (2001)
4. Bergman, R.N., Ider, Y.Z., Bowden, C.R., Cobelli, C.: Quantitative estimation of insulin sensitivity. Am. J. Physiol.—Endocrinol. Metab. **236**(6), E667–E677 (1979)
5. Bhattacharyya, S.P., Chapellat, H., Keel, L.: Robust Control: The Parametric Approach, 1st edn. Prentice Hall PTR, Upper Saddle River (1995)
6. Bondia, J., Dassau, E., Zisser, H., Calm, R., Vehí, J., Jovanovič, L., Doyle, F.J.: Coordinated basal-bolus infusion for tighter postprandial glucose control in insulin pump therapy. J. Diabetes Sci. Technol. **3**(1), 89–97 (2009)
7. Calm, R., García-Jaramillo, M., Bondia, J., Sainz, M.A., Vehí, J.: Comparison of interval and Monte Carlo simulation for the prediction of postprandial glucose under uncertainty in type 1 diabetes mellitus. Comput. Methods Programs Biomed. **104**(3), 325–332 (2011)
8. Calm, R., García-Jaramillo, M., Vehí, J., Bondia, J., Tarín, C., García-Gabín, W.: Prediction of glucose excursions under uncertain parameters and food intake in intensive insulin therapy for type 1 diabetes mellitus. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1770–3 (2007)
9. Christiansen, M., Bailey, T., Watkins, E., Liljenquist, D., Price, D., Nakamura, K., Boock, R., Peyser, T.: A new-generation continuous glucose monitoring system: improved accuracy and reliability compared with a previous-generation system. Diabetes Technol. Ther. **15**(10), 881–888 (2013)
10. Dalla Man, C., Rizza, R., Cobelli, C.: Meal simulation model of the glucose-insulin system. IEEE Trans. Biomed. Eng. **54**(10), 1740–1749 (2007)
11. Doyle, F.J., Huyett, L.M., Lee, J.B., Zisser, H.C., Dassau, E.: Closed-loop artificial pancreas systems: engineering the algorithms. Diabetes Care **37**(5), 1191–1197 (2014)
12. Finan, D.A., Doyle, F.J., Palerm, C.C., Bevier, W.C., Zisser, H.C., Jovanovic, L., Seborg, D.E.: Experimental evaluation of a recursive model identification technique for type 1 diabetes. J. Diabetes Sci. Technol. **3**(5), 1192–1202 (2009)
13. Galvanin, F., Barolo, M., Macchietto, S., Bezzo, F.: Optimal design of clinical tests for the identification of physiological models of type 1 diabetes mellitus. Ind. Eng. Chem. Res. **48**, 1989–2002 (2009)
14. Garcia, A., Rack-Gomer, A.L., Bhavaraju, N.C., Hampapuram, H., Kamath, A., Peyser, T., Facchinetti, A., Zecchin, C., Sparacino, G., Cobelli, C.: Dexcom G4AP: an advanced continuous glucose monitor for the artificial pancreas. J. Diabetes Sci. Technol. (Online) **7**(6), 1436–1445 (2013)
15. Garcia-Jaramillo, M.: 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. Ph.D. thesis, University of Girona, Girona, Spain (2011)
16. García-Jaramillo, M., Calm, R., Bondia, J., Tarín, C., Vehí, J.: Computing the risk of postprandial hypo- and hyperglycemia in type 1 diabetes mellitus considering intrapatient variability and other sources of uncertainty. J. Diabetes Sci. Technol. **3**(4), 895–902 (2009)
17. García-Jaramillo, M., Calm, R., Bondia, J., Tarín, C., Vehí, J.: Insulin dosage optimization based on prediction of postprandial glucose excursions under uncertain parameters and food intake. Comput. Methods Programs Biomed. **105**(1), 61–69 (2012)
18. García-Jaramillo, M., Calm, R., Bondia, J., Vehí, J.: Prediction of postprandial blood glucose under uncertainty and intra-patient variability in type 1 diabetes: a comparative study of three interval models. Comput. Methods Programs Biomed. **108**(1), 224–233 (2012)

19. Gardenyes, E., Mielgo, H., Trepat, A.: Modal intervals: Reasons and ground semantics. In: Nickel, K. (ed.) Interval Mathematics. Lectures Notes in Computer Science, vol. 212, pp. 27–35. Springer, Berlin (1985)
20. Haidar, A., Elleri, D., Kumareswaran, K., Leelarathna, L., Allen, J.M., Caldwell, K., Murphy, H.R., Wilinska, M.E., Acerini, C.L., Evans, M.L., Dunger, D.B., Nodale, M., Hovorka, R.: Pharmacokinetics of insulin aspart in pump-treated subjects with type 1 diabetes: reproducibility and effect of age, weight, and duration of diabetes. Diabetes Care **36**(10), e173–e174 (2013)
21. Hansen, E.: Global Optimization using Interval Analysis. Marcel Dekker, New York (1992)
22. Herrero, P., Calm, R., Vehí, J., Armengol, J., Georgiou, P., Oliver, N., Tomazou, C.: Robust fault detection system for insulin pump therapy using continuous glucose monitoring. J. Diabetes Sci. Technol. **6**(5), 1131–1141 (2012)
23. Herrero, P., Georgiou, P., Toumazou, C., Delaunay, B., Jaulin, L.: An efficient implementation of the sivia algorithm in a high-level numerical programming language. Reliab. Comput. **16**, 239–251 (2012)
24. Hirsch, I.B., Armstrong, D., Bergenstal, R.M., Buckingham, B., Childs, B.P., Clarke, W.L., Peters, A., Wolpert, H.: Clinical application of emerging sensor technologies in diabetes management: consensus guidelines for continuous glucose monitoring (CGM). Diabetes Technol. Ther. **10**(4), 232–246 (2008)
25. Hovorka, R., Canonico, V., Chassin, L.J., Haueter, U., et al.: Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. Physiological Measurement (2004)
26. Jaulin, L., Kieffer, M., Didrit, O., Walter, E.: Applied Interval Analysis: with Examples in Parameter and State Estimation, Robust Control and Robotics. Springer, Berlin (2012)
27. Kay, P.: Refining imprecise models and their behaviors. Ph.D. thesis, University of Texas at Austin, Austin, TX, USA (1996)
28. Kirchsteiger, H., Johansson, R., Renard, E., Del Re, L.: Continuous-time interval model identification of blood glucose dynamics for type 1 diabetes. Int. J. Control **87**(7), 1454–1466 (2014)
29. Kowalski, A.J.: Can we really close the loop and how soon? Accelerating the availability of an artificial pancreas: a roadmap to better diabetes outcomes. Diabetes Technol. Ther. **11**(Suppl 1), S113–S119 (2009)
30. Kühn, W.: Rigorously computed orbits of dynamical systems without the wrapping effect. Computing. Arch. Sci. Comput. **61**(1), 47–67 (1998)
31. Kuipers, B.: Qualitative simulation. Artif. Intell. **29**(3), 289–338 (1986)
32. Laguna, A.J.: Uncertainty in postprandial model identification in type 1 diabetes. Ph.D. thesis, Universitat Politècnica de València, Valencia, Spain (2014)
33. Laguna, A.J., Rossetti, P., Ampudia-Blasco, F.J., Vehí, J., Bondia, J.: Optimal design for individual model identification based on ambulatory continuous glucose monitoring in patients with type 1 diabetes. In: Conference proceedings: UKACC International Conference on Control, pp. 1–6 (2010)
34. Laguna, A.J., Rossetti, P., Ampudia-Blasco, F.J., Vehí, J., Bondia, J.: Experimental blood glucose interval identification of patients with type 1 diabetes. J. Process Control **24**(1), 171–181 (2014)
35. Laguna, A.J., Rossetti, P., Ampudia-Blasco, F.J., Vehí, J., Bondia, J.: Identification of intrapatient variability in the postprandial response of patients with type 1 diabetes. Biomed. Signal Process. Control **12**, 39–46 (2014)
36. Lohner, R.J.: Enclosing the solutions of ordinary initial and boundary value problems. Computer Arithmetic, Scientific Computation and Programming Languages, pp. 255–286 (1987)
37. Makino, K., Berz, M.: Suppression of the wrapping effect by Taylor model-based verified integrators: the single step. Int. J. Pure Appl. Math. (2007)
38. Mazze, R.S., Strock, E., Borgman, S., Wesley, D., Stout, P., Racchini, J.: Evaluating the accuracy, reliability, and clinical applicability of continuous glucose monitoring (CGM): Is CGM ready for real time? Diabetes Technol. Ther. **11**(1), 11–18 (2009)

39. Medtronic: Medtronic Launches MiniMed 640G System, Breakthrough in Artificial Pancreas Technology, Outside the U.S. http://newsroom.medtronic.com/phoenix.zhtml?c=251324&p=irol-newsArticle&ID=2009123. Accessed 18 May 2015
40. Moore, R.E.: Interval Analysis. Prentice-Hall, Englewood Cliffs (1966)
41. Moore, R.E.: Methods and Applications of Interval Analysis. SIAM, Philadelphia (1979)
42. Neumaier, A.: The wrapping effect, ellipsoid arithmetic, stability and confidence regions. Valid. Numer. Comput. Suppl. **9**, 175–190 (1993)
43. Palerm, C.C., Rodríguez-Fernández, M., Bevier, W.C., Zisser, H., Banga, J.R., Jovanovic, L., Doyle, F.J.: Robust parameter estimation in a model for glucose kinetics in type 1 diabetes subjects. In: Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE Engineering in Medicine and Biology Society Conference, pp. 319–322 (2006)
44. Percival, M.W., Bevier, W.C., Wang, Y., Dassau, E., Zisser, H.C., Jovanovic, L., Doyle, F.J.: Modeling the effects of subcutaneous insulin administration and carbohydrate consumption on blood glucose. J. Diabetes Sci. Technol. **4**(5), 1214–1228 (2010)
45. de Pereda, D., Romero-Vivo, S., Ricarte, B., Bondia, J.: On the prediction of glucose concentration under intra-patient variability in type 1 diabetes: a monotone systems approach. Comput. Methods Programs Biomed. **108**(3), 993–1001 (2012)
46. Puig, V., de Oca, S.M., Blesa, J.: Adaptive threshold generation in robust fault detection using interval models: time-domain and frequency-domain approaches. Int. J. Adap. Signal Process. **27**(10), 873–901 (2013)
47. Puig, V., Stancu, A., Quevedo, J.: Simulation of uncertain dynamic systems described by interval models: a survey. In: 16th IFAC World Congress (2005)
48. Rossetti, P., Ampudia-Blasco, F.J., Laguna, A., Revert, A., Vehí, J., Ascaso, J.F., Bondia, J.: 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), 1043–1052 (2012)
49. Sainz, M.A., Armengol, J., Calm, R., Herrero, P., Jorba, L., Vehi, J.: Modal Interval Analysis: New Tools for Numerical Information. Springer, New York (2014)
50. Smith, H.L.: Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems. American Mathematical Society, Providence (2008)
51. Sontag, E.D.: Monotone and near-monotone biochemical networks. Syst. Synth. Biol. **1**(2), 59–87 (2007)
52. Ståhl, F., Johansson, R.: Diabetes mellitus modeling and short-term prediction based on blood glucose measurements. Math. Biosci. **217**, 101–117 (2009)
53. Zschornack, E., Schmid, C., Pleus, S., Link, M., Klötzer, H.M., Obermaier, K., Schoemaker, M., Strasser, M., Frisch, G., Schmelzeisen-Redeker, G., Haug, C., Freckmann, G.: Evaluation of the performance of a novel system for continuous glucose monitoring. J. Diabetes Sci. Technol. **7**(4), 815–823 (2013)

Modeling and Prediction Using Stochastic Differential Equations

Rune Juhl, Jan Kloppenborg Møller, John Bagterp Jørgensen and Henrik Madsen

Abstract Pharmacokinetic/pharmakodynamic (PK/PD) modeling for a single subject is most often performed using nonlinear models based on deterministic ordinary differential equations (ODEs), and the variation between subjects in a population of subjects is described using a population (mixed effects) setup that describes the variation between subjects. The ODE setup implies that the variation for a single subject is described by a single parameter (or vector), namely the variance (covariance) of the residuals. Furthermore the prediction of the states is given as the solution to the ODEs and hence assumed deterministic and can predict the future perfectly. A more realistic approach would be to allow for randomness in the model due to e.g., the model be too simple or errors in input. We describe a modeling and prediction setup which better reflects reality and suggests stochastic differential equations (SDEs) for modeling and forecasting. It is argued that this gives models and predictions which better reflect reality. The SDE approach also offers a more adequate framework for modeling and a number of efficient tools for model building. A software package (CTSM-R) for SDE-based modeling is briefly described.

1 Introduction

Pharmacokinetic/pharmakodynamic (PK/PD) modeling is often performed using nonlinear mixed effects models based on deterministic ordinary differential equations (ODEs), [24]. The ODE models the dynamics of the system as

$$\begin{aligned} \frac{dX}{dt} &= f(X(t), t) \\ y_k &= X(t_k) + e_k, \end{aligned}$$

R. Juhl (✉) · J.K. Møller · J.B. Jørgensen · H. Madsen
Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark
e-mail: ruju@dtu.dk

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*, Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_10

where $X(t)$ is the state of the system, $f(\cdot)$ the model, y_k the discrete observations, and e_k the measurement errors which are assumed independent and identically distributed (iid) Gaussian. Given an initial value, the solution to the ODE $X(t)$ is a perfect prediction of all future values. The ODE model is an input–output model, where the residuals are the difference between the solution to the ODE and the observations. In the population setup, this implies that the total variation in data for a population of individuals is split into inter- and intraindividual variation. However, due to the ODE framework, the interindividual variation can only come from the covariance of the iid residuals, i.e., there must be no autocorrelation in the residuals.

The ODE framework is built on the assumption that future values of the states $X(t)$ can be predicted exactly and that the residual error is independent of the prediction horizon. This is often too simplistic and implies that the uncertainty about future values of the states and observations is not adequately described. This again has consequences for the design of model-based controllers and proper planning of medical treatments in general.

The ODE-based model class has a restricted residual error structure, as it assumes serially uncorrelated prediction residuals. There are several reasons why this assumption is violated: (1) misspecification or approximations of the structural model due to the complexity of the biological system, (2) unrecognized inputs, and (3) unpredictable random behavior of the process due to measurement errors for the input variables (e.g., specification of meals or physical exercise; both factors are known to influence future values of the blood glucose). In addition to these issues, the intraindividual (residual) variability also accounts for various environmental errors such as those associated with assay, dosing, and sampling. Since most of these errors cannot be considered as uncorrelated measurement errors, the description of the total individual error should preferably be separated (see also [7, 13]). Furthermore, [8] describe three types of residual error models to population PK/PD data analysis to account for more complicated residual error structures.

Neglecting the correlated residuals in the model description not only leads to serious issues when the model is used for forecasting and control as mentioned above, but it also disables a possibility for using proper methods for statistical model validation, parameter testing, and model identification (see e.g., [16], pp. 46–47).

In this chapter, stochastic differential equations (SDEs) are introduced to address these issues. SDEs facilitate the ability to split the intraindividual error into two fundamentally different types: (1) serially uncorrelated measurement error typically caused by assay error and (2) system error caused by model and input misspecifications. The concept will first be studied for a single subject and later on in a mixed effects setup with a population of individuals.

The use of SDEs opens up for new tools for model development, as it quantifies the amount of system and measurement noise. Specifically the approach allows for tracking of unknown inputs and parameters overtime by modeling them as random walk processes. These principles lead to efficient methods for pinpointing model deficiencies, and subsequently for identifying model improvements. The SDE approach also provides methods for model validation. This modeling framework is often called gray box modeling [25].

In this study, we will use maximum likelihood techniques both for parameter estimation and for model identification, and both for a single subject and in the population setting. It is known that parameter estimation in nonlinear mixed effects models with SDEs is most effectively carried out by considering an approximation to the population likelihood function. The population likelihood function is then approximated using the first-order conditional estimation (FOCE) method, which is based on a second-order Taylor expansion of each individual likelihood function at its optimum—see [17]. Like in [12], the extended Kalman filter is used for evaluating the single subject likelihood function.

This algorithm introduces a two-level numerical optimization, since not only the population likelihood function has to be maximized, but also for each value of the population likelihood all the individual likelihood functions must be maximized. This makes estimation computationally demanding, but the algorithm facilitates parallelization at several places to reduce the estimation time. The method is implemented in the R-package CTSM-R (continuous time stochastic modeling in R) [2], which is used in the DIACON project [3] focusing on technologies for semi- and fully-automatic insulin administration for treatment of type 1 diabetes. This project takes advantage of the fact that the SDE approach provides probabilistic forecasts for future values of the system states, which is crucial for reliable semi- and fully-automatic (closed-loop) insulin administration using model predictive control.

Section 2 describes various scenarios for data (single subject, repeated experiments, and populations of subjects), and how the likelihood function is formulated for each of these scenarios. Section 3 describes the approach used for population data from an experiment conducted in DIACON. Some practical issues related to SDE-based modeling are discussed in Sect. 4, and finally Sect. 5 summarizes. Both simulated and real-life experimental data are used throughout the chapter for illustrating the modeling and prediction framework.

2 Data and Modeling

Experiments can be conducted in various ways and the appropriate modeling approach depends on this. The basics start with a single experiment (solid ellipse in Fig. 1) which results in a series of data points $\mathscr{Y}$ sampled, possibly irregularly, at times $t_1 < t_2 < \cdots < t_N$. This single time series and how it is modeled are described in Sect. 2.1. Repeating the same experiment multiple times (dashed ellipse in Fig. 1) may be modeled as independent data series assuming no random effects between the runs. This is described in Sect. 2.2. When an experiment is done using several subjects (dotted ellipse in Fig. 1), then it is normal to include random effects between them. This is the so called population extension which is described in Sect. 2.3. In addition to the structure of data prior information may be available or used as a modeling technique. This is described in Sect. 2.4.

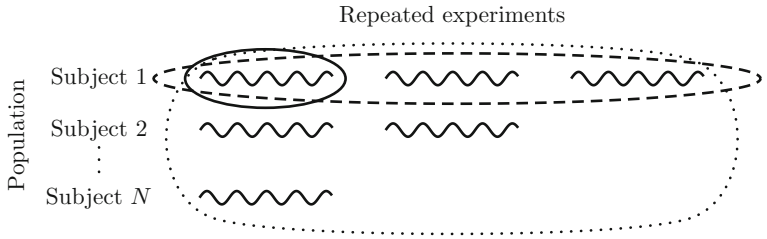

Fig. 1 A scenario of experiments in a study. The *solid ellipse* is a single time series trial. The *dashed ellipse* is a collection of three possibly independent repeated trials of subject 1. The *dotted ellipse* is a collection of subjects with random variation from a population

2.1 Single Data Series

This section begins by introducing a fundamental framework describing how to model physical phenomena. The aim is to provide a probabilistic model for a discrete time series $\mathscr{Y}_N = Y_1, Y_2, \ldots, Y_N$. The formulation in this section is a general framework which is useful for all types of correlated time series data and not just physiological data.

The natural extension to the ODE framework is SDE's. We begin by introducing the stochastic process $\mathbf{x}_t$ which satisfies an Itô SDE

$$d\mathbf{x}_t = \mathbf{f}\left(\mathbf{x}_t, \mathbf{u}_t, t, \boldsymbol{\theta}\right) dt + \boldsymbol{\sigma}\left(\mathbf{x}_t, \mathbf{u}_t, t, \boldsymbol{\theta}\right) d\boldsymbol{\omega}_t \,, \tag{1}$$

where $\mathbf{x}_t$ is the state, $\mathbf{u}_t$ is an exogenous input, and $\boldsymbol{\theta}$ the parameters of the model. $\mathbf{f}()$ and $\boldsymbol{\sigma}()$ are possibly nonlinear functions called the drift and diffusion terms. $\boldsymbol{\omega}$ is the Wiener process driving the stochastic part of the process. (1) describes the dynamics and is called the system equation. Note that the ODE model is contained within the SDE when removing the diffusion term $\boldsymbol{\sigma}\left(\mathbf{x}_t, \mathbf{u}_t, t, \boldsymbol{\theta}\right) d\boldsymbol{\omega}_t$.

The solution to the SDE (1) is not in general known except for linear and a few other SDE's. Many methods for solving SDE's have been proposed, e.g., Hermite expansions, simulation-based methods and Kalman filtering, see [4]. This chapter focuses on the Kalman filter using CTSM-R. The Kalman filter restricts the diffusion to being independent of the states because the approximations required to integrate an SDE with state-dependent diffusion give undesirable results or performance. However, some SDEs with state-dependent diffusion can be transformed to an SDE with unit diffusion by the Lamperti transform (see Sect. 4.1).

The stochastic process is observed discretely and possibly partially with independent noise via the measurement equation

$$\mathbf{y}_k = \mathbf{h}\left(\mathbf{x}_k, \mathbf{u}_k, t_k, \boldsymbol{\theta}, \mathbf{e}_k\right), \tag{2}$$

where $\mathbf{h}()$ is a possibly nonlinear function of the states and inputs. $\mathbf{e}_k$ is an independent noise term attributed by the imperfect measurements. Due to the Kalman filter, the

measurement model is restricted to additive noise in CTSM-R

$$\mathbf{y}_k = \mathbf{h}\left(\mathbf{x}_k, \mathbf{u}_k, t_k, \boldsymbol{\theta}\right) + \mathbf{e}_k, \tag{3}$$

where $\mathbf{e}_k$ is Gaussian with $\mathcal{N}(0, \mathbf{S}(\mathbf{u}_k, t_k))$.

The combination (1) and (3) is the state space model formulation used in this paper to understand data. This is a gray box model as it bridges the gap between data driven black box models and pure physical white box models.

Example 1 As an example to illustrate the methods, we will use a simulation example (see Fig. 2). A linear 3 compartment transport model [15] similar to the real-data modeling example presented in Sect. 3 is used. We can think of the response (y) as venous glucose concentration in the blood of a patient, and the input (u) as exogenous glucagon.

The data are simulated according to the model

$$d\mathbf{x}_t = \left(\begin{bmatrix} u_t \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -k_a & 0 & 0 \\ k_a & -k_a & 0 \\ 0 & k_a & -k_e \end{bmatrix} \mathbf{x}_t \right) dt + \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} d\boldsymbol{\omega}_t \tag{4}$$

$$y_k = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{t_k} + e_k, \tag{5}$$

where $\mathbf{x} \in \mathbb{R}^3$, $e_k \sim \mathcal{N}(0, s^2)$, $t_k = \{1, 11, 21, \ldots\}$, and the specific parameters ($\boldsymbol{\theta}$) used for simulation are given in Table 1 (first column).

The structure of the model (4) will of course usually be hidden, and we will have to identify the structure based on the measurements as given in Fig. 2. As a general

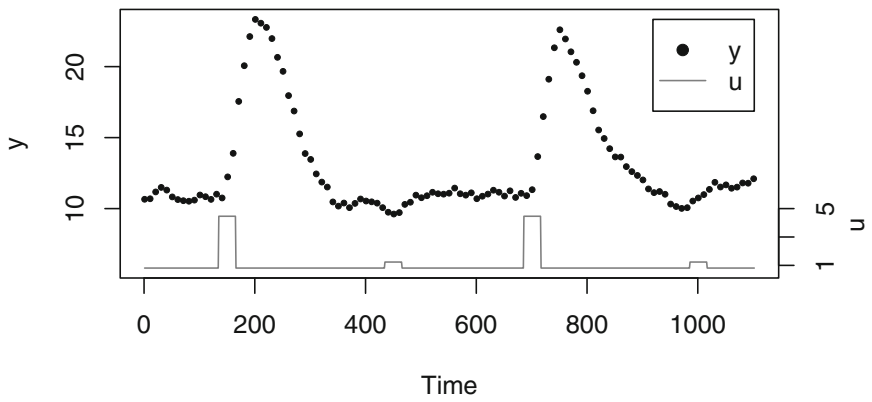

Fig. 2 Simulated data for the example (Eqs. (4), (5), and Table 1)

principle simple models are preferred over more complex models, and therefore a first hypothesis could be (Model 1)

$$dx_t = (u_t - k_e x_t)\, dt + \sigma_3 d\omega_t \tag{6}$$

$$y_k = x_{t_k} + e_k. \tag{7}$$

In this approach, the estimation is based on the likelihood function as defined in the following section.

2.1.1 Likelihood

Given a sequence of measurements

$$\mathscr{Y}_N = [\mathbf{y_0}, \mathbf{y_1}, \ldots, \mathbf{y_k}, \ldots, \mathbf{y}_N] \tag{8}$$

the likelihood of the unknown parameters θ given the model formulated as (1)–(3) is the joint probability density function (pdf)

$$L(\boldsymbol{\theta}, \mathscr{Y}_N) = \mathrm{p}(\mathscr{Y}_N|\boldsymbol{\theta}), \tag{9}$$

where the likelihood L is the probability density function given θ. The joint probability density function is partitioned as the product of the one-step conditional probability functions

$$L(\boldsymbol{\theta}, \mathscr{Y}_N) = \left(\prod_{k=1}^{N} \mathrm{p}\left(\mathbf{y}_k|\mathscr{Y}_{k-1}, \boldsymbol{\theta}\right)\right) \mathrm{p}(\mathbf{y}_0|\boldsymbol{\theta}). \tag{10}$$

The solution to a linear SDE driven by a Brownian motion is a Gaussian process. Nonlinear SDEs do not result in a Gaussian process and thus the marginal probability is not Gaussian. By sampling, the nonlinearities fast enough in some sense then it is reasonable to assume that the conditional density is Gaussian.

The Gaussian density is fully described by the first and second-order moments

$$\hat{\mathbf{y}}_{k|k-1} = E\left[\mathbf{y}_k|\mathscr{Y}_{k-1}, \boldsymbol{\theta}\right] \tag{11}$$

$$\Sigma_{k|k-1} = V\left[\mathbf{y}_k|\mathscr{Y}_{k-1}, \boldsymbol{\theta}\right]. \tag{12}$$

Introducing the innovation error

$$\boldsymbol{\varepsilon}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}, \tag{13}$$

the likelihood (10) becomes

$$L(\boldsymbol{\theta}, \mathscr{Y}_N) = \left(\prod_{k=1}^{N} \frac{\exp\left(-\frac{1}{2}\varepsilon_k^T \Sigma_{k|k-1}^{-1} \varepsilon_k\right)}{\sqrt{|\Sigma_{k|k-1}|\sqrt{2\pi}^l}} \right) \mathrm{p}(\mathbf{y}_0|\boldsymbol{\theta}). \tag{14}$$

The probability density of the initial observation $p(y_0|\boldsymbol{\theta})$ is parameterized through the probability density of the initial state $p(x_0|\boldsymbol{\theta})$. The mean $\hat{\mathbf{y}}_{k|k-1}$ and covariance $\Sigma_{k|k-1}$ are computed recursively using the extended Kalman filter, see Appendix A for a brief description, or [10] for a detailed description.

The unknown parameters are estimated by maximizing the likelihood function using an optimization algorithm. The likelihood (14) is a product of probability densities all less than 1, which causing numerical problems. Taking the logarithm of the likelihood (14) turns the product into a summation and cancels the exponentials thus stabilizing the calculation. The parameters are now found by maximizing the log-likelihood or by convention minimize the negative log-likelihood

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \left(-\ln\left(L(\boldsymbol{\theta}, \mathscr{Y}_N)\right)\right). \tag{15}$$

The uncertainty of the maximum likelihood parameter estimate $\hat{\boldsymbol{\theta}}$ is related to the curvature of the likelihood function. An estimate of the asymptotic covariance of $\hat{\boldsymbol{\theta}}$ is the inverse of the observed Fisher information matrix

$$V\left[\hat{\boldsymbol{\theta}}\right] = \left[\mathbf{I}\left(\hat{\boldsymbol{\theta}}\right)\right]^{-1}, \tag{16}$$

where $\mathbf{I}\left(\hat{\boldsymbol{\theta}}\right)$ is the observed Fisher information matrix, that is the negative Hessian matrix (curvature) of the likelihood function evaluated at the maximum likelihood estimate [17, 21].

Example 2 We continue with the simulated data from Example 1. As noted above, a first approach to model the data could be a first-state model (Eqs. (6)–(7)). The result of the estimation ($\hat{\boldsymbol{\theta}}_1$) is given in Table 1, the initial value of the state (x_{30}) and the time constant ($1/k_e$) are both captured quite well, while the uncertainty parameters are way off, the diffusion is too large and the observation variance is too small (with extremely large uncertainty).

The parameters in the model are all assumed to be greater than zero, and it is therefore advisable to estimate parameters in the log-domain, and then transform back to the original domain before presenting the estimates. The log-domain estimation is also the explanation for the nonsymmetric confidence intervals in Table 1, the confidence intervals are all based on the Hessian of the likelihood at the optimal parameter values, and confidence intervals are based on the Wald confidence interval in the transformed (log) domain [21]. Such intervals could be refined using profile likelihood-based confidence intervals [21] (see also Sect. 4.4).

Table 1 Parameter estimates from simulation example and confidence intervals for the individual parameters are given in parenthesis below the estimates

	θ	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$
x_{10}	40.000	–		38.819 (29.172, 48.466)
x_{20}	35.000	–	107.960 (75.211, 140.710)	33.421 (29.778, 37.064)
x_{30}	11.000	10.657 (6.606, 14.708)	10.641 (10.392, 10.889)	10.604 (10.281, 10.927)
k_a	0.025	–	0.006 (0.0038, 00778)	0.026 (0.025, 0.027)
k_e	0.080	0.081 (0.071, 0.094)	0.056 (0.0418, 0.0743)	0.080 (0.078, 0.083)
σ_1	1.000	–	–	0.5500 (0.224, 1.353)
σ_2	0.200	–	3.616 (2.670, 4.898)	0.282 (0.113, 0.704)
σ_3	0.050	2.206 (1.848, 2.634)	0.001 $(2 \cdot 10^{-55}, 3 \cdot 10^{48})$	0.001 $(9 \cdot 10^{-56}, 1 \cdot 10^{49})$
s	0.025	0.0002 $(2 \cdot 10^{-33}, 2.6 \cdot 10^{25})$	0.016 (0.0065, 0.0388)	0.031 (0.020, 0.049)
$l(\hat{\theta}, \mathbf{y})$	–	–343.68	–67.85	–19.70
df	–	4	7	9

θ is the true values, $\hat{\theta}_1$, $\hat{\theta}_2$, and $\hat{\theta}_3$ are the estimated for the first-, second-, and third-state models, respectively. Last two rows present the log-likelihood and the number of degrees of freedom

In order to validate the model and suggest further development, we should inspect the innovation error. When the model is not time homogeneous, the standard error of the prediction will not be constant and the innovation error should be standardized

$$r_k = \frac{\varepsilon_k}{\sqrt{\Sigma_{k|k-1}}}, \tag{17}$$

where the innovation error (ε_k) is given in (13). All numbers needed to calculate the standardized residuals can be obtained directly from CTSM-R using the function `predict`. Both the autocorrelation and partial autocorrelation (Fig. 3) are significant in lag 1 and 2. This suggests a second-state model for the innovation error, and hence a third-state model should be used. Consequently we can go directly from the first-state model to the true structure (a third-state model).

Now we have assumed that a number of the parameters are actually zero, in a real-life situation, we might test these parameters using likelihood ratio tests, or indeed identify them through engineering principles. The parameter estimates are given in Table 1 ($\hat{\theta}_3$); in this case, the diffusion parameter (σ_3) has an extremely

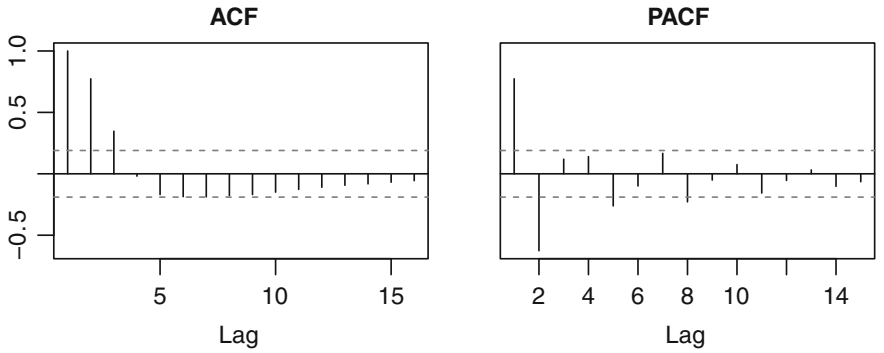

Fig. 3 Autocorrelation and partial autocorrelation from a simple (1 state) model

wide confidence interval, and it could be checked if these parameters should indeed be zero (again using likelihood ratio test), but for now we will proceed with the residual analysis which is an important part of model validation (see e.g., [16]). The autocorrelation and partial autocorrelation for the third-state model are shown in Fig. 4. We see that there are no values outside the 95 % confidence interval, and we can conclude that there is no evidence against the hypothesis of white noise residuals, i.e., the model sufficiently describes the data.

Autocorrelation and partial autocorrelations are based on short-term predictions (in this case 10 min) and hence we check the local behavior of the model. Depending on the application of the model, we might be interested in longer-term behavior of the model. Prediction can be made on any horizon using CTSM-R. In particular, we can compare deterministic simulation in CTSM-R (meaning conditioning only on the initial value of the states). Such a simulation plot is shown in Fig. 5, here we

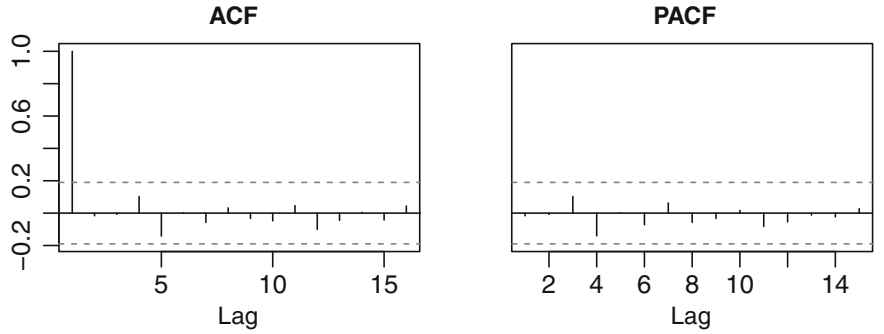

Fig. 4 Autocorrelation and partial autocorrelation from the third-state model (i.e., the correct model)

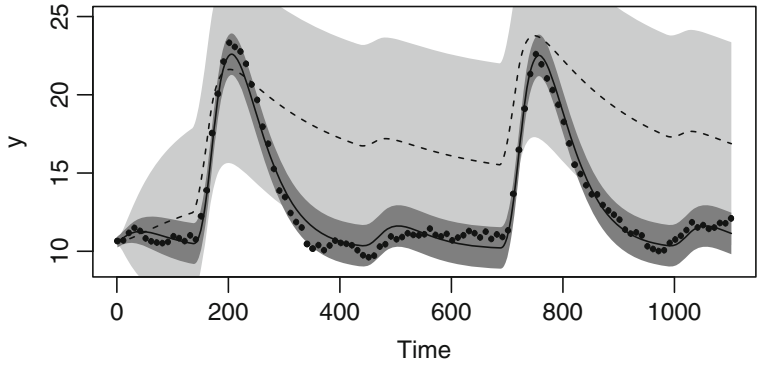

Fig. 5 Simulation with model 2 and 3, *dashed gray line* expectation of model 2, *black line* expectation of model 3, *light gray area* 95 % prediction interval for model 2, *dark gray area* 95 % prediction interval for model 3, and *black dots* are the observations

compare a second-state model (see Table 1) with the true third-state model. It is quite evident that model 2 is not suited for simulation, with the global structure being completely off, while "simulation" with a third-state model (with the true structure, but estimated parameters), gives narrow and reasonable simulation intervals. In the case of linear SDE-models with linear observation, this "simulation" is exact, but for nonlinear models it is recommended to use real simulations, e.g., using a Euler scheme.

The step from a second-state model (had we initialized our model development with a second-state model) to the third-state model is not at all trivial. However, Fig. 5 shows that simulation of model 2 does not contain the observations and thus model 2 will not be well suited for simulations. Also the likelihood ratio test (or AIC/BIC) supports that model 3 is far better than model 2, further it would be reasonable to fix σ_3 at zero (in practice a very small number).

2.2 Independent Data Series

An experiment may be repeated several times without expecting variation in the underlying parameters. Given S sequences of possibly varying length

$$\mathbf{Y} = \left[\mathcal{Y}_{N_1}^1, \mathcal{Y}_{N_2}^2, \ldots, \mathcal{Y}_{N_i}^i, \ldots, \mathcal{Y}_{N_S}^S\right], \tag{18}$$

the likelihood is the product of the likelihood (10) for each sequence

$$L(\boldsymbol{\theta}, \mathbf{Y}) = \prod_{i=1}^{S} \left(\left(\prod_{k=1}^{N} \frac{\exp\left(-\frac{1}{2}\varepsilon_k^T \Sigma_{k|k-1}^{-1} \varepsilon_k\right)}{\sqrt{|\Sigma_{k|k-1}|}\sqrt{2\pi}^{l}} \right) \mathrm{p}(\mathbf{y}_{0,i}|\boldsymbol{\theta}) \right). \tag{19}$$

The unknown parameters are again estimated by minimizing the negative log-likelihood

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \left(-\ln\left(L(\boldsymbol{\theta}, \mathbf{Y})\right)\right). \tag{20}$$

If the independence assumption is violated and the parameters vary between the time series, then the model performance would be lowered as the parameter estimates will be a compromise. The natural extension is to include a population effect.

2.3 Population Extension

The gray box model can be extended to include a hierarchical structure to model variation occurring between data series where each series has its own parameter set. This is useful for describing data from a number of individuals belonging to a population of individuals. The hierarchical modeling is also called mixed effects and population extension in pharmaceutical science. Nonlinear mixed effects modeling has long been used in pharmacokinetic/pharmacodynamic studies to account for variation from the natural grouping: multiple centers, multiple days, age and BMI of subjects, etc. Mixed effects modeling combines fixed and random effects [17]. The fixed effect is the average of that effect over the entire population while the random effect allows for variation around that average.

Consider N subjects in a clinical study. This is a single level grouping. The model for the ith subject is

$$d\mathbf{x}_{i,t} = \mathbf{f}\left(\mathbf{x}_{i,t}, \mathbf{u}_{i,t}, t, \boldsymbol{\theta}_i\right) dt + \boldsymbol{\sigma}\left(\mathbf{u}_{i,t}, t, \boldsymbol{\theta}_i\right) d\boldsymbol{\omega}_t \tag{21}$$

$$\mathbf{y}_{i,k} = \mathbf{h}\left(\mathbf{x}_{i,k}, \mathbf{u}_{i,k}, t_{i,k}, \boldsymbol{\theta}_i\right) + \mathbf{e}_{i,k}, \tag{22}$$

which is the general model extended with subscript i. The individual parameters $\boldsymbol{\theta}_i$ are

$$\boldsymbol{\theta}_i = z(\boldsymbol{\theta}_f, \mathbf{Z}_i, \boldsymbol{\eta}_i), \tag{23}$$

where z maps from subject covariates such (i.e., BMI and age) $\mathbf{Z}_i$, fixed effects parameters θ_f, and the random effects $\eta_i \in R^k \sim \mathcal{N}(0, \Omega)$ to subject parameters. The subject parameters are typically modeled as either normally or log-normally distributed by combining the fixed effect parameters and the random effects in either an additive $\theta_i = \theta_f + \eta_i$ or an exponential transform $\theta_i = \theta_f e^{\eta_i}$.

The likelihood of the fixed effects is the product of the marginal probability densities for each subject

$$L(\boldsymbol{\theta}_f, \Omega) = \prod_{i=1}^{N} p(\mathscr{Y}_i|\boldsymbol{\theta}_i, \Omega), \tag{24}$$

where the marginal density is found by integrating over the random effects $\boldsymbol{\eta}_i$

$$p(\mathscr{Y}_i|\boldsymbol{\theta}_i, \Omega) = \int p_1(\mathscr{Y}_i|\boldsymbol{\theta}_i, \boldsymbol{\eta}_i) p_2(\boldsymbol{\eta}_i|\Omega) d\boldsymbol{\eta}_i. \tag{25}$$

$p_1(\mathscr{Y}_i|\boldsymbol{\theta}, \boldsymbol{\eta})$ is the probability of the individual subject which given by (10). $p_2(\eta_i|\Omega)$ is the probability of the second-stage model where the random effects describe the interindividual variation.

2.3.1 Approximation of the Marginal Density

The integral in (25) rarely has a closed-form solution and thus must be approximated in a computationally feasible way. This can be done in two ways: approximating (a) the integrand by Laplacian or (b) the entire integral by Gaussian quadrature.

Gaussian quadrature can approximate the integral by a weighted sum of the integrand evaluated at specific nodes. The accuracy of Gaussian quadrature increases as the order (number of nodes) increases. With adaptive Gaussian quadrature, the accuracy can be improved even further at higher cost. The computational complexity of Gaussian quadrature suffers from the curse of dimensionality and becomes infeasible even for few dimensions.

Now consider the Laplacian approximation which is widely used approximation to integrals [17]. Observe that the integrand in (25) is nonnegative such that

$$\begin{aligned} p_1(\mathscr{Y}_i|\boldsymbol{\theta}_i, \boldsymbol{\eta}_i) p_2(\boldsymbol{\eta}_i|\Omega) &= e^{\log(p_1(\mathscr{Y}_i|\boldsymbol{\theta}_i, \boldsymbol{\eta}_i) p_2(\boldsymbol{\eta}_i|\Omega))} \\ &= e^{g_i(\eta_i)}, \end{aligned} \tag{26}$$

where $g_i(\boldsymbol{\eta}_i)$ is the log-posterior distribution for the ith subject. Now consider the second-order Taylor expansion of $g_i(\boldsymbol{\eta}_i)$ around its mode $\hat{\boldsymbol{\eta}}_i$

$$g_i(\boldsymbol{\eta}_i) \approx g_i(\hat{\boldsymbol{\eta}} i) + \frac{1}{2} \left(\boldsymbol{\eta}_i - \hat{\boldsymbol{\eta}}_i\right)^T \Delta g_i\left(\hat{\boldsymbol{\eta}}_i\right) \left(\boldsymbol{\eta}_i - \hat{\boldsymbol{\eta}}_i\right), \tag{27}$$

since $\nabla g_i\left(\hat{\boldsymbol{\eta}}_i\right) = 0$ at the mode. By inserting (27) and (26) in (25), the Laplacian approximation of the marginal probability density is defined as

$$\begin{aligned} p(\mathscr{Y}_i|\boldsymbol{\theta}_f, \Omega) &\approx \int e^{g_i(\hat{\boldsymbol{\eta}}_i)+\frac{1}{2}(\boldsymbol{\eta}_i-\hat{\boldsymbol{\eta}}_i)^T \Delta g_i(\hat{\boldsymbol{\eta}}_i)(\boldsymbol{\eta}_i-\hat{\boldsymbol{\eta}}_i)} d\boldsymbol{\eta}_i \\ &= e^{g_i(\hat{\boldsymbol{\eta}}_i)} \int e^{\frac{1}{2}(\boldsymbol{\eta}_i-\hat{\boldsymbol{\eta}}_i)^T \Delta g_i(\hat{\boldsymbol{\eta}}_i)(\boldsymbol{\eta}_i-\hat{\boldsymbol{\eta}}_i)} d\boldsymbol{\eta}_i, \end{aligned} \tag{28}$$

where the integral is recognized as the scaled integral over a multivariate Gaussian[1] distribution with covariance $\Sigma = (-\Delta g(\eta_i))^{-1}$. The marginal density becomes

$$p(\mathscr{Y}_i|\boldsymbol{\theta}_i, \Omega) \approx e^{g_i(\hat{\boldsymbol{\eta}}_i)} \sqrt{\frac{(2\pi)^k}{|-\Delta g\left(\hat{\boldsymbol{\eta}}_i\right)|}}. \tag{29}$$

Inserting (29) in (24) the likelihood becomes

$$L(\theta_f, \Omega) \approx \prod_{i=1}^{N} e^{g_i(\hat{\boldsymbol{\eta}}_i)} \sqrt{\frac{(2\pi)^k}{|-\Delta g\left(\hat{\boldsymbol{\eta}}_i\right)|}}. \tag{30}$$

The Hessian $\Delta g(\hat{\boldsymbol{\eta}}_i)$ is found by analytically differentiating the expression for the log-posterior $g(\boldsymbol{\eta})$. After some derivation, the Hessian is

$$\begin{aligned} \Delta g(\boldsymbol{\eta}_i) \; = \sum_{k=1}^{N} \Bigg[& \frac{\partial^2 \mathbf{y}^T}{\partial \boldsymbol{\eta}_i \partial \boldsymbol{\eta}_i} \Sigma_{k|k-1}^{-1} \left(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}\right) + 2 \frac{\partial \hat{\mathbf{y}}_{k|k-1}}{\partial \boldsymbol{\eta}_i} \frac{\partial \left[\Sigma_{k|k-1}^{-1}\right]}{\partial \boldsymbol{\eta}_i} \left(\mathbf{y} - \hat{\mathbf{y}}_{k|k-1}\right) \\ & - \frac{\partial \hat{\mathbf{y}}_{k|k-1}}{\partial \boldsymbol{\eta}_i} \Sigma_{k|k-1}^{-1} \frac{\partial \hat{\mathbf{y}}_{k|k-1}}{\partial \boldsymbol{\eta}_i} - \frac{1}{2} \left(\mathbf{y} - \hat{\mathbf{y}}_{k|k-1}\right) \frac{\partial^2 \left[\Sigma_{k|k-1}^{-1}\right]}{\partial \boldsymbol{\eta}_i \partial \boldsymbol{\eta}_i} \left(\mathbf{y} - \hat{\mathbf{y}}_{k|k-1}\right) \\ & + \operatorname{tr}\left(\frac{\partial \left[\Sigma_{k|k-1}^{-1}\right]}{\partial \boldsymbol{\eta}_i} \frac{\partial \Sigma_{k|k-1}}{\partial \boldsymbol{\eta}_i} + \Sigma_{k|k-1}^{-1} \frac{\partial \Sigma_{k|k-1}}{\partial \boldsymbol{\eta}_i \partial \boldsymbol{\eta}_i} \right) \Bigg] - \Omega^{-1}, \end{aligned} \tag{31}$$

where tr is the trace of a matrix. The second-derivative terms are generally complicated or inconvenient to compute. At the mode $\hat{\boldsymbol{\eta}}_i$, the contribution of the second-derivative terms is usually negligible and thus an approximation for the Hessian is

$$\Delta g(\hat{\boldsymbol{\eta}}_i) \approx -\sum_{k=1}^{N} \left(\frac{\partial \hat{\mathbf{y}}_{k|k-1}}{\partial \boldsymbol{\eta}_i}\bigg|_{\boldsymbol{\eta}_i=\hat{\boldsymbol{\eta}}_i} \Sigma_{k|k-1}^{-1} \frac{\partial \hat{\mathbf{y}}_{k|k-1}}{\partial \boldsymbol{\eta}_i}\bigg|_{\boldsymbol{\eta}_i=\hat{\boldsymbol{\eta}}_i} \right) - \Omega^{-1}. \tag{32}$$

[1] The integral over the multivariate Gaussian density is $\frac{1}{\sqrt{(2\pi)^k|\Sigma|}} \int e^{\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)} d\mathbf{x} = 1$.

This approximation is similar to the Gauss–Newton and NONMEM's first-order conditional estimation (FOCE) approximations of the Hessian where only first partial derivatives are included [9, 17].

The parameters are found by iteratively minimizing the first- and second-stage model. For a trial set of fixed effect parameters, an optimization of g must be done for all subjects. When all η_i have been found, the Laplacian and FOCE approximations can be computed to obtain the population likelihood. The population likelihood can then be optimized.

2.4 Prior Information

Bayesian analysis combines the likelihood of the data and already known information which is called a prior. When the prior probability density function is updated, it becomes the posterior probability density function. In true, Bayesian analysis the prior may be any distribution, although conjugated priors are used in practice to simplify the computations.

In the view of CTSM-R, priors are mainly used as (a) empirical prior or for (b) regularizing the estimation.

An empirical prior is a result from a previous estimation. Imagine an experiment has been analyzed and followed by rerunning the experiment. These two data series are stochastically independent sets and should be analyzed as in Sect. 2.2. However, using the results from the first analysis as a prior, only the new data series has to be analyzed. If the quadratic Wald approximation holds this prior is Gaussian.

Regularizing one or more parameters is sometimes required to achieve a feasible estimation of the parameters. State equations describe a physical phenomenon and as such the modeler often has knowledge (possibly partly subjective) about the parameters from, e.g., another study. The reported values are often a mean and a standard deviance. Thus a Gaussian prior is reasonable.

Updating the prior probability density function $p(\boldsymbol{\theta})$ forms the posterior probability density function through Bayes' rule

$$p(\boldsymbol{\theta}|\mathscr{Y}_N) = \frac{p(\mathscr{Y}_N|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathscr{Y}_N)} \propto p(\mathscr{Y}_N|\boldsymbol{\theta})p(\boldsymbol{\theta}), \tag{33}$$

where the probability density $p(\mathscr{Y}_N|\boldsymbol{\theta})$ is proportional to the likelihood of a single data series given in (10). No information is called a diffuse prior which is uniform over the entire domain. The posterior then reduces to the likelihood of the data.

Let the prior be described by a Gaussian distribution $\mathscr{N}(\boldsymbol{\mu}_{\boldsymbol{\theta}}, \Sigma_{\boldsymbol{\theta}})$ where

$$\boldsymbol{\mu}_{\boldsymbol{\theta}} = E[\boldsymbol{\theta}] \tag{34}$$

$$\Sigma_{\boldsymbol{\theta}} = V[\boldsymbol{\theta}], \tag{35}$$

and let

$$\boldsymbol{\varepsilon}_{\boldsymbol{\theta}} = \boldsymbol{\theta} - \boldsymbol{\mu}_{\boldsymbol{\theta}}, \tag{36}$$

then the posterior probability density function is

$$p(\boldsymbol{\theta}|\mathscr{Y}_N) \propto \left(\left(\prod_{k=1}^{N} \frac{\exp\left(-\frac{1}{2}\varepsilon_k^T \Sigma_{k|k-1}^{-1} \varepsilon_k\right)}{\sqrt{|\Sigma_{k|k-1}|\sqrt{2\pi}^l}} \right) \mathrm{p}(\mathbf{y}_0|\boldsymbol{\theta}) \right) \times \frac{\exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}_{\boldsymbol{\theta}}^T \Sigma_{\boldsymbol{\theta}}^{-1} \boldsymbol{\varepsilon}_{\boldsymbol{\theta}}^T\right)}{\sqrt{|\Sigma_{\boldsymbol{\theta}}|\sqrt{2\pi}^p}}. \tag{37}$$

The parameters are estimated by maximizing the posterior density function (37), i.e., maximum a posteriori (MAP) estimation. The MAP parameter estimate is found by minimizing the negative logarithm of (37)

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \left(-\ln\left(p(\boldsymbol{\theta}|\mathscr{Y}_N, \mathbf{y}_0)\right)\right). \tag{38}$$

When there is no prior the MAP estimate reduces to the ML estimate.

3 Example: Modeling the Effect of Exercise on Insulin Pharmacokinetics in "Continuous Subcutaneous Insulin Infusion" Treated Type 1 Diabetes Patients

The artificial pancreas is believed to ease substantially the burden of constant management of type 1 diabetes for patients. An important aspect of the artificial pancreas development is the mathematical models used for control, prediction, and simulation. A major challenge to the realization of the artificial pancreas is the effect of exercise on the insulin and plasma glucose dynamics. This is the first step towards a population model of exercise effects in type 1 diabetes. The focus is on the effect on the insulin pharmacokinetics in continuous subcutaneous insulin infusion (CSII)-treated patients by modeling the absorption rate as a function of exercise. This example is described in detail in [5].

3.1 Data

The insulin data for this study originates from a clinical study on 12 subjects with type 1 diabetes treated with continuous subcutaneous insulin infusion (CSII). Each subject did two study days separated by at least three weeks. The insulin was observed by drawing blood nonequidistantly over the course of the trial. A detailed description of the data is found in [23].

Natural considerations toward the subjects limits how frequent the insulin can be sampled. This limits the amount of observations per time series and often care-

ful nonequidistant sampling becomes necessary. Both issues makes estimation of parameters more difficult. However, using all the subjects collectively increases the amount of data and improves estimation. The repeated trials per subject are considered independent trials, i.e., no random variation on the parameters. The subjects are assumed to have interindividual variation for several of the parameters.

3.2 The Gray Box Insulin Model

A linear three-compartment ODE model is used as basis to describe the pharmacokinetics of subcutaneous infused insulin in a single subject as suggested by [26]. The model is illustrated in Fig. 6.

The absorption is characterized by the rate parameter k_a between all three compartments. The two compartments Isc_2 and Ip are modeled with diffusion. Only the third-state Ip is being observed.

The compartment model is formulated as the following SDE

$$d\begin{bmatrix} Isc_1 \\ Isc_2 \\ Ip \end{bmatrix} = \left(\begin{bmatrix} -k_a & 0 & 0 \\ k_a & -k_a & 0 \\ 0 & \frac{k_a}{V_I} & -k_e \end{bmatrix} \begin{bmatrix} Isc_1 \\ Isc_2 \\ Ip \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} I_{pump} \right) dt + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sigma_{Isc} & 0 \\ 0 & 0 & \sigma_{Ip} \end{bmatrix} d\omega_t, \tag{39}$$

where Isc_1 [mU] and Isc_2 [mU] represent the subcutaneous layer and deeper tissues, respectively, and Ip [mU/L] represents plasma. I_{pump} is the input from the pump [mU/min]. k_a [min^{-1}] is the absorption rate and k_e [min^{-1}] is the clearance rate of insulin from plasma. V_I is the volume of distribution [L]. σ_{Isc} and σ_{Ip} are the standard deviation of the diffusion processes.

The observation equation is formulated through a transformation of the third-state I_p. The log transformation used here is a natural choice since I_p is a concentration which is a nonnegative number. Transformations are discussed in Sect. 4.1. The observation equation is

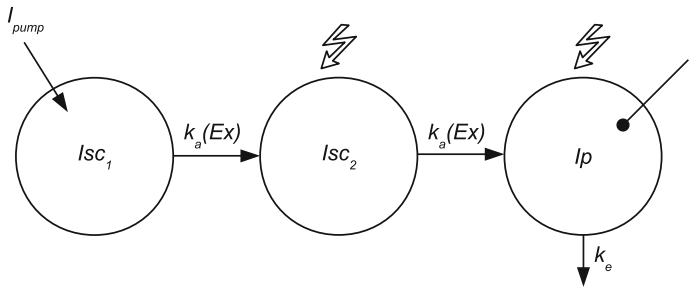

Fig. 6 Illustration of a three-compartment model describing the pharmacokinetics of insulin delivered continuously from an insulin pump. Lightning bolts indicate diffusion terms

$$\log(y_k) = \log(Ip_k) + e_k, \tag{40}$$

where y_k is the observed plasma insulin concentration and $e_k \sim N(0, \xi)$ is the measurement noise. The variance is further modeled such that $\xi = S_{\min} + S$, where $S_{\min}$ is a known hardware specific measurement error variance of the equipment [5]. Note that the measurement error multiplicative in the natural domain of y_k. This works as an approximation of a proportional error model.

The full gray box model is the SDE system equation (39) and the observation equation (40).

Population Parameters

The individual parameters are modeled as a combination of fixed population effects and random individual effects

$$\theta_i = h(\theta_{pop}, Z_i) \cdot e^{\eta_i}, \tag{41}$$

where θ_i is the parameter value for individual i, $h(\cdot)$ is a possibly nonlinear function, θ_{pop} is the overall population parameter (fixed effect), Z_i are covariates (age, weight, gender etc.), and $\eta_i \sim N(0, \Omega)$ is the individual random effect.

For this model, four parameters were modeled with a random effect. The initial values of the two subcutaneous layer states are assumed to be affected by the same variation from the population mean

$$Isc_{1_0,i} = Isc_{1_0} \cdot e^{\eta_{i,1}} \quad Isc_{2_0,i} = Isc_{2_0} \cdot e^{\eta_{i,1}}.$$

The absorption rate k_a and the clearance rate k_e have separate random effects

$$k_{a,i} = k_a \cdot e^{\eta_{i,2}} \quad k_{e,i} = k_e \cdot e^{\eta_{i,3}}.$$

The volume of distribution V_I is scaled by the weight (kg) of the subject. The weight is a covariate

$$V_{I_i} = V_I \cdot \text{weight}_i.$$

The random effects are assumed Gaussian with

$$\eta_i = [\eta_{i1}, \eta_{i2}, \eta_{i3}] \sim \mathcal{N}\left(0, \text{diag}\left(\omega_{Isc}, \omega_{k_a}, \omega_{k_e}\right)\right).$$

3.3 Exercise Effects

The model is further extended by making the absorption rate k_a dependent on exercising. Two extensions are investigated.

Model A

The first extension specifies k_a as

$$k_a = \bar{k}_a + \alpha \cdot \mathrm{Ex}, \tag{42}$$

where $\bar{k}_a$ is the basal rate and α is the effect of exercise. Ex is a binary input which is 1 when the subject is exercising and otherwise 0.

Model B

The subjects were exercising at two intensities and this extends (42) to

$$k_a = \bar{k}_a + \alpha_{\mathrm{mild}} \cdot \mathrm{Ex}_{\mathrm{mild}} + \alpha_{\mathrm{moderate}} \cdot \mathrm{Ex}_{\mathrm{moderate}}, \tag{43}$$

where $\bar{k}_a$ is the basal rate, α_{mild} and $\alpha_{\mathrm{moderate}}$ are the effects of mild and moderate exercise. $\mathrm{Ex}_{\mathrm{mild}}$ and $\mathrm{Ex}_{\mathrm{moderate}}$ are binary inputs which is 1 during either mild or moderate exercising.

3.4 Model Comparison

The best model is selected by comparing the ML estimates with the likelihood ratio test, AIC, and BIC in Table 2. The base model is nested in both model A and B and model A is nested in B. The nested models can be compared with the likelihood ratio test. Both models A and B explain significantly more of the variability in the data than the base model. Model A is the prefered model based on the likelihood ratio test. The additional improvement in the likelihood with model B is not enough to justify the extra parameter. The difference in AIC and BIC between model A and B relatively small but indicate that model B is to be prefered. The relative likelihood between model A and B is $\exp(0.5 \cdot (1815 - 1817)) = 0.37$ and suggests that model A is 37 % as probable as model B [1].

The parameter estimates for all three models are seen in Table 3. For model B, the moderate intensity exercise results in a larger absorption rate than mild exercise.

Table 2 Model comparison using likelihood ratio test, AIC and BIC

Model	df	$-\log(L)$	LRT (p)	AIC	BIC
Base	10	927	–	1878	1799
Model A versus Base	11	897	$<10^{-7}$	1817	1729
Model B versus A	12	895	0.16	1815	1720

Table 3 Parameter estimates from the three models: base, A and C

	Base	Model A	Model B
Isc_{1_0}	87.4	58.3	61.8
Isc_{2_0}	35.9	56.3	52.2
k_a	0.023	0.026	0.024
k_e	0.079	0.077	0.076
σ_{Isc}	2.94	2.61	2.48
σ_{Ip}	0.030	0.027	0.026
S	0.00028	0.00034	0.00075
ω_{Isc}	0.379	0.226	0.299
ω_{ka}	0.122	0.112	0.112
ω_{ke}	0.142	0.150	0.146
α		0.00762	
α_{mild}			0.00961
α_{moderate}			0.00515

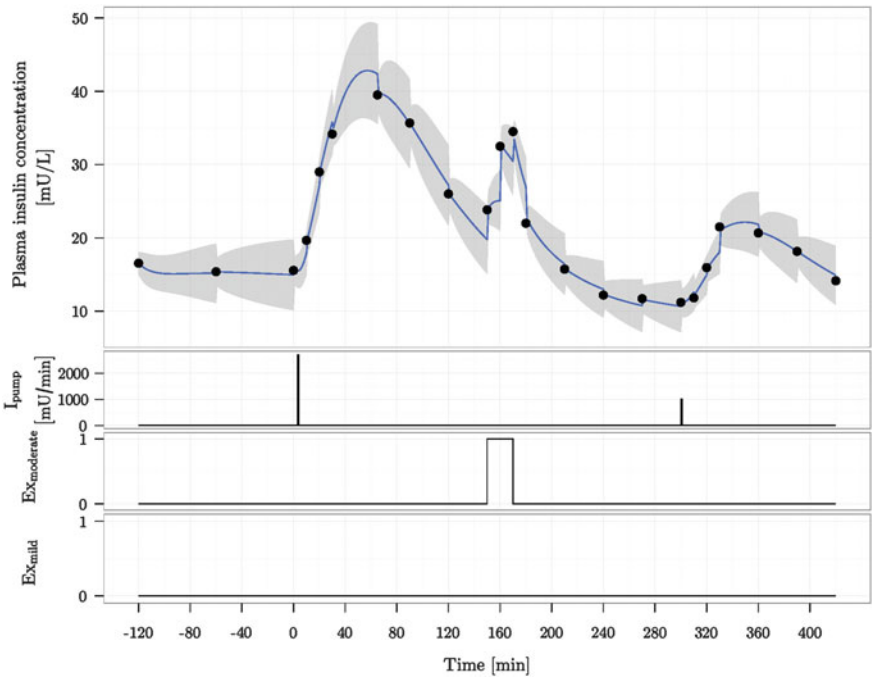

Fig. 7 *Top* One-step predictions from model A (*Blue line*). The observations are represented by dots. The *gray area* indicates 95% prediction interval. *Middle* and *bottom* Insulin and exercise inputs

3.5 Predictions

From the three models tried here, model A with a single absorption rate is the best to explain the data. One-step predictions using model A using a single trial of one subject are shown in Fig. 7. In general, the predictions are acceptable and the model does seem to capture the increase related to exercise. Especially, in Fig. 7, the compliance between the predictions and the observations is good. The width of the prediction interval is, however, large in this case. k-step predictions can also easily be calculated using CTSM-R and the `predict` function. A more detailed account of the exercise dependence analysis using population modeling is found in [5].

4 Other Topics

4.1 Transformations

In general, transformations should be applied whenever appropriate, and as all inference with CTSM-R assumes Gaussian random output, this should be ensured by transformations. Transformations can be applied in three different levels (1) state transformations, (2) transformation of observations, and (3) transformation of the parameters. We will briefly discuss each of these types of transformations and refer the interested reader to appropriate literature.

If there are natural restrictions of the state space, e.g., the natural state space is the positive real axis, or some interval, then these restrictions should be included in the SDE description. This implies a formulation of the form

$$dx_t = f(x_t, u_t)dt + \sigma(x_t)dw_t. \tag{44}$$

However, the Kalman filter requires the diffusion term to be independent of the state and therefore we should apply the Lamperti transform;

$$z_t = \int \frac{d\xi}{\sigma(\xi)}\bigg|_{\xi=x_t} \tag{45}$$

and use Itô's Lemma to obtain a description where the SDE description is independent of the state (see [18], Paper D for a tutorial on the Lamperti transform, and [19] for a nontrivial application).

The usual comments on transformation of the observations also apply to the SDE models, i.e., the standardized residuals should have constant variance, this should be checked and if the residuals do not have constant variance the observations should be transformed (e.g., using log transformation).

As already discussed in the examples in this chapter, the parameters should be estimated on the real axis (implying e.g., log transformation of positive parameters).

4.2 Identification

We have already seen that the autocorrelation function and the partial autocorrelation functions can be used for identification. If data are not equivalently sampled, one might use linear SDE models on the residuals to identify model order (number of states).

For nonlinear models the usual autocorrelation function is also relevant. Nonlinear dependence in the residuals will almost always include a linear dependence which will appear in the autocorrelation function. It can be shown that some nonlinear functions does not have linear dependence and the autocorrelation functions will fail. Generalizations in the form of lag-dependent and partial lag-dependent functions might then be used instead [20].

Finally identification can be based on random walk identification, where one parameter is formulated as a random walk process and the reconstruction or smoothed parameter is compared with state estimates and/or input to identify possible model extensions (see also [18, 19], paper F, and [11]).

4.3 Simulation/Prediction Models

As we have already seen in the simulation example, misspecification of a model can lead to very poor performance in simulation (long-term prediction) performance of models. A way to ensure reasonable performance in long-term predictions is by forcing the diffusion parameters to be small. This is done by fixing diffusion parameters, see [14] for a discussion about simulation and multistep predictions in SDE-models.

4.4 Testing and Confidence Intervals

Often, in particular for data-rich situations, the standard Wald confidence intervals, as presented directly from CTSM-R, are good approximations of the "true" confidence intervals. These are, however, approximations, and conclusions regarding individual parameters should be based on likelihood ratio tests rather than confidence intervals. In cases where models are not nested, it is recommended to use likelihood-based information criteria (AIC or BIC) for model selection.

Still, confidence intervals provide useful information that should always be reported, also when parameters are significant. But as we saw in the simulation examples, the Wald confidence interval might fail completely (e.g., σ_3 in Models 2 and 3). The problem is that the Wald standard error uses the local curvature of the likelihood (the Hessian), to approximate the uncertainty, and e.g., if the curvature

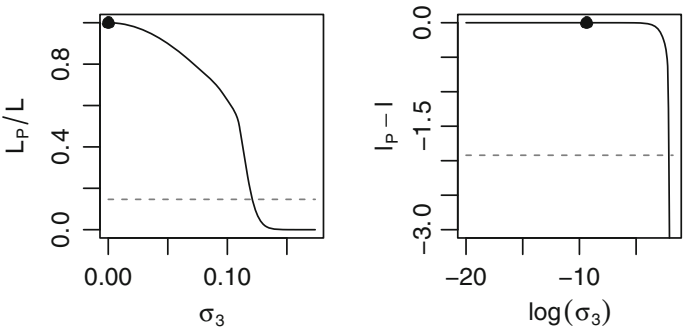

Fig. 8 Profile likelihood for σ_3 in the third-state simulation model of Examples 1–2

is close to zero (see Fig. 8), then the variance of the parameter estimates becomes infinite (as we saw in the examples).

As an alternative, we can calculate profile likelihood confidence intervals (see Fig. 8), we will not go into detail with the calculation of such intervals, but note that the profile likelihood confidence interval is based on the same statistical properties of the likelihood ratio as the likelihood ratio test. In the case of Model 3 of the simulation example, the profile likelihood confidence interval for σ_3 is $[0, 0.12]$, which seems much more reasonable than the values obtained by the Wald approximation. For further reading see [17, 21].

5 Summary

A general framework for modeling physical dynamical systems using stochastic differential equations has been demonstrated. CTSM-R is an efficient and parallelized implementation in the statistical language R. R facilitates easy data handling, visualization, and statistical tests essential for any modeling task. CTSM-R uses maximum likelihood and thus known techniques for model identification and selection can also be used for this framework as demonstrated.

This chapter has demonstrated the principles using linear models with transformations. CTSM-R has been used for a number of nonlinear problems see e.g., [18, 22].

CTSM-R has been extended to include hierarchical modeling. A study of exercise dependence in insulin absorption was modeled with a random effect between the subjects. This is an example of commonly used population modeling in PK/PD.

A detailed user guide and additional examples are available from http://ctsm.info.

Appendix A: Extended Kalman Filtering

For nonlinear models the innovation vectors $\boldsymbol{\varepsilon}_k$ (or $\boldsymbol{\varepsilon}_k^i$) and their covariance matrices $\Sigma_{k|k-1}^{yy}$ (or $\Sigma_{k|k-1}^{yy,i}$) can be computed recursively by means of the extended Kalman filter (EKF) as outlined in the following.

Consider first the linear time-varying model

$$d\mathbf{X}_t = (\mathbf{A}(\mathbf{u}_t, t, \boldsymbol{\theta})\mathbf{X}_t + B(\mathbf{u}_t, t, \boldsymbol{\theta}))\, dt + \boldsymbol{\sigma}(\mathbf{u}_t, t, \boldsymbol{\theta}) d\boldsymbol{\omega}_t \tag{46}$$

$$\mathbf{Y}_k = \mathbf{C}(\mathbf{u}_k, t_k, \boldsymbol{\theta})\mathbf{X}_k + \mathbf{e}_k \tag{47}$$

in the following we will use $\mathbf{A}(t), \mathbf{B}(t)$, and $\boldsymbol{\sigma}(t)$ as short-hand notation for $\mathbf{A}(\mathbf{u}_t, t, \boldsymbol{\theta})$, $B(\mathbf{u}_t, t, \boldsymbol{\theta})$, and $\boldsymbol{\sigma}(\mathbf{u}_t, t, \boldsymbol{\theta})$.

We will restrict ourselves to the initial value problem; solve (46) for $t \in [t_k, t_{k+1}]$ given that the initial condition $X_{t_k} \sim N(\hat{\mathbf{x}}_{k|k}, \Sigma_{k|k}^{xx})$. This is the kind of solution we would get from the ordinary Kalman filter in the update step.

Now if we consider, the transformation

$$\mathbf{Z}_t = e^{-\int_{t_k}^t \mathbf{A}(s)ds}\mathbf{X}_t, \tag{48}$$

then by Itô's Lemma, it can be shown that the process $\mathbf{Z}_t$ is governed by the Itô stochastic differential equation

$$d\mathbf{Z}_t = e^{-\int_{t_k}^t \mathbf{A}(s)ds}\mathbf{B}(t)dt + e^{-\int_{t_k}^t \mathbf{A}(s)ds}\boldsymbol{\sigma}(t)d\boldsymbol{\omega}_t \tag{49}$$

with initial conditions $\mathbf{Z}_{t_k} \sim N(\hat{\mathbf{x}}_{k|k}, \Sigma_{k|k}^{xx})$. The solution to (49) is given by the integral equation

$$\mathbf{Z}_t = \mathbf{Z}_{t_k} + \int_{t_k}^t e^{-\int_{t_k}^u \mathbf{A}(u)du}\mathbf{B}(s)ds + \int_{t_k}^t e^{-\int_{t_k}^s \mathbf{A}(u)du}\boldsymbol{\sigma}(s)d\boldsymbol{\omega}_s \tag{50}$$

Now inserting the inverse of the transformation (48) gives

$$\begin{aligned}\mathbf{X}_t = e^{\int_{t_k}^t \mathbf{A}(s)ds}\mathbf{X}_0 + e^{\int_{t_k}^t \mathbf{A}(s)ds}\int_{t_k}^t e^{-\int_{t_k}^u \mathbf{A}(u)du}\mathbf{B}(s)ds \\ + e^{\int_{t_k}^t \mathbf{A}(s)ds}\int_{t_k}^t e^{-\int_{t_k}^s \mathbf{A}(u)du}\boldsymbol{\sigma}(s)d\boldsymbol{\omega}_s\end{aligned} \tag{51}$$

Taking the exception and variance on both sides of (51) gives

$$E[\mathbf{X}_t] = e^{\int_{t_k}^{t} \mathbf{A}(s)ds} E[\mathbf{X}_{t_k}] + e^{-\int_{t_k}^{t} \mathbf{A}(s)ds} \int_{t_k}^{t} e^{-\int_{t_k}^{u} \mathbf{A}(u)du} \mathbf{B}(s)ds \tag{52}$$

$$\begin{aligned} V[\mathbf{X}_t] &= e^{\int_{t_k}^{t} \mathbf{A}(s)ds} V[\mathbf{X}_{t_k}] e^{\int_{t_k}^{t} \mathbf{A}(s)^T ds} + e^{\int_{t_k}^{t} \mathbf{A}(s)ds} V\left[\int_{t_k}^{t} e^{-\int_{t_k}^{s} \mathbf{A}(u)du} \boldsymbol{\sigma}(s) d\boldsymbol{\omega}_s\right] e^{\int_{t_k}^{t} A(s)^T ds} \\ &= e^{\int_{t_k}^{t} \mathbf{A}(s)ds} V[\mathbf{X}_0] e^{\int_{t_k}^{t} \mathbf{A}(s)^T ds} \\ &\quad + e^{\int_{t_k}^{t} \mathbf{A}(s)ds} \int_{t_k}^{t} e^{-\int_{t_k}^{s} \mathbf{A}(u)du} \boldsymbol{\sigma}(s)\boldsymbol{\sigma}(s)^T e^{-\int_{t_k}^{s} \mathbf{A}(u)^T du} ds e^{\int_{t_k}^{t} \mathbf{A}^T(s)ds}, \end{aligned} \tag{53}$$

where we have used Itô isometry in the second equation for the variance. Now differentiation the above expression w.r.t. time gives

$$\frac{dE[\mathbf{X}_t]}{dt} = \mathbf{A}(t)E[\mathbf{X}_t] + \mathbf{B}(t) \tag{54}$$

$$\frac{dV[\mathbf{X}_t]}{dt} = \mathbf{A}(t)V[\mathbf{X}_t] + V[\mathbf{X}_t]\mathbf{A}(t)^T + \boldsymbol{\sigma}(t)\boldsymbol{\sigma}(t)^T, \tag{55}$$

with initial conditions given by $E[\mathbf{X}_{t_k}] = \hat{\mathbf{x}}_{k|k}$ and $V[\mathbf{X}_{t_k}] = \Sigma^{xx}_{k|k}$.

For the nonlinear case

$$d\mathbf{X}_t = \mathbf{f}(\mathbf{X}_t, \mathbf{u}_t, t, \boldsymbol{\theta})dt + \boldsymbol{\sigma}(\mathbf{u}_t, t, \boldsymbol{\theta})d\boldsymbol{\omega}_t \tag{56}$$

$$\mathbf{Y}_k = \mathbf{h}(\mathbf{X}_k, \mathbf{u}_k, t_k, \boldsymbol{\theta}) + \mathbf{e}_k, \tag{57}$$

we introduce the Jacobian of $\mathbf{f}$ around the expectation of $\mathbf{X}_t$ ($\hat{\mathbf{x}}_t = E[\mathbf{X}_t]$), we will use the following short hand notation

$$\mathbf{A}(t) = \left.\frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}_t, t, \boldsymbol{\theta})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}_{t|k}}, \quad \mathbf{f}(t) = \mathbf{f}(\hat{\mathbf{x}}_{t|k}, \mathbf{u}_t, t, \boldsymbol{\theta}) \tag{58}$$

where $\hat{\mathbf{x}}_t$ is the expectation of $\mathbf{X}_t$ at time t, this implies that we can write the first-order Taylor expansion of (56) as

$$d\mathbf{X}_t \approx \left[\mathbf{f}(t) + \mathbf{A}(t)(\mathbf{X}_t - \hat{\mathbf{x}}_{t|k})\right]dt + \boldsymbol{\sigma}(t)d\boldsymbol{\omega}_t. \tag{59}$$

Using the results from the linear time-varying system above, we get the following approximate solution to the (59)

$$\frac{dE[\mathbf{X}_t]}{dt} \approx \mathbf{f}(t) \tag{60}$$

$$\frac{dV[\mathbf{X}_t]}{dt} \approx \mathbf{A}(t)V[\mathbf{X}_t] + V[\mathbf{X}_t]\mathbf{A}^T(t) + \boldsymbol{\sigma}(t)\boldsymbol{\sigma}^T(t), \tag{61}$$

with initial conditions $E[\mathbf{X}_{t_k}] = \hat{\mathbf{x}}_{k|k}$ and $V[\mathbf{X}_{t_k}] = \Sigma^{xx}_{k|k}$. Equations (60) and (61) constitute the basis of the prediction step in the Extended Kalman Filter, which for completeness is given below

Theorem 1 (Continuous-discrete time extended Kalman filter) *With given initial conditions for the* $\hat{\mathbf{x}}_{1|0} = \mathbf{x}_0$ *and* $\Sigma^{xx}_{1|0} = \Sigma^{xx}_0$ *the extended Kalman filter approximations are given by; the output* prediction *equations:*

$$\hat{\mathbf{y}}_{k|k-1} = \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}, \mathbf{u}_k, t_k, \boldsymbol{\theta}); \quad \Sigma^{yy}_{k|k-1} = \mathbf{C}_k \Sigma^{xx}_{k|k-1} \mathbf{C}_k^T + \mathbf{S}_k \tag{62}$$

the innovation *and Kalman* gain *equation:*

$$\boldsymbol{\varepsilon}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}; \quad \mathbf{K}_k = \Sigma^{xx}_{k|k-1} \mathbf{C}^T \left(\Sigma^{yy}_{k|k-1} \right)^{-1} \tag{63}$$

the updating *equations:*

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \boldsymbol{\varepsilon}_k; \quad \Sigma^{xx}_{k|k} = \Sigma^{xx}_{k|k-1} - \mathbf{K}_k \Sigma^{yy}_{k|k-1} \mathbf{K}_k^T \tag{64}$$

and the state prediction *equations:*

$$\frac{d\hat{\mathbf{x}}_{t|k}}{dt} = \mathbf{f}(\hat{\mathbf{x}}_{t|k}, \mathbf{u}_t, t, \boldsymbol{\theta}) \, , \, t \in [t_k, t_{k+1}[\tag{65}$$

$$\frac{d\Sigma^{xx}_{t|t_k}}{dt} = \mathbf{A}(t)\Sigma^{xx}_{t|t_k} + \Sigma^{xx}_{t|t_k}\mathbf{A}(t)^T + \boldsymbol{\sigma}(t)\boldsymbol{\sigma}(t)^T \, , \, t \in [t_k, t_{k+1}[\tag{66}$$

where the following short-hand notation has been applied:

$$\mathbf{A}(t) = \left. \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u}_t, t, \boldsymbol{\theta})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_{t|k-1}}, \quad \mathbf{C}_k = \left. \frac{\partial \mathbf{h}(\mathbf{x}, \mathbf{u}_{t_k}, t_k, \boldsymbol{\theta})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k-1}} \tag{67}$$

$$\boldsymbol{\sigma}(t) = \boldsymbol{\sigma}(\mathbf{u}_t, t, \boldsymbol{\theta}) \, , \quad \mathbf{S}_k = \mathbf{S}(\mathbf{u}_k, t_k, \boldsymbol{\theta}) \tag{68}$$

The prediction step was covered above and the updating step can be derived from linearization of the observation equation and the projection theorem [6]. From the construction above, it is clear that the approximation is only likely to hold if the nonlinearities are not too strong. This implies that the sampling frequency is fast enough for the prediction equations to be a good approximation and that the accuracy in the observation equation is good enough for the Gaussian approximation to hold approximately. Even though "simulation" through the prediction equations is available in CTSM-R, it is recommended that simulation results are verified (or indeed performed), by real-stochastic simulations (e.g., by simple Euler simulations).

References

1. Burnham, K., Anderson, D.: Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer (2002)
2. CTSM-R (Continuous Time Stochastic Modelling in R). www.ctsm.info
3. DIACON Project. www.diacongroup.org. New Technologies for treatment of Type 1 diabetes
4. Donnet, S., Samson, A.: A review on estimation of stochastic differential equations for pharmacokinetic/pharmacodynamic models. In: Advanced Drug Delivery Reviews (2013). doi:10.1016/j.addr.2013.03.005. http://www.sciencedirect.com/science/article/pii/S0169409X13000501
5. Duun-Henriksen, A., Juhl, R., Schmidt, S., Nørgaard, K., Madsen, H.: Modelling the effect of exercise on insulin pharmacokinetics in "continuous subcutaneous insulin infusion" treated type 1 diabetes patients. Technical report DTU Compute-Technical Report-2013, Technical University of Denmark (2013)
6. Jazwinski, A.H.: Stochastic processes and flitering theory. Dover publications, Inc. (1970)
7. Jellife, R., Schumitzky, A., Van Guilder, M.: Population pharmacokinetics/pharmacodynamics modeling; parametric and nonparametric methods. Ther. Drug Monit. **22**, 354–365 (2000)
8. Karlsson, M., Beal, S., Sheiner, L.: Three new residual error models for population pk/pd analysis. J. Pharmacokinet. Pharmacodyn. **23**, 651–672 (1995)
9. Klim, S., Mortensen, S.B., Kristensen, N.R., Overgaard, R.V., Madsen, H.: Population stochastic modelling (PSM)-an R package for mixed-effects models based on stochastic differential equations. Comput. Methods Progr. Biomed. **94**(3), 279–289 (2009). doi:10.1016/j.cmpb.2009.02.001. http://www.sciencedirect.com/science/article/pii/S0169260709000455
10. Kristensen, N.R., Madsen, H.: Continuous time stochastic modelling–CTSM 2.3 Mathamatics guide. Technical University of Denmark, DTU Informatics, Building 321 (2003). www.ctsm.info
11. Kristensen, N.R., Madsen, H., Jørgensen, S.B.: A method for systematic improvement of stochastic gray-box models. Comput. Chem. Eng. **116**, 1431–1449 (2004)
12. Kristensen, N.R., Madsen, H., Jørgensen, S.B.: Parameter estimation in stochastic grey-box models. Automatica **40**(2), 225–237 (2004)
13. Lindsey, J., Jones, B., Jarvis, P.: Some statistical issues in modelling pharmacokinetic data. Stat. Med. **20**, 2775–2783 (2001)
14. Löwe, R., Mikkelsen, P., Madsen, H.: Stochastic rainfall-runoff forecasting: parameter estimation, multi-step prediction, and evaluation of overflow risk. Stoch. Environ. Res. Risk Assess. 28(3), 505-516 (2014). doi:10.1007/s00477-013-0768-0. (Offprint, no public access)
15. Lv, D., Breton, M.D., Farhy, L.S.: Pharmacokinetics modeling of exogenous glucagon in Type 1 diabetes mellitus patients. Diabet. Technol. Ther. **15**(11), 935–941 (2013). doi:10.1089/dia.2013.0150. http://online.liebertpub.com.globalproxy.cvt.dk/doi/abs/10.1089/dia.2013.0150
16. Madsen, H.: Time Series Analysis. Chapman and Hall (2008)
17. Madsen, H., Thyregod, P.: Introduction to general and generalized linear models. Chapman and Hall (2011)
18. Møller, J., Madsen, H.: Stochastic state space modelling of nonlinear systems–with application to marine ecosystems. In: IMM-PHD-2010-246. Technical University of Denmark, DTU Informatics, Building 321 (2010)
19. Møller, J., Phillipsen, K.R., Christensen, L.E., Madsen, H.: Development of a restricted state space stochastic differential equation model for bacterial growth in rich media. J. Theor. Biol. **305**, 78–87 (2012). doi:10.1016/j.jtbi.2012.04.015
20. Nielsen, H.A., Madsen, H.: A generalization of some classical time series tools. Comput. Stat. Data Anal. **37**, 13–31 (2001)
21. Pawitan, Y.: In all likelihood: Statistical modelling and inference using likelihood. Oxford Science Publications (2001)
22. Philipsen, K.R., Christiansen, L.E., Hasman, H., Madsen, H.: Modelling conjugation with stochastic differential equations. J. Theor. Biol. **263**(1), 134–142 (2010). doi:10.1016/j.jtbi.2009.11.011

23. Schmidt, S., Finan, D.A., Duun-Henriksen, A.K., Jørgensen, J.B., Madsen, H., Bengtsson, H., Holst, J.J., Madsbad, S., Nørgaard, K.: Effects of everyday life events on glucose, insulin, and glucagon dynamics in continuous subcutaneous insulin infusion treated type 1 diabetes: Collection of clinical data for glucose modeling. Diabet. Technol. Ther. **4**(3), 210–217 (2012). doi:10.1089/dia.2011.0101. http://online.liebertpub.com/doi/abs/10.1089/dia.2011.0101
24. Tornøe, C.W., Agersø, H., Jonsson, E.N., Madsen, H., Nielsen, H.A.: Non-linear mixed-effects pharmacokinetic/pharmacodynamic modelling in nlme using differential equations. Computer Methods and Programs in Biomedicine. Comput. Methods Progr. Biomed. **76**(1), 31–40 (2004). doi:10.1016/j.cmpb.2004.01.001
25. Tornøe, C.W., Jacobsen, J., Pedersen, O., Hansen, T., Madsen, H.: Grey-box modelling of pharmacokinetic/pharmacodynamic systems. J. Pharmacokinet. Pharmacodyn. **31**(5), 401–417 (2004)
26. Wilinska, M.E., Chassin, L.J., Acerini, C.L., Allen, J.M., Dunger, D.B., Hovorka, R.: Simulation environment to evaluate closed-loop insulin delivery systems in type 1 diabetes. J. Diabet. Sci. Technol. **4**(1), 132–144 (2010)

Uncertainties and Modeling Errors of Type 1 Diabetes Models

Levente Kovács and Péter Szalay

Abstract Modeling and control are tightly connected if we want to guarantee safety and reliability. These are minimum requirements in the medical field. The more sophisticated methods usually require information beyond the available measurements, and one way or another incorporate all a priori knowledge. This can manifest in state estimation, model-based prediction, or robust design assuming the worst case, among others. The better the model the better the achievable control; however, all aspects of modeling are more difficult in the case of physiological systems compared to regular engineering applications. In the following, we will investigate how various errors resulting from modeling inaccuracies affect the prediction of the behavior in case of blood glucose prediction. Sigma-point filters are used to efficiently support Kalman filtering, while the error sources are introduced in a single uncertainty block.

1 Introduction

Diabetes mellitus is the dysfunction of the human glucose regulation that is currently incurable, but treatable. Normally, the concentration of plasma glucose is kept in a narrow range (3.9–7.8 mmol/L or 70–110 mg/dL) by a complex endocrine system, where insulin plays a key role in this process. Type 1 (insulin dependent), Type 2 (non-insulin dependent), gestational, and special types like genetic deflections are the types of diabetes. Classical therapy of type 1 diabetes mellitus (T1DM) consists of insulin injections administered by the patient. However, even the most cooperative patient can be subjected to the chronic complications, such as neuropathy, nephropathy, or retinopathy, among others [14]. Existing therapies could greatly benefit from an

L. Kovács (✉)
Physiological Controls Group, Applied Informatics Institute, John von Neumann Faculty of Informatics, Obuda University, Budapest, Hungary
e-mail: kovacs.levente@nik.uni-obuda.hu

P. Szalay
Department of Control Engineering and Information Technology, Budapest University of Technology and Economics, Budapest, Hungary

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*, Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_11

accurate prediction of the plasma glucose concentration, as well as optimized dosage of insulin based on these predictions.

Complete automation of the treatment is researched in the literature (for T1DM) under the artificial pancreas (AP) problem. AP consists on three components: a continuous glucose monitoring (CGM) sensor for the subcutaneous measurement of glucose concentrations, an insulin pump for the subcutaneous delivery of insulin, and a control algorithm that based on CGM measurements is able to determine the necessary insulin dosage to be injected by the insulin pumps [7, 8].

There are various control methods already developed: classical PID [25]; run-to-run control [39]; exact linearization-based nonlinear control [26]; $\mathscr{H}_\infty$ control [10, 13, 27, 29]; model predictive control (MPC) [16, 22, 24]; linear parameter varying (LPV)-based robust control [9, 21]. Soft computing-based methods, such as fuzzy logic control [28] and model-free soft computing-based control [36] are gaining popularity as well. Most of these techniques require signals beyond what is physically measurable; hence, accurate estimation of the state variables and precise modeling is needed.

In this chapter, various error sources related to modeling will be investigated using the widely used T1DM model of the literature [16]. Section 2 focuses on the general aspects of modeling the human metabolism, briefly presenting the used model for our investigations. Model reduction is examined in Sect. 3. State estimation in terms of novel versions of Kalman filtering is detailed in Sect. 4, while the modeling errors resulting from estimation is investigated in Sect. 5. Section 6 concludes the paper. Simulations have been performed using the in silico simulator of the University of Cambridge version 2.2 (SimEdu) [35].

2 Modeling Diabetes

Various models appeared in the literature to describe the normal or impaired human metabolism. One of the earliest and simplest can capture the dynamics using merely three state variables [4]. On the other hand, one of the most sophisticated models [31]—despite being potentially highly accurate—contains way too many states and parameters to be useful in the clinical practice. Furthermore, the representation of subcutaneous compartments is usually neglected in the early models. However, the ones presented in [24, 35] corrected these shortcomings by balancing complexity and manageable size and including the subcutaneous route for both sensor and actuator. Even if we consider ordinary differential equations only, the available models which are successfully used for various purposes are beyond count; but, they do share certain similarities: they contain some form of nonlinearity, most frequently the multiplication of two state variables, Michaelis–Menten functions, and saturation. In general, transfer rate between two states (representing compartments) are a function of other state variables. Most of the state variables cannot have negative value. There are inputs which cannot be measured and only vaguely represented, such as stress or

physical activity. Finally, the model parameters are assumed to be changing in time. In the followings, these models will be referred to as T1DM models.

The model investigated in this paper is described by the following differential equations [35]:

$$
\begin{aligned}
\dot{C}(t) &= -k_{a,int}C(t) + \tfrac{k_{a,int}}{V_G}Q_1(t) \\
\dot{Q}_1(t) &= -\left(\tfrac{F_{01}}{Q_1(t)+V_G} + x_1(t)\right)Q_1(t) + k_{12}Q_2(t) \\
&\quad -R_{cl}max\{0, Q_1(t) - R_{thr}V_G\} - Phy(t) \\
&\quad +EGP_0max\{0, 1 - x_3(t)\} + min\left\{U_{G,ceil}, \tfrac{G_2(t)}{t_{max}}\right\} \\
\dot{Q}_2(t) &= x_1(t)Q_1(t) - \Big(k_{12} + x_2(t)\Big)Q_2(t) \\
\dot{x}_1(t) &= -k_{b1}x_1(t) + S_{IT}k_{b1}I(t) \\
\dot{x}_2(t) &= -k_{b2}x_2(t) + S_{ID}k_{b2}I(t) \\
\dot{x}_3(t) &= -k_{b3}x_3(t) + S_{IE}k_{b3}I(t) \\
\dot{I}(t) &= \tfrac{k_a}{V_I}S_2(t) - k_eI(t) \\
\dot{S}_2(t) &= -k_aS_2(t) + k_aS_1(t) \\
\dot{S}_1(t) &= -k_aS_1(t) + u(t) \\
\dot{G}_2(t) &= \tfrac{G_1(t)-G_2(t)}{max\left\{t_{max}, \frac{G_2(t)}{U_{G,ceil}}\right\}} \\
\dot{G}_1(t) &= -\tfrac{G_1(t)}{max\left\{t_{max}, \frac{G_2(t)}{U_{G,ceil}}\right\}} + D(t)
\end{aligned}
\tag{1}
$$

where the state variables are: $C(t)$ glucose concentration in the subcutaneous tissue [mmol/L], $Q_1(t)$ and $Q_2(t)$ the masses of glucose in accessible and nonaccessible compartments [mmol], $x_1(t)$, $x_2(t)$ and $x_3(t)$ remote effect of insulin on glucose distribution, disposal, and endogenous glucose production, respectively, [1/min], $I(t)$ insulin concentration in plasma [mU/L], $S_1(t)$, and $S_2(t)$ insulin masses in the accessible and nonaccessible compartments [mU], $G_1(t)$, and $G_2(t)$ glucose masses in the accessible and nonaccessible compartments [mmol]. $u(t)$-injected insulin flow of rapid-acting insulin [mU/min] is the input of the system, while $D(t)$ amount of ingested carbohydrates [mmol/min], and $Phy(t)$ effect of physical activity [mmol/min] are considered as disturbances.

The parameters of the model are $k_{a,int}$ transfer rate constant between the plasma and the subcutaneous compartment [1/min], V_G distribution volume of glucose in the accessible compartment [L], F_{01} parameter of the total non-insulin dependent glucose flux [mmol/min], k_{12} transfer rate constant from the nonaccessible to the accessible compartment [1/min], R_{cl} renal clearance constant [1/min], R_{thr} glucose threshold [mmol/L], EGP_0 endogenous glucose production extrapolated to the zero insulin concentration [mmol/min], t_{max} time-to-maximum appearance rate of glucose in the accessible compartment [min], $U_{G,ceil}$ maximum glucose flux from the gut [mmol/kg/min], k_{b1} and k_{b2} deactivation rate constants [$\frac{min^{-2}}{mU/L}$], k_{b3} deactivation rate constant for the insulin effect on endogenous glucose production [$\frac{min^{-1}}{mU/L}$], S_{IT},

S_{ID} and S_{IE} insulin sensitivities for transport, distribution, and endogenous glucose production $[\frac{min^{-1}}{mU/L}]$ and $[\frac{1}{mU/L}]$, k_a insulin absorption rate constant [1/min], V_I volume of distribution of rapid-acting insulin [L], k_e fractional elimination rate from plasma [1/min]. The following parameters are assumed to be time-varying with ±5 % deviation: $k_{a,int}$, F_{01}, k_{12}, EGP_0, k_{b1}, k_{b2}, k_{b3}, S_{IT}, S_{ID}, S_{IE}, k_a, and k_e.

This can be extended with a continuous glucose monitor (CGM) model, with $C(t)$ as input and measured glucose concentration as output. There are various CGM models to choose from [5, 6, 12]. The sampling time associated with sensor is usually around 5 min.

2.1 Linear Parameter Varying Model

There are several ways to handle the nonlinearity of the model. The classical nonlinear methodology focuses on the differential geometric approach [17], while a more recent methodology is represented by linear parameter varying (LPV) systems [3, 23]. LPV is an acceptable compromise between the model's complexity and the developed control algorithm, as LPV systems can be seen as an extension of linear time invariant (LTI) systems, where the relations are considered to be linear, but model parameters are assumed to be functions of a time-varying signal.

Most T1DM models can be represented with a linear parameter varying (LPV) model [21, 32]. These models are essentially an extension of linear models, and have the following form:

$$\begin{aligned}
\dot{\mathrm{x}}(t) &= \mathbf{A}(\rho(t))\mathrm{x}(t) + \mathbf{B}(\rho(t))\mathrm{u}(t) \\
\mathrm{y}(t) &= \mathbf{C}(\rho(t))\mathrm{x}(t) + \mathbf{D}(\rho(t))\mathrm{u}(t) \\
\mathbf{A}(\rho(t)) &= \mathbf{A}_0 + \prod_{i=1}^{m} \rho_i(t)\mathbf{A}_i\mathbf{B}(\rho(t)) = \mathbf{B}_0 + \prod_{i=1}^{m} \rho_i(t)\mathbf{B}_i \\
\mathbf{C}(\rho(t)) &= \mathbf{C}_0 + \prod_{i=1}^{m} \rho_i(t)\mathbf{C}_i\mathbf{D}(\rho(t)) = \mathbf{D}_0 + \prod_{i=1}^{m} \rho_i(t)\mathbf{D}_i
\end{aligned} \tag{2}$$

where the scheduling variables $(\rho(t))$ are assumed to be bounded, as well as their first time derivatives. Although different configurations are possible [32], the following scheduling parameters have been chosen for model (1):

$$\rho(t) = \begin{pmatrix} \rho_1(t) \\ \rho_2(t) \\ \rho_3(t) \end{pmatrix} = \begin{pmatrix} Q_1(t) \\ F_{01}(Q_1(t) + V_G)^{-1} \\ Q_2(t) \end{pmatrix} \tag{3}$$

3 Model Reduction

As model (1) is an 11th order model, it is rather difficult to handle for anything beyond simulations. A decreased number of state variables would be more practical for tasks with memory and computation power requirements rising exponentially with the model order. Based on the parameter bounds presented in [16, 35] the speed of transfer between certain compartments is comparable to the sampling time of the CGM. Hence, the states associated with them can be eliminated. The resulting reduced model is as follows:

$$\begin{aligned}
\dot{Q}_1(t) &= -\left(\tfrac{F_{01}}{Q_1(t)+V_G} + x_1(t)\right) Q_1(t) + k_{12} Q_2(t) \\
&\quad -R_{cl} max\{0, Q_1(t) - R_{thr} V_G\} - Phy(t) \\
&\quad +EGP_0 max\left\{0, 1 - \tfrac{k_a S_{IE}}{V_I k_e} S_2(t)\right\} + min\left\{U_{G,ceil}, \tfrac{G_2(t)}{t_{max}}\right\} \\
\dot{Q}_2(t) &= x_1(t) Q_1(t) - \left(k_{12} + \tfrac{k_a S_{ID}}{V_I k_e} S_2(t)\right) Q_2(t) \\
\dot{x}_1(t) &= k_{b1}\left(\tfrac{k_a S_{IT}}{V_I k_e} S_2(t) - x_1(t)\right) \\
\dot{S}_2(t) &= -k_a S_2(t) + k_a S_1(t) \\
\dot{S}_1(t) &= -k_a S_1(t) + u(t) \\
\dot{G}_2(t) &= \tfrac{G_1(t) - G_2(t)}{max\left\{t_{max}, \frac{G_2(t)}{U_{G,ceil}}\right\}} \\
\dot{G}_1(t) &= -\tfrac{G_1(t)}{max\left\{t_{max}, \frac{G_2(t)}{U_{G,ceil}}\right\}} + D(t)
\end{aligned} \tag{4}$$

where the output is $C(t) \approx Q_1(t)/V_G$.

The equations for state variables $S_1(t)$, $S_2(t)$ and $x_1(t)$ can be regarded as a third-order linear system with injected insulin $u(t)$ as the only input and three outputs: $x_1(t)$, $x_2(t) \approx (k_a S_{ID} S_2(t))/(V_I k_e)$ and $x_3(t) \approx (k_a S_{IE} S_2(t))/(V_I k_e)$. Graphical representation of state elimination is presented in Fig. 1.

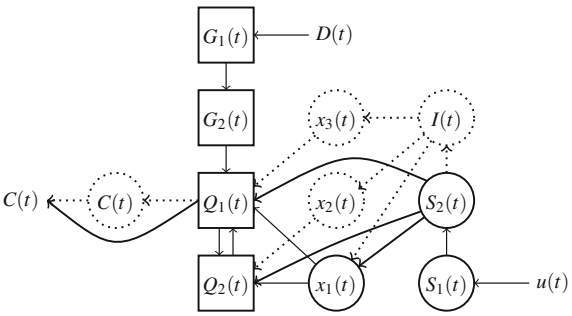

Fig. 1 Eliminating state variables. *Circle* nodes represent linear, *square* nodes represent nonlinear equations. Disturbance input $Phy(t)$ is not displayed

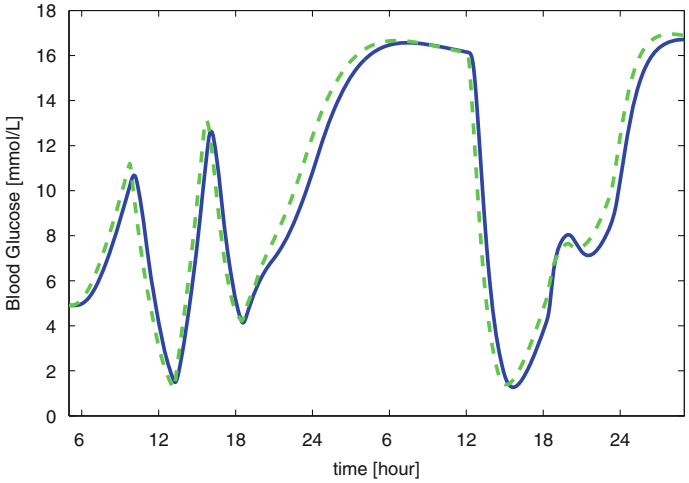

Fig. 2 Output of the original (*solid line*) and reduced system (*dashed line*)

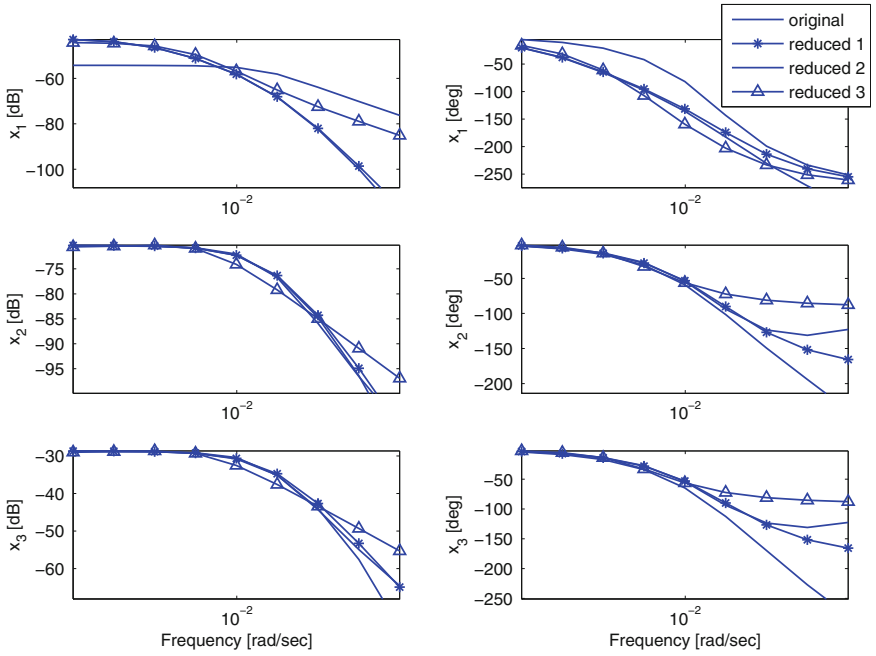

Fig. 3 Bode diagram for different types of reduction. *Reduced 1* represents the effect of state elimination only (4). *Reduced 2* displays the Bode diagram in case the subsystem is reduced to a 2nd-order one, but no weighting is used, unlike in case of *reduced 3* where weighting has been also applied

Further reduction of this third-order system is possible using frequency-weighted balanced reduction presented in [11] or [37]. If weighting is needed, one might use the nonlinear subsystem of the model—described by the equations belonging to states $Q_1(t)$ and $Q_2(t)$—linearized at a chosen working point. However, when the system is used in other working points, this may not be the most effective choice. In Fig. 2, the outputs of the original 11th-order model (1) is compared with the reduced 7th-order model (4), where four state variables were neglected. The insulin, meal, and physical activity inputs are randomly generated [32], assuming time-invariant parameters.

Figure 3 presents the Bode diagrams of the above-mentioned third-order subsystem for all three outputs in the case of different types of reduction.

The parameters changing in time can alter the dynamics of the model quite significantly, which can lead to errors in prediction as well as instability in closed-loop control. The $\pm 5\,\%$ deviation can be considered a rather optimistic assumption. Figure 4 shows the comparison of the output of the time-invariant model with an example in the case of time-varying parameters, as well as the envelope for the maximum deviation from the nominal behavior during simulation. Note, that not even these figures represent the theoretically possible highest deviation, for randomized simulations cannot guarantee a deterministic worst case scenario [33].

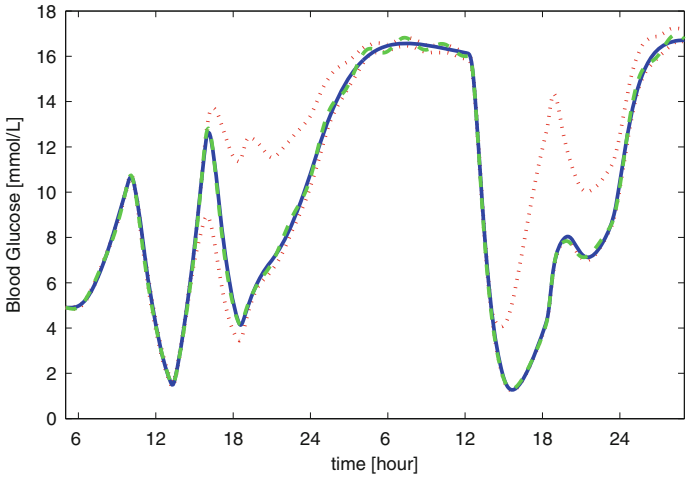

Fig. 4 Time-varying parameters. The *solid line* is the output of the nominal model, the *dashed line* is an example output in case of changing parameters, while the bounds marked with *dotted lines* are the results of numerous simulations with randomized settings

4 State Estimation

Prediction usually needs information about all state variables of the system, as well as some of the more advanced control techniques. Hence, since they are not available for measurement, reliable estimation is needed. Kalman filter is a popular choice for such tasks. They are especially useful in the case of prediction, since Kalman filters provide not only the estimated values of the states, but the variance of the estimation error as well. When the CGM has significant drift, this information can be used for recalibration. Let us approximate the combination of T1DM and sensor models with the following discrete-time nonlinear system:

$$\begin{aligned} \mathrm{x}_{k+1} &= \mathrm{f}(\mathrm{x}_k, \mathrm{u}_k, \mathrm{w}_k) \\ y_k &= h(\mathrm{x}_k) + n_k \end{aligned} \tag{5}$$

where $\mathrm{x}_k \in \mathbb{R}^{n_x}$ is the vector of state variables, $y_k \in \mathbb{R}$ denotes the measured output disturbed with $n_k \sim \mathcal{N}(0, R_k)$ additive white noise. $\mathrm{u}_k \in \mathbb{R}^{n_u}$ is the vector of known deterministic inputs, while $\mathrm{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k)$ is the vector of disturbances affecting the states, with assumingly zero mean Gaussian distribution and $n_w \times n_w$ real positive semidefinite covariance matrix $\mathbf{Q}_k$. $\mathrm{f} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \to \mathbb{R}^{n_x}$ and $h : \mathbb{R}^{n_x} \to \mathbb{R}$ are piecewise continuous nonlinear mappings. The errors resulting from the assumptions made about the disturbances and measurement noise are not investigated in this paper.

Once the new model (5) is available, it is possible to use Kalman filter for state estimation or prediction. Extended Kalman filter is the usual choice in the case of nonlinear systems. However, the precision of this filter is only satisfactory in case of mild nonlinearity and disturbances, since it is based on first-order linearization [15].

A more effective alternative are the use of sigma-point filters. They use a number of deterministic samples, called sigma points, to represent the probability distribution of the system state needed in the Kalman filter algorithm [19]. There are several versions which differ mainly on how these sigma points are selected. Cubature Kalman filter (CKF) is based on the cubature rule [1] and is one of the most straightforward approaches. Unscented Kalman Filter (UKF) relies on the unscented transformation with parameters that can be tuned for each filtering problem in order to achieve better performance [19]. Gauss–Hermite quadrature filter (GHQF) can be used if Gaussian distribution is guaranteed offering the highest accuracy [2]. However, it also requires a large amount of sigma points and hence increased computational power which can be undesirable in certain practical applications. As a result, our choice went on sparse-grid quadrature filtering method (SGQF) developed in order to overcome the dimensionality problem of GHQF [18]. Moreover, CKF and UKF can be considered as a special form of SGQF.

4.1 Sigma-Point Selection

Let us introduce the notation χ for a set of sigma points. This set contains N sigma points denoted as $\xi_i, i = 1, \ldots, N$. The sigma points represent the stochastic variable μ with $\hat{\mu}$ mean and $\boldsymbol{\Sigma}$ covariance matrix, and can be written in the following form:

$$\xi_i = \boldsymbol{\Sigma}^{\frac{1}{2}} \varphi_i + \hat{\mu} \tag{6}$$

$\Sigma^{\frac{1}{2}}$ is the factor of $\boldsymbol{\Sigma}$ so that $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\frac{1}{2}} (\Sigma^{\frac{1}{2}})^T$, and since $\boldsymbol{\Sigma}$ is positive definite, Cholesky decomposition is commonly used. In case $\boldsymbol{\Sigma}$ is close to being singular, or nondefinite due to sigma-point collapse [34], singular value decomposition can be used as well [18]. μ is not limited to state variables only, it contains the disturbances and measurement noises as well [30], so that:

$$\hat{\mu} = \begin{pmatrix} \hat{\mathrm{x}} \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & \mathbf{Q} & 0 \\ 0 & 0 & \mathbf{R} \end{bmatrix} \tag{7}$$

where $\mathbf{Q}$ and $\mathbf{R}$ denote covariance matrices of the disturbances and measurement noise, just like earlier. Using these sigma points, one can estimate the mean and covariance of the distribution of $f(\chi)$ as a weighted sum, where $f(.)$ is a nonlinear function:

$$\begin{aligned} \mathrm{E}\{f(\mu)\} = \bar{\mathrm{f}}_\mu &\approx \textstyle\sum_{i=1}^{N} \omega_i^{(m)} f(\xi_i) \\ cov\{f(\mu)\} &\approx \textstyle\sum_{i=1}^{N} \omega_i^{(c)} \left((f(\xi_i) - \bar{\mathrm{f}}_\mu)(f(\xi_i) - \bar{\mathrm{f}}_\mu)^T \right) \end{aligned} \tag{8}$$

There are various strategies to choose φ_i and the weights $\omega_i^{(m)}$ and $\omega_i^{(c)}$. In case of CKF, there are $2L$ sigma points, where L is the dimension of μ. The weights and basis functions $\omega^{(m)}$, $\omega^{(c)}$, and φ in the case of a CKF are:

$$\begin{aligned} \varphi_i &= \begin{cases} \mathrm{e}_i \sqrt{L} & i = 1, \ldots, L \\ -\mathrm{e}_i \sqrt{L} & i = L+1, \ldots, 2L \end{cases} \\ \omega_i^{(m)} &= \omega_i^{(c)} = \frac{1}{\sqrt{L}} \end{aligned} \tag{9}$$

where e_i denotes the unit vector in $\mathbb{R}^L$ with the $(i-1)$th element being 1. Note that CKF does not have any adjustable parameter, opposed to UKF which has three: κ, α, and β. The weights and basis functions ω and ϕ in the case of a CKF are:

$$\varphi_i = \begin{cases} 0 & i = 1 \\ \mathrm{e}_i\sqrt{L+\lambda} & i = 2,\ldots,L+1 \\ -\mathrm{e}_i\sqrt{L+\lambda} & i = L+2,\ldots,2L+1 \end{cases}$$
$$\omega_i^{(m)} = \begin{cases} \frac{\lambda}{n+\lambda} & i = 1 \\ \omega_i^{(m)} = \frac{1}{2(n+\lambda)} & i = 2,\ldots,2L+1 \end{cases} \tag{10}$$
$$\omega_1^{(c)} = \begin{cases} \frac{\lambda}{n+\lambda} + 1 - \alpha^2 + \beta & i = 1 \\ \omega_i^{(c)} = \frac{1}{2(n+\lambda)} & i = 2,\ldots,2L+1 \end{cases}$$

where $\lambda = \alpha^2(L+\kappa) - L$ is a scaling parameter [15]. The constant α determines the spread of sigma points around μ, and is usually set to a small positive value (e.g., $1 \geq \alpha \geq 10^{-4}$). The constant κ is a second scaling parameter usually set to $3 - L$ so that the kurtosis of the sigma points agrees with that of the Gaussian distribution [20]. β is used to incorporate prior knowledge of the distribution of μ and usually set to 2 for Gaussian distribution. In case $\alpha = 1$ and $\beta = 0$, both CKF and UKF can be seen as a special case of level-2 SGQF.

$$\varphi_i = \begin{cases} 0 & i = 1 \\ \mathrm{e}_i p_1 & i = 2,\ldots,L+1 \\ -\mathrm{e}_i p_1 & i = L+2,\ldots,2L+1 \\ \mathrm{e}_i p_2 & i = 2L+2,\ldots,3L+1 \\ -\mathrm{e}_i p_2 & i = 3L+2,\ldots,4L+1 \\ \mathrm{e}_i p_3 & i = 4L+2,\ldots,5L+1 \\ -\mathrm{e}_i p_3 & i = 5L+2,\ldots,6L+1 \\ \mathrm{e}_i p_1 + \mathrm{e}_j p_1, & i = 6L+2,\ldots,6L+1+C \quad j \neq i \\ -\mathrm{e}_i p_1 + \mathrm{e}_j p_1, & i = 6L+2+C,\ldots,6L+1+2C \quad j \neq i \\ \mathrm{e}_i p_1 - \mathrm{e}_j p_1, & i = 6L+2+2C,\ldots,6L+1+3C \quad j \neq i \\ -\mathrm{e}_i p_1 - \mathrm{e}_j p_1, & i = 6L+2+3C,\ldots,6L+1+4C \quad j \neq i \end{cases} \tag{11}$$

$$\omega_i = \begin{cases} \left(\frac{(L-2+L\hat{\omega}_1^2)}{2} - L\hat{\omega}_1\right)(L-1) + L\hat{\omega}_3 & i = 1 \\ (L-1)\hat{\omega}_2(\hat{\omega}_1 - 1) & i = 2,\ldots,2L+1 \\ \hat{\omega}_4 & i = 2L+2,\ldots,4L+1 \\ \hat{\omega}_5 & i = 4L+2,\ldots,6L+1 \\ \hat{\omega}_2^2 & i = 6L+2,\ldots,6L+1+4C \end{cases}$$

The level-3 SGQF requires $2L^2 + 4L + 1$ or less sigma points. The exact number depends on how the three free parameters—p_1, p_2, and p_3—are chosen. Similarly to the GHKF, these parameters are selected from the perspective of an univariate estimation, where the points $\mu + \{-p_1, 0, p_1\}$ and $\mu + \{-p_3, -p_2, 0, p_2, p_3\}$ are used to estimate certain moments of an univariate Gaussian distribution transformed by a nonlinear function. If all parameters are different, the sigma points used in the level-3 SGQF are shown in Eq. (11), where $C = L(L-1)/2$, while $\hat{\omega}_1, \ldots, \hat{\omega}_5$ are

defined from the parameters p_1, p_2, p_3 using moment matching method. Furthermore, $\omega_i^{(c)} = \omega_i^{(m)}$ and $j \neq i$.

5 Model Uncertainty

Control methods using state feedback, like the popular MPC [24, 35] or exact linearization-based control [26], require all state variables. LPV modeling (and control) is less strict requiring only that the scheduling parameters should be available for measurement. However, since these signals cannot be measured directly, the estimation error must be taken into consideration. Consequently, it can be regarded as additional parameter inaccuracy or a virtual disturbance. For model (1), a combination of the two approaches have been chosen. For the scheduling variable $\rho_i(t)$, we can assume that the estimated $\hat{\rho}_i(t)$ is the correct value, but additive $\Delta\rho_{i,0}$ and multiplicative $\Delta\rho_{i,1}$ error is present, as well as a disturbance $d_{\rho,i}(t)$ with Gaussian distribution and zero mean. Hence:

$$\rho_i(t) \approx \hat{\rho}_i(t)(1 + \Delta\rho_{i,1}) + \Delta\rho_{i,0} + d_{\rho,i}(t) \tag{12}$$

The values of $\Delta\rho_{i,0}$ and $\Delta\rho_{i,1}$ can be defined using least squares (LS) or weighted least squares (WLS) method based on a series of simulations with carefully chosen randomized inputs, initial conditions, and measurement noises. Once these two parameters are set, the worst case value of the variance of $d_{\rho,i}(t)$ can be defined (Fig. 5). Figure 6 displays how the additive and multiplicative error of the scheduling parameters can change the behavior of the nominal model in worst case scenario.

5.1 Error Weighting Function

Due to the high number of error sources and inaccuracies, estimating the worst case behavior of the system can be demanding for computations. However, it is possible to capture them in a single weighting function $W(s)$ [38] in an offline way (Fig. 7). Once this function is available, the online tasks of prediction and control can consider worst case deviation from the nominal model easier. An example is displayed in Fig. 8 obtained by simulations on the in silico simulator of the University of Cambridge version 2.2 (SimEdu) [35]. $W(s)$ incorporates parameters changing in time, reduction using state elimination and scheduling variable estimation error.

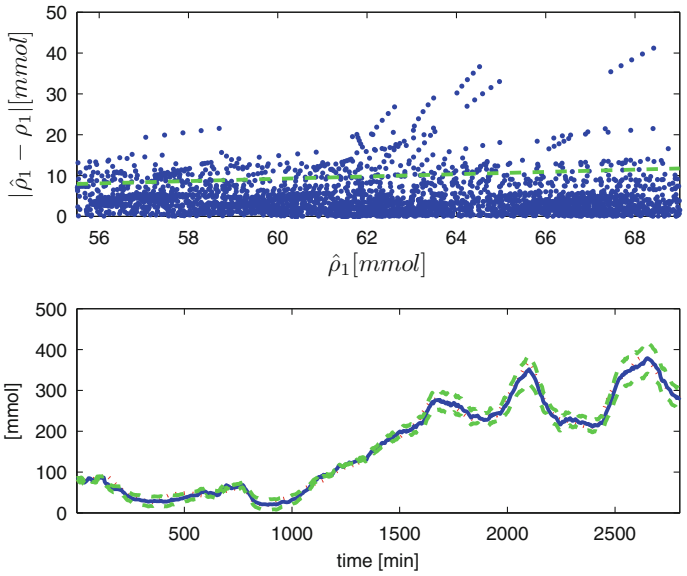

Fig. 5 Defining parameter uncertainty parameters for $\rho_1(t)$. In the *upper* subplot, the absolute difference between the estimated ($\hat{\rho}_1$) and the real (ρ_1) value of a scheduling variable is displayed as a function of the estimated value. The *dashed line* represents the chosen additive and multiplicative uncertainty. Everything that is not covered by the uncertainty is considered disturbance. The *lower* subplot shows the time function of a scheduling variable (*dotted line*), the estimation (*solid line*), and the uncertainty bounds (*dashed line*)

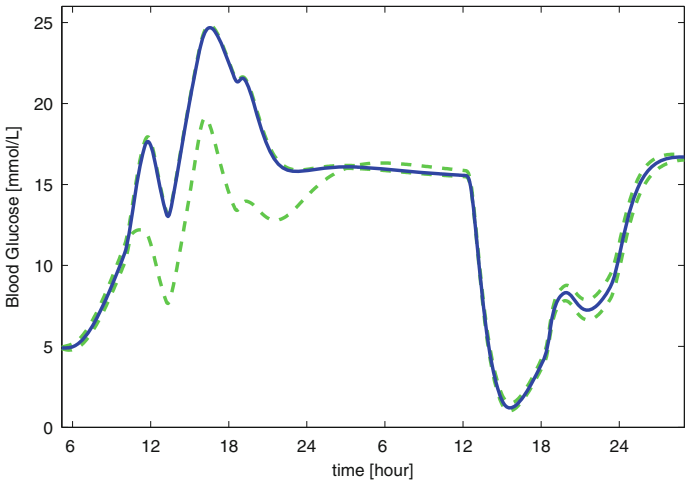

Fig. 6 Effects of scheduling variable estimation error. The *solid line* is the output of the nominal model; the *dashed lines* are the bounds for how much the response of the actual system can deviate

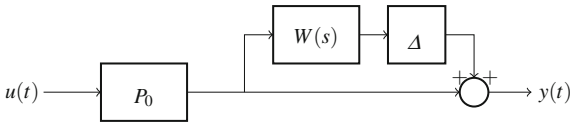

Fig. 7 Output multiplicative uncertainty. Δ represents an unstructured uncertainty block, which is an unknown linear system assumed to have $\mathscr{H}_\infty$ norm smaller than 1, stable and minimal phase

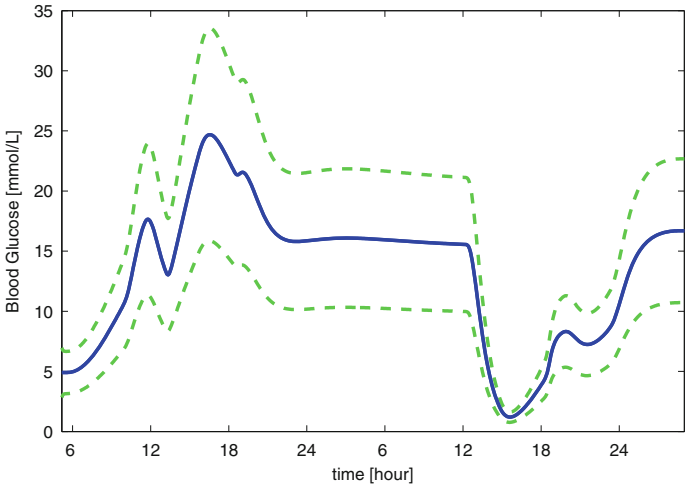

Fig. 8 Uncertainties and errors represented with a single weighting function. The *solid line* is the output of the nominal model, the *dashed lines* are the bounds for how much the response of the actual system can deviate

6 Conclusion

In this chapter, the effects of changing parameters, model reduction, and state estimation errors were examined on the widely used T1DM model of [16] to show the challenges they impose on model-based prediction and control. Sigma-point filters have been used as an extension of the nonlinear Extended Kalman Filtering technique, offering higher accuracy in this particular problem.

Further work will focus on parameter sensitivity of each sigma-point method, but a more sophisticated sensor model could be used as well. Filters requiring less sigma points should be included also in the comparison. The question of robustness has to be investigated as well, since it is safe to assume that the used T1DM model is inaccurate. Finally, additional safety measures must be taken in case the patient would be less cooperative, and occasionally neglect the requested sensor calibration or be simply late with the measurement.

Acknowledgments Levente Kovács is Bolyai Fellow of the Hungarian Academy of Sciences. The work is partially supported by the Hungarian National Development Agency GOP-1.1.1.-11-2012-0055 project and by the European Union TÁMOP-4.2.2.A-11/1/KONV-2012-0073 project.

References

1. Arasaratnam, I., Haykin, S.: Cubature kalman filters. IEEE Trans. Autom. Control **54**(6), 1254–1269 (2009)
2. Arasaratnam, I., Haykin, S., Elliot, R.: Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature. Proc. IEEE **95**, 953–977 (2007)
3. Becker, G.S., Packard, A.: Robust performance of LPV systems using parametrically dependent linear feedback. Syst. Control Lett. **23**, 205–215 (1994)
4. Bergman, B.N., Ider, Y.Z., Bowden, C.R., Cobelli, C.: Quantitative estimation of insulin sensitivity. Amer. J. Physiol. **236**, 667–677 (1979)
5. Breton, M., Kovatchev, B.: Analysis, modeling, and simulation of the accuracy of continuous glucose sensors. J. Diabetes Sci. Technol. **2**(5), 853–862 (2008)
6. Cengiz, E., Tamborlane, W.V.: A tale of two compartments: interstitial versus blood glucose monitoring. Diabetes Technol. Ther. **11**(1), S11–S16 (2009)
7. Chee, F., Fernando, T.: Closed–Loop Control of Blood Glucose, Springer LNCIS 386. Springer, New York (2007)
8. Cobelli, C., Renard, E., Kovatchev, B.: Artificial pancreas: past, present and future. Diabetes **60**(11), 2672–2682 (2011)
9. Colmegna, P., Sanchez-Pena, R.S.: LPV control to minimize risks in Type 1 diabetes. In: Proceedings of the 19th IFAC World Congress, Cape Town, South Africa (2014)
10. Colmegna, P., Sanchez-Pena, R.S., Gondhalekar, R., Dassau, E., Doyle III, F.J.: Reducing risk in Type 1 diabetes using H_∞ control. IEEE Trans. Biomed. Eng. **61**(12), 2939–2947 (2014)
11. Enns, D.F.: Model reduction with balanced realizations: an error bound and a frequency weighted generalization. In: 23rd IEEE Conference on Decision and Control, pp. 127–132 (1984)
12. Facchinetti, A., Sparacino, G., Cobelli, C.: Modeling the error of continuous glucose monitoring sensor data: critical aspects discussed through simulation studies. J. Diabetes Sci. Technol. **4**(1), 4–14 (2010)
13. Femat, R., Ruiz-Velazquez, E., Quiroz, G.: Weighting restriction for intravenous insulin delivery on T1DM patient via $\mathscr{H}_\infty$ control. IEEE Trans. Autom. Sci. Eng. **6**, 239–247 (2009)
14. Fonyo, A., Ligeti, E.: Physiology, 3rd edn. Medicina, Budapest (2008)
15. Haykin, S.: Kalman Filtering and Neural Networks. Wiley, New York (2002)
16. Hovorka, R., Canonico, V., Chassin, L., Haueter, U., Massi-Benedetti, M., Federici, M.O., Pieber, T., Schaller, H., Schaupp, L., Vering, T., Wilinska, M.: Nonlinear model predictive control of glucose concentration in subjects with Type 1 diabetes. Physiol. Meas. **25**, 905–920 (2004)
17. Isidori, A.: Nonlinear Control Systems. Springer, London (1995)
18. Jia, B., Xin, M., Cheng, Y.: Sparse-grid quadrature nonlinear filtering. Automatica **48**, 327–341 (2012)
19. Julier, S.J., Uhlmann, J.K.: Unscented filtering and nonlinear estimation. Proc. IEEE **92**(3), 401–422 (2004)
20. Julier, S., Uhlmann, J., Durrant-Whyte, H.: A new approach for filtering nonlinear systems. Proc. Amer. Control Conf. **3**, 1628–1632 (1995)
21. Kovács, L., Benyó, B., Bokor, J., Benyó, Z.: Induced $\mathscr{L}$-norm minimization of glucose-insulin system for Type I diabetic patients. Comput. Methods. Progr. Biomed. **102**, 105–118 (2011)
22. Kovatchev, B., Cobelli, C., Renard, E.: Multi-national study of subcutaneous model-predictive closed-loop control in Type 1 diabetes: summary of the results. J. Diabetes Sci. Technol. **4**, 1374–1381 (2010)

23. Lee, L.: Identification and robust control of linear parameter-varying systems. Ph.D. thesis, University of California at Berkeley, USA (1997)
24. Magni, L., Raimondo, D.M., Dalla Man, C., Nicolao, G., Kovatchev, B., Cobelli, C.: Model predictive control of glucose concentration in Type I diabetic patients: an In Silico trial. Biomed. Signal Process Control **4**(4), 338–346 (2009)
25. Palerm, C.: Physiologic insulin delivery with insulin feedback: a control systems perspective. Comput. Methods. Progr. Biomed. **102**, 130–137 (2011)
26. Palumbo, P., Pizzichelli, G., Panunzi, S., Pepe, P., De Gaetano, A.: Model-based control of plasma glycemia: tests on populations of virtual patients. Math. Biosci. **257**, 2–10 (2014)
27. Parker, R., Doyle, F., Ward, J., Peppas, N.: Robust $\mathscr{H}_\infty$ glucose control in diabetes using a physiological model. AIChE J. **46**, 2537–2549 (2000)
28. Phillip, M., Battelino, T., Atlas, E., Kordonouri, O., Bratina, N., Miller, S., Biester, T., Stefanija, M., Muller, I., Nimri, R., Danne, T.: Nocturnal glucose control with an artificial pancreas at a diabetes camp. N. Engl. J. Med. **368**, 824–833 (2013)
29. Ruiz-Velazquez, E., Femat, R., Campos-Delgado, D.: Blood glucose control for Type 1 diabetes mellitus: a robust tracking H_∞ problem. Control Eng. Pract. **12**, 1179–1195 (2004)
30. Rutten, M.: Square-root unscented filtering and smoothing. In: IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing, pp. 294–299 (2013)
31. Sorensen, J.T.: A physiologic model of glucose metabolism in man and its use to design and assess improved insulin therapies for diabetes. Ph.D. Thesis, Department of Chemical Engineering Massachusetts Institute of Technology, USA (1985)
32. Szalay, P., Eigner, G., Kozlovszky, M., Rudas, I., Kovács, L.: The significance of LPV modeling of a widely used T1DM model. In: EMBC 35th Annual International Conference of the IEEE, pp. 3531–3534, Osaka, Japan (2013)
33. Tempo, R., Calafiore, G., Dabbene, F.: Randomized Algorithms for Analysis and Control of Uncertain Systems. Springer, London (2013)
34. Turner, R., Rasmussen, C.: Model based learning of sigma points in unscented kalman filtering. In: IEEE International Workshop on Machine Learning for Signal Processing (MLSP), pp. 178–183 (2010)
35. Wilinska, M., Chassin, L., Acerini, C., Allen, J., Dunger, D., Hovorka, R.: Simulation environment to evaluate closed-loop insulin delivery systems in Type 1 diabetes. J. Diabetes Sci. Technol. **4**, 132–144 (2010)
36. Zarkogianni, K., Vazeou, A., Mougiakakou, S., Prountzou, A., Nikita, K.: An insulin infusion advisory system based on autotuning nonlinear model-predictive control. IEEE Trans. Biomed. Eng. **58**, 2467–2477 (2011)
37. Zhou, K.: Frequency-weighted $\mathscr{L}_\infty$ norm and optimal hankel norm model reduction. IEEE Trans. Autom. Control **40**(10), 1687 (1995)
38. Zhou, K.: Robust and Optimal Control. Prentice Hall, New Jersey (1996)
39. Zisser, H., Palerm, C.C., Bevier, W.C., Doyle III, F.J., Jovanovic, L.: Clinical update on optimal prandial insulin dosing using a refined run-to-run control algorithm. J. Diabetes Sci. Technol. **3**, 487–491 (2009)

Recent Results on Glucose–Insulin Predictions by Means of a State Observer for Time Delay Systems

Pasquale Palumbo, Pierdomenico Pepe, Simona Panunzi and Andrea De Gaetano

Abstract To achieve accurate and affordable predictions of glucose and insulin plasma concentrations is of paramount importance, especially in the field of the artificial pancreas, where real-time measurements could be properly exploited in model-based glucose control algorithms. This note focuses on a recently developed research line that makes use of a state observer to estimate insulin in real-time from glucose measurements, since it is known that insulin measurements are slower and more cumbersome to obtain, more expensive and also less accurate. Based on these predictions, glucose control algorithms can be designed and can be exploited for both intravenous and subcutaneous insulin infusions. The safety, robustness, and efficacy of the observer-based control algorithms have been validated on a population of rather heterogenous virtual patients, modeled by a different, comprehensive model of the glucose–insulin system, recently accepted by the *Food and Drug Administration* as a substitute of animal trials.

1 Introduction

Diabetes mellitus (DM) is a worldwide disease with an alarming increase, especially in the developing countries; the diabetic population, estimated to be around 171 million people in 2000, has been predicted to double within 2030, [38]. To diagnose and contain the spread of DM disease, glucose and insulin predictions are of great importance, especially when required in closed-loop, real-time algorithms for the artificial pancreas. Differently from plasma glycemia, which can be straightforwardly measured with relative low-cost devices and affordable algorithms, plasma insulinemia

P. Palumbo (✉) · S. Panunzi · A. De Gaetano
BioMatLab, IASI "A. Ruberti", National Council of Researches, UCSC Largo A. Gemelli 8, 00168 Rome, Italy
e-mail: pasquale.palumbo@iasi.cnr.it

P. Pepe
Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, Via Vetoio, 67100 L'aquila, Italy

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*, Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_12

is slower and more cumbersome to obtain, more expensive and also less accurate. This fact stimulates the study of algorithms capable of providing the plasma insulin concentration by processing a stream of glycemia measurements. When these algorithms exploit the available measurements as coming from a known deterministic dynamical model generating them, we deal with *state observers*. The importance of these powerful tools is due to the great variety of observer-based control laws applicable, at least in theory, to the glucose control problem, with exogenous insulin administration playing the role of the control input. Observer-based closed-loop control laws belong to the field of *model-based* strategies, that means the regulator is synthesized by explicitly exploiting the structure of the model equations [4].

Differently from the great majority of model-based approaches, which use nonlinear ordinary differential equation (ODE) models, a nonlinear discrete-delay differential equation (DDE) model of the glucose/insulin system is considered [20, 32]. Motivation to use DDE models is that they provide a better representation of the pancreatic insulin delivery rate (IDR) (e.g., [12, 17] and references therein), therefore allowing to treat in a unified fashion both Type 1 and Type 2 diabetic patients, these latter with a not negligible IDR. More in detail, the adopted glucose/insulin DDE model has been shown to exhibit a number of desirable characteristics, such as to conform to established physiological concepts (e.g., pancreatic insulin secretion rate is limited), to exhibit satisfactory properties of the solutions (e.g., positivity and boundedness of solutions, local attractivity of a single positive equilibrium, [20]), and to be statistically robust, in that its parameters are identifiable with very good precision, fitting the model onto observations from standard perturbation experiments, such as the intravenous glucose tolerance test (IVGTT), [31, 32].

Different results are here reported, dealing with theoretical design (feedback linearization with delay cancelation, state observers, asymptotic convergence, local input-to-state stability) and practical issues (discretization of the control algorithm, measurement uncertainties, insulin pump malfunctioning, intrapatient variability) according to both intravenous [22, 23, 25, 26] and subcutaneous [21, 24, 28] insulin delivery. Although the proposed control laws require the knowledge of the whole state of the system (i.e., both glycemia and insulinemia, possibly also at retarded instants) only glucose measurements have been exploited, leaving the task to provide real-time insulin estimates to an observer for nonlinear time delay systems. Safety, efficacy, and robustness have been recently validated [30] on a population of *Virtual Patients* (VP) generated by a different, comprehensive model of the glucose–insulin system, recently accepted by the *Food and Drug Administration* as a substitute of animal trials, [15].

The chapter continues as follows. Next Section is devoted to introduce the DDE model of the glucose–insulin system adopted for the state prediction task and, consequently, for the synthesis of the closed-loop control. Sections 3 and 4 deal with the intravenous and subcutaneous insulin delivery modes, detailing with the most important results achieved. Conclusions follow.

2 The DDE Model of the Glucose–Insulin System

The model-based algorithms here reported for the artificial pancreas have been synthesized according to the following DDE model of the glucose–insulin system [20, 32]. Symbols 'mmol' and 'pmol' stand for *millimoles* and *picomoles*, with 'mM' and 'pM' denoting *mmoles/liter* and *pmol/liter*, respectively. The equations are written with respect to plasma glycemia, $G(t)$, [mM], and insulinemia, $I(t)$, [pM]:

$$\begin{aligned}\frac{dG(t)}{dt} &= -K_{xgi}G(t)I(t) + \frac{T_{gh}}{V_G} + \frac{d(t)}{V_G},\\ \frac{dI(t)}{dt} &= -K_{xi}I(t) + \frac{T_{iGmax}}{V_I}f\big(G(t-\tau_g)\big) + \frac{u(t)}{V_I},\end{aligned} \tag{1}$$

with K_{xgi}, [min^{-1} pM^{-1}], the rate of (insulin-dependent) glucose uptake by tissues per pM of plasma insulin concentration; T_{gh}, [(mmol/kgBW)/min], the net balance between hepatic glucose output and insulin-independent zero-order glucose tissue uptake; V_G, V_I, [L/kgBW], the apparent distribution volumes for glucose and insulin, respectively; K_{xi}, [min^{-1}], the apparent first-order disappearance rate constant for insulin; T_{iGmax}, [(pmol/kgBW)/min], the maximum rate of second-phase insulin release; τ_g, [min], the apparent delay with which the pancreas varies secondary insulin release in response to varying plasma glucose concentrations.

The following Hill function is chosen for the nonlinear map $f(\cdot)$, modeling the endogenous pancreatic insulin delivery rate:

$$f(G) = \frac{(G/G^*)^\gamma}{1+(G/G^*)^\gamma}, \tag{2}$$

with γ the Hill coefficient and G^*, [mM], the glycemia at which the insulin release is half of its maximum rate.

The exogenous signals $u(t)$, [(pmol/kgBW)/min], and $d(t)$, [(mmol/kgBW)/min], are the control input (i.e., the external insulin delivery rate) and a disturbance in the glucose dynamics (e.g., the glucose intake from a meal). As a matter of fact, model (1) does not consider the subcutaneous depot assuming that insulin is straightforwardly delivered intravenously. On the other hand, in case of subcutaneous insulin administration, Eq. (1) are modified as follows:

$$\begin{aligned}\frac{dG}{dt} &= -K_{xgi}G(t)I(t) + \frac{T_{gh}}{V_G} + \frac{d(t)}{V_G},\\ \frac{dI}{dt} &= -K_{xi}I(t) + \frac{T_{iGmax}}{V_I}f\big(G(t-\tau_g)\big) + \frac{S_2(t)}{V_I t_{max,I}},\\ \frac{dS_2}{dt} &= \frac{1}{t_{max,I}}S_1(t) - \frac{1}{t_{max,I}}S_2(t),\\ \frac{dS_1}{dt} &= -\frac{1}{t_{max,I}}S_1(t) + u(t),\end{aligned} \tag{3}$$

with S_1, S_2 [pmol/kgBW] the insulin mass in the accessible and not accessible subcutaneous depot, respectively, and $t_{max,I}$, [min], the time-to-maximum insulin absorption. In this framework, the control input $u(t)$, [pmol/kgBW/min], is the exogenous insulin infusion rate, delivered subcutaneously. The model of insulin absorption here adopted (third and fourth equations in (3)) refers to [35] with no insulin degradation at the injection site. It has been recently analyzed in [39], and it has been exploited with the aim of glucose control in [11], according to which, here we assume the same notation.

As far as the initial conditions, the subject is supposed to be at rest before the insulin therapy starts, so that plasma glycemia and insulinemia are equal to the constant (hyperglycemic for uncontrolled diabetic patients) basal levels (G_b, I_b):

$$G(\tau) = G_b, \quad I(\tau) = I_b, \qquad \tau \in [-\tau_g, 0]. \tag{4}$$

In case of model (3), besides (4) we have that the subcutaneous depots are empty; therefore,

$$S_1(\tau) = 0, \quad S_2(\tau) = 0, \qquad \tau \in [-\tau_g, 0]. \tag{5}$$

It has to be stressed that models (1–3) may represent equally well healthy subjects and insulin-resistant or severely insulin-deficient diabetic patients, by appropriately changing the parameter values. Moreover, it does belong to the class of "minimal models," in the sense that according to a "minimal" set of independent parameters, it allows to very well resemble the physiology of the glucose/insulin kinetics, and it is identifiable from data with very good precision, according to IVGTT standard perturbation experiments (see [31, 32]).

3 Observer-Based Control by Means of Intravenous Insulin Infusion

Exogenous insulin administration is the basic procedure to cope with diabetes; for Type 1 patients only exogenous insulin is available, while for Type 2 exogenous insulin complements pancreatic production. The use of intravenous insulin administration, delivered by automatic, variable speed pumps under the direct supervision of a physician, provides a wide range of possible strategies and ensures a rapid delivery with negligible delays. As a matter of fact, control algorithms based on intravenous infusions (we can cite, among the others, [5, 7, 13, 14, 25, 33, 36]) are directly applicable so far only to problems of glycemia stabilization in critically ill subjects, such as in surgical intensive care units after major procedures, [37].

The aim of the proposed control law therapies is to reduce a high basal plasma glucose concentration to a lower level, according to a smooth reference glucose trajectory $G_{\text{ref}}(t)$. The equations of model (1) will be taken into account. The control law is synthesized disregarding any disturbances in the glucose dynamics, like meals.

The ability of the controller to reject these disturbances will play a crucial role in the evaluation of the artificial pancreas.

3.1 Synthesis of the Glucose Control Law

Define the tracking error as follows:

$$e(t) = \begin{bmatrix} e_1(t) \\ e_2(t) \end{bmatrix} = Z(t) - Z_{\text{ref}}(t) \tag{6}$$

with:

$$Z(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} G(t) \\ -K_{xgi}G(t)I(t) + \frac{T_{gh}}{V_G} \end{bmatrix}, \qquad Z_{\text{ref}}(t) = \begin{bmatrix} G_{\text{ref}}(t) \\ \dot{G}_{\text{ref}}(t) \end{bmatrix}, \tag{7}$$

where $G_{\text{ref}}(t)$ is the glucose reference signal to be tracked. It is supposed to be bounded (with lower bound strictly positive), twice continuously differentiable, with bounded first and second derivatives. Since $G_{\text{ref}}(t)$ provides the desired plasma glucose concentration, it clearly comes that these constraints readily match with a physiologically meaningful choice. Examples of suitable $G_{\text{ref}}(t)$ can be found in [22, 25, 30], where a function exponentially decreasing from the hyperglycemic basal state down to a safe euglycemic level is chosen.

In [22], according to the theory of input–output feedback linearization with delay cancelation (see [8, 10, 19]), with respect to the output $y(t) = G(t)$ and the input $u(t)$, it is shown that, by applying the control input

$$\frac{u(t)}{V_I} = \frac{S\big(G(t), I(t), G(t-\tau_g)\big) - v(t)}{K_{xgi}G(t)}, \qquad t \geq 0, \tag{8}$$

with

$$\begin{aligned} S\big(G(t), I(t), G(t-\tau_g)\big) = &-K_{xgi}I(t)\left(-K_{xgi}I(t)G(t) + \frac{T_{gh}}{V_G}\right) \\ &- K_{xgi}G(t)\left(-K_{xi}I(t) + \frac{T_{iGmax}}{V_I}f\big(G(t-\tau_g)\big)\right) \end{aligned} \tag{9}$$

and

$$v(t) = \ddot{G}_{\text{ref}}(t) + Re(t), \tag{10}$$

the tracking error dynamics can be written as:

$$\dot{e}(t) = He(t), \qquad H = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} R. \tag{11}$$

Because of the structure of H, the control matrix $R \in \mathbb{R}^{1\times 2}$ can be designed such that matrix H has prescribed eigenvalues in the left half complex plane, thus ensuring the exponential convergence to zero of the tracking error, that implies $G(t) \mapsto G_{\text{ref}}(t)$.

Such a control law (8–10) requires both glucose and insulin measurements; on the other hand, insulin measurements are slower and more cumbersome to obtain, more expensive, and also less accurate than glucose measurements: a need exists, therefore, to construct a control law avoiding the measurements of insulin serum. For these reasons, a state observer for system (1) has been considered in [25], in order to estimate the plasma insulin concentration and design a feedback control law based on only glucose measurements. To this end, denote with $\widehat{G}(t), \widehat{I}(t)$ the glucose and insulin estimates and consider the following equations for the observer:

$$\begin{aligned} \begin{bmatrix} d\widehat{G}/dt \\ d\widehat{I}/dt \end{bmatrix} = & \begin{bmatrix} -K_{xgi}\widehat{G}(t)\widehat{I}(t) + \frac{T_{gh}}{V_G} \\ -K_{xi}\widehat{I}(t) + \frac{T_{iGmax}}{V_I} f\big(\widehat{G}(t-\tau_g)\big) + \frac{u(t)}{V_I} \end{bmatrix} \\ & + Q^{-1}\big(\widehat{G}(t), \widehat{I}(t)\big) W\big(G(t) - \widehat{G}(t)\big), \end{aligned} \tag{12}$$

where Q^{-1} is the inverse matrix of the matrix function $Q(x_1, x_2) \in \mathbb{R}^{2\times 2}$ defined as

$$Q(x_1, x_2) = \begin{bmatrix} 1 & 0 \\ -K_{xgi}x_2 & -K_{xgi}x_1 \end{bmatrix}, \tag{13}$$

and the observer gain matrix $W \in \mathbb{R}^{2\times 1}$ is designed to ensure that

$$\widehat{H} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - W \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{14}$$

is Hurwitz with prescribed eigenvalues in the left half complex plane. The initial conditions for (12) are formally given by

$$\widehat{G}(\tau) = \widehat{G}_0(\tau), \quad \widehat{I}(\tau) = \widehat{I}_0(\tau), \qquad \tau \in [-\tau_g, 0]. \tag{15}$$

According to [25] and references therein, it comes that the observer can be designed such that, if the estimation error at zero is sufficiently small and the input signal is suitably bounded, the estimation error converges exponentially to zero, with arbitrary

decay rate fixed by means of a suitable choice of W. Moreover, substitute the real glucose and insulin values in (8–9) with the estimates coming from the observer:

$$\frac{u(t)}{V_I} = \frac{S\big(\widehat{G}(t), \widehat{I}(t), \widehat{G}(t-\tau_g)\big) - v(t)}{K_{xgi}\widehat{G}(t)}, \qquad t \geq 0, \tag{16}$$

and consider

$$v(t) = \ddot{G}_{\text{ref}}(t) + R\hat{e}(t) \tag{17}$$

with the target error $\hat{e}(t)$ defined by $\hat{e}(t) = \widehat{Z}(t) - Z_{\text{ref}}(t)$, with:

$$\widehat{Z}(t) = \begin{bmatrix} \hat{z}_1(t) \\ \hat{z}_2(t) \end{bmatrix} = \begin{bmatrix} \widehat{G}(t) \\ -K_{xgi}\widehat{G}(t)\widehat{I}(t) + \frac{T_{gh}}{V_G} \end{bmatrix}. \tag{18}$$

It is proven in [25] that there exist gain matrices R, W such that the closed-loop system given by (1) with $d(t) = 0$, (12), (16–17) ensures that the plasma glycemia is controlled to track the reference trajectory, with the error tracking asymptotically converging to zero, provided that the initial tracking and observer errors are suitably small. It worths noticing that such a result is ensured by exploiting both the estimated glucose and insulin concentrations in the control law. To use the real (and available) glucose measurements, instead of the estimated glycemia, would not ensure improvements nor (and more important) would it ensure to maintain the aforementioned theoretical results.

Remark 1 Since the theory does not explicitly take into account the impossibility to release negative insulin, in this cases, the regulator would temporarily switch off, leaving the patient without control for an unpredictable period. This is clearly an undesirable situation, to be avoided. Another important requirement is to prevent glucose oscillations, possibly determining dangerous hypoglycemia. Both these issues need to be addressed in the setting of the control parameters (i.e., matrices R and W), as well as in the setting of the reference glucose trajectory, realizing a tradeoff between closed-loop fast asymptotic stability (suitably negative real part eigenvalues) and transient behavior (possibly smooth trajectories without oscillations). Simulations reported in [30] (see Figs. 2 and 3) show that the issue to prevent the switching off of the regulator cannot be ensured, at least when implementing the control law on a virtual environment, accounting for uncertainties affecting sensors and actuators devices as well as model parameter uncertainties; nevertheless, undesired oscillations can successfully be avoided.

3.2 Evaluation Criteria and Validation

Numerical simulations run by closing the control loop on (1) allowed to set the control parameters in order to track safely the desired glycemia within a couple of hour of

insulin administration (see [25] for the details). Of course, a realistic validation requires to build up an environmental framework which comprises measurement uncertainties, insulin pump malfunctioning, intrapatient variability as well as the discretization of the algorithm, because of real-time devices involving digital glucose sensors and insulin pumps: the former provide quite reliable measurements of plasma glycemia at given sample times, whose frequency is limited by the time needed to analyze plasma glucose on a bed-side analyzer, [3]; the latter are used to administer insulin by means of piecewise-constant infusions. To this end, simulations have been carried out by discretizing the proposed control law at suitable sampling period Δ, according to the following scheme ($k = 0, 1, \dots$):

1. at time $k\Delta$ the measurement of $G(k\Delta)$ is delivered by the sensor;
2. at time $k\Delta$ the control input is computed by means of the available state estimates $\widehat{G}(k\Delta)$, $\widehat{G}(k\Delta - \tau_g)$, $\widehat{I}(k\Delta)$, see (16):

$$\frac{u(k\Delta)}{V_I} = \frac{S\left(\widehat{G}(k\Delta), \widehat{I}(k\Delta), \widehat{G}(k\Delta - \tau_g)\right) - v(k\Delta)}{K_{xgi}\widehat{G}(k\Delta)}; \tag{19}$$

   if $u(k\Delta) < 0$, then $u(k\Delta)$ is forced to be 0;
3. the constant infusion $u(k\Delta)$ is administered to the patient in the time interval $[k\Delta, (k+1)\Delta)$;
4. simultaneously with item [3], the controller device integrates numerically the following equation, see (12), in the time interval $[k\Delta, (k+1)\Delta)$, using the available measurement $G(k\Delta)$ and past estimations:

$$\begin{bmatrix} d\widehat{G}/dt \\ d\widehat{I}/dt \end{bmatrix} = \begin{bmatrix} -K_{xgi}\widehat{G}(t)\widehat{I}(t) + \frac{T_{gh}}{V_G} \\ -K_{xi}\widehat{I}(t) + \frac{T_{iGmax}}{V_I} f\left(\widehat{G}(t - \tau_g)\right) + \frac{u(k\Delta)}{V_I} \end{bmatrix} + Q^{-1}\left(\widehat{G}(t), \widehat{I}(t)\right) W\left(G(k\Delta) - \widehat{G}(k\Delta)\right); \tag{20}$$

5. the value of k is incremented by 1.

Notice that (20) provides the state estimation in the prediction interval $[k\Delta, (k+1)\Delta)$. A crucial point in validating the synthesized control law has been to exploit two distinct models of the glucose–insulin system [30]. The minimal model (1), easy-to-handle from the control perspective, though physiologically meaningful and mathematically coherent, is exploited to synthesize the control law. A different model, more comprehensive, multicompartmental, is exploited to simulate a *Virtual Patient* onto close the artificial pancreas. To this end, the model published in [6], recently accepted by the *Food and Drug Administration* as a substitute of animal trials [15], has been considered. The idea is sketched in Fig. 1. The *Virtual Patient* stands for the chosen model for validation (and, in a hopefully not too far future, could be replaced by a real patient). The *model identification* dashed arrow pointing to the *DDE model-based controller* from the Virtual Patient block refers to a methodology to identify the DDE model parameters on the basis of a chosen virtual procedure (for instance, in [30] we considered a virtual IVGTT). The *control*

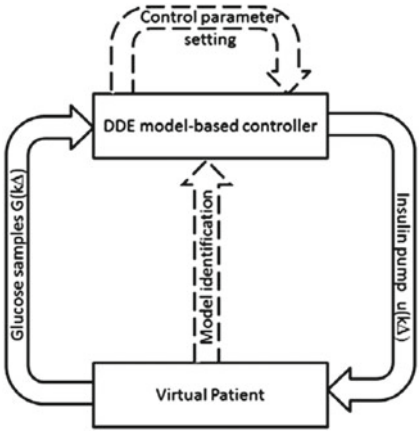

Fig. 1 Sketch of the virtual environment built up to validate the model-based control law

parameter setting dashed arrow working on the DDE model-based controller block stands for the setting of the control parameters (i.e., matrices R and W), a procedure achieved by means of *trial and error* simulations run on the DDE model without assuming failures or malfunctioning. The solid arrows refer to the closed-loop system, which is implemented once the aforementioned preliminary two tasks have been properly fulfilled: the Virtual Patient provides sampled glucose measurements to the controller, and the DDE model-based controller provides the proper insulin infusion rate to be applied in the time interval $[k\Delta, (k+1)\Delta)$.

In [30] the control parameters have been set once and for all the subjects of a population generated by sampling the many parameters of the virtual patient provided by [6], each with a 5 % of coefficient of variation (CV). Moreover, glucose measurements are supposed to be affected by a 5 % of CV as well as the insulin infusion is supposed to be affected by a 15 % of CV.

Safety and efficacy criteria have been set according to the ones reported in [3]. The application of these criteria to a population of VPs, with an average basal glycemia of 9 mM, provides very interesting results. In case of absence of meals, no disturbance signal $d(t)$ in (1), on a population of 10,000 VPs, no hypoglycemia cases are reported (glycemia never reduces below 3.3 mM), in favor of very good efficacy results: more than 99.8 % of diabetic VPs definitely reduce glycemia below 7 mM within the first 3 h of simulated experiments. These results are shown to be robust with respect to a discretization period $\Delta \in [5, 15]$ min. Analogous very good results are obtained by applying the same control law on a 24 h temporal period, accounting for the administration of three meals (treated as unknown disturbances). Indeed, besides the complete absence of hypoglycemia cases, glycemia is constrained below 11 mM within the 2 h from the meal administration and during the period before the successive meal, for a set of diabetic VPs which exceeds 95 % of the population (see [30] for the details).

Remark 2 It is well known that disturbances like meals are hard to anticipate in timing, in amount and in the rate of effective absorption of the nutrient. By treating

the meal $d(t)$ in (1) as a completely unknown disturbance, it is shown in [26] that the aforementioned closed-loop system (1), (12), (16–18) enjoys the local input-to-state stability property with respect to the unknown disturbance $d(t)$.

Remark 3 Copying with the need of an implementable discrete-time control law, a different philosophy would be to work on a discretized glucose–insulin system and then synthesize the control law according to the discrete framework. In [23], the DDE model in (1) has been discretized according to [1] and the digital control law is designed according to [34] where the dummy output proposed in [1] is used, in order to preserve under sampling the full relative degree of the continuous time model, up to an approximation of order 3 in the sampling period. The digital control law proposed in [23] guarantees the asymptotic stability of the suitably approximated sampled glucose–insulin system.

4 Observer-Based Control by Means of Subcutaneous Insulin Infusion

Glucose control strategies actuated by means of subcutaneous insulin administration (see [2] and references therein) are easier to realize and, indeed, they can be managed nowadays by the patients themselves. However, in order to design closed-loop control algorithms, the insulin absorption from the subcutaneous depot needs also to be considered (see, among the others [11, 16, 21]).

In order to synthesize a subcutaneously delivered insulin therapy, the model-based approach requires to endow the glucose–insulin DDE model (1) with the subcutaneous insulin absorption compartment, see the complete model (3). The goal is to achieve a desired euglycemic glucose level for patients with a basal hyperglycemic state. No exogenous glucose disturbances are here considered (i.e., $d(t) = 0$ in (3)).

According to [24, 28], by applying the theory of exact input–output feedback linearization with delay cancellation (see [8, 10, 18, 19]), with respect to the input $u(t)$ and the output $G(t) - G_d$ (G_d is the desired level of glycemia), the following control law is found:

$$u(t) = V_I t_{max,I}^2 \frac{\alpha(\cdot) - v(t)}{K_{xgi} G(t)}, \tag{21}$$

where $\alpha(\cdot)$ is a function of the system variables at the present time $G(t)$, $I(t)$, $S_1(t)$, $S_2(t)$, and of some of them at delayed times (namely of $G(t - \tau_g)$, $I(t - \tau_g)$, $S_2(t - \tau_g)$ and $G(t - 2\tau_g)$), see [24] for the explicit formulation. By applying (21)

to (3), it comes that the closed-loop system may be written using new variables $z(t) = [G(t)\ \ G^{(1)}(t)\ \ G^{(2)}(t)\ \ G^{(3)}(t)]^T$ as:

$$\dot{z}(t) = A_b z(t) + B_b v(t), \qquad t \geq 0, \tag{22}$$

with A_b, B_b the fourth-order Brunowski pair and the brackets $^{(i)}$ denoting the ith time derivative. By setting the outer input $v(t) = \Gamma e(t)$, with

$$e(t) = z(t) - z_d, \qquad z_d = [G_d\ \ 0\ \ 0\ \ 0]^T, \tag{23}$$

the error dynamics becomes:

$$\dot{e}(t) = He(t), \qquad H = A_b + B_b \Gamma. \tag{24}$$

Since the Brunowski pair is controllable, we design the gain matrix $\Gamma \in \mathbb{R}^{1\times 4}$ in order to make Hurwitz H and, therefore, the tracking error between plasma glycemia and its reference signal (the first component of $e(t)$) converges exponentially to zero.

Besides the drawback of real-time insulin measurements involved in the aforementioned glucose control law (shared also by the intravenous insulin delivery therapies dealts with in Sect. 3), here also subcutaneous insulin measurements are required, which are quite impossible to obtain, especially in a real-time closed-loop framework. In order to overcome such problems, in [21] a state observer for system (3) is considered, with the aim of estimating the insulin on the basis of continuous time glucose measurements.

In order to design the state observer for (3), define $X(t) = [G(t)\ \ I(t)\ \ S_2(t)\ \ S_1(t)]^T \in \mathbb{R}^4$ such that the DDE system (3) with $d(t) = 0$ can be formally written in the more compact form:

$$\dot{X}(t) = \mathscr{F}\big(X(t), X(t-\tau_g)\big) + B_b u(t), \tag{25}$$

where $\mathscr{F} : \mathbb{R}^4 \times \mathbb{R}^4 \mapsto \mathbb{R}^4$ is defined for $X = [X_1\ X_2\ X_3\ X_4]^T, Y = [Y_1\ Y_2\ Y_3\ Y_4]^T \in \mathbb{R}^4$ as:

$$\mathscr{F}(X, Y) = \begin{pmatrix} -K_{xgi} X_1 X_2 + \frac{T_{gh}}{V_G} \\ -K_{xi} X_2 + \frac{T_{iGmax}}{V_I} f(Y_1) + \frac{1}{t_{max,I}} X_3, \\ \frac{1}{t_{max,I}} X_4 - \frac{1}{t_{max,I}} X_3 \\ -\frac{1}{t_{max,I}} X_4 \end{pmatrix} \tag{26}$$

The measured output is, then, given by $y(t) = G(t) = C_b X(t)$, where $C_b = [1\ \ 0\ \ 0\ \ 0]$.

The observer for system (25–26) adopted in [21] is the one developed in [9], given by the following neutral system, with $\widehat{X}(t), w(t) \in \mathbb{R}^4$:

$$\dot{\widehat{X}}(t) = \mathscr{F}\big(\widehat{X}(t), \widehat{X}(t-\tau_g)\big) + B_b u(t) + w(t), \qquad t \geq 0, \tag{27}$$

with

$$w(t) = Q^{-1}\big(\widehat{X}(t), \widehat{X}(t-\tau_g)\big)\Big(W\big(y(t) - C_b\widehat{X}(t)\big) - Q_1\big(\widehat{X}(t), \widehat{X}(t-\tau_g)\big)w(t-\tau_g)\Big). \tag{28}$$

Matrices

$$\begin{aligned} Q\big(\widehat{X}(t), \widehat{X}(t-\tau_g)\big) &= \frac{\partial\Theta\left(\widehat{X}(t), \widehat{X}(t-\tau_g)\right)}{\partial\widehat{X}(t)} \\ Q_1\big(\widehat{X}(t), \widehat{X}(t-\tau_g)\big) &= \frac{\partial\Theta\left(\widehat{X}(t), \widehat{X}(t-\tau_g)\right)}{\partial\widehat{X}(t-\tau_g)} \end{aligned} \tag{29}$$

are obtained from the partial derivatives of the function $\Theta(\cdot,\cdot)$ (see [9]), which is formally defined as the aggregate of the output $G(t)$ and its first three time derivatives, obtained according to (3). $\Theta(\cdot,\cdot)$ is a function of the system variables at the present time (i.e., $G(t)$, $I(t)$, $S_1(t)$, $S_2(t)$) and of some of them at the delayed time (namely $G(t-\tau_g)$ and $I(t-\tau_g)$) (see [21] for the explicit expression of $\Theta(\cdot,\cdot)$). When the function Θ is used for computations in (29), $\widehat{X}(t)$ takes the place of $[G(t)\ I(t)\ S_2(t)\ S_1(t)]^T$, for any required time t. The gain matrix $W \in \mathbb{R}^{4\times 1}$ is chosen such that the matrix $\widehat{H} = A_b - WC_b$ is Hurwitz.

It has been shown in [9] that the gain matrix W can be properly designed to ensure the global asymptotic convergence to zero of the observation error, provided that proper conditions are satisfied. Such conditions are not completely satisfied by the system at hand (for instance, the functions involved in (3) are not globally Lipschitz). However, we are not interested in the convergence of the observation error to zero for any initial state and input signal in bounded sets, as in [9], but only in the convergence, at least locally, of the state variables of the closed-loop system to the desired equilibrium. To this end, we exploit the observer equations to close the loop from the observed state $\widehat{X}(t)$, so that the state variables in (21) are replaced by their estimates. Then, the control law becomes:

$$u(t) = V_I t_{max,I}^2 \frac{\alpha(\hat{\cdot}) - v(t)}{K_{xgi} C_b \widehat{X}(t)}, \qquad v(t) = \Gamma\big(\hat{z}(t) - z_d\big), \tag{30}$$

with $\hat{z}(t) = \Theta(\widehat{X}(t), \widehat{X}(t-\tau_g)$ and $\alpha(\hat{\cdot})$ denoting the function $\alpha(\cdot)$ computed in the observed state variables $\widehat{X}(t)$, $\widehat{X}(t-\tau_g)$, $\widehat{X}(t-2\tau_g)$, instead of the real ones $X(t)$, $X(t-\tau_g)$, $X(t-2\tau_g)$.

The main result provided by [21] is that there exist matrices W and Γ such that, as long as the initial estimation error and the initial tracking error are sufficiently small, the evolution of the closed-loop system (25)–(30) asymptotically converges to $\mathscr{X}_E = [X_E^T\ \ X_E^T] \in \mathbb{R}^8$, where $X_E = [G_d\ \ I_d\ \ S_{2,d}\ \ S_{1,d}]^T$ is the equilibrium point of (3) when the control input $u(t) = u_d = S_{1,d}/t_{max,I}$ (see next remark) is provided.

Remark 4 The constant terms $G_d, I_d, S_{2,d}, S_{1,d}$ are such that, if at a given time instant $\bar{t} \geq 0$, it is $G(\tau) = G_d$, $I(\tau) = I_d$, $S_2(\tau) = S_{2,d}$, $S_1(\tau) = S_{1,d}$, for $\tau \in [\bar{t} - 2\tau_g, \bar{t}]$, and the control law $u(t)$ designed as in (21) is applied, then the solution of (3) with

$d(t) = 0$ is $G(t) = G_d, I(t) = I_d, S_2(t) = S_{2,d}, S_1(t) = S_{1,d}$, for $t \geq \bar{t}$, and the control law becomes $u(t) = u_d$, $t \geq \bar{t}$. In other words, once G_d has been chosen, we may compute I_d, $S_{2,d}$, $S_{1,d}$ and u_d as the reference levels of the system variables and of the control input that asymptotically correspond to a perfect tracking of G_d. As a matter of fact, the state $X_E = [G_d \;\; I_d \;\; S_{2,d} \;\; S_{1,d}]^T \in \mathbb{R}^4$ is the equilibrium point of the closed loop system (3)–(21).

5 Conclusions

The common denominator of the proposed insulin administration therapies are given by (i) the model-based approach, (ii) the use of a state observer to predict real-time insulin measurements. Indeed, only glucose measurements have been considered, with insulin administered both intravenously and subcutaneously. The resulting control laws are theoretically appealing and, at least for the intravenous case, robust with respect to many sources of uncertainties, measurement errors and actuator malfunctioning, according to a virtual environment making use of a different and more comprehensive model of the glucose–insulin system. The same philosophy has been recently applied also to a different clinical framework such as the one of the Euglycemic Hyperinsulinemic Clamp (EHC) [27, 29], a nontrivial perturbation experiment during which large amounts of insulin are administered intravenously to the subject, and exogenous glucose is administered, according to given protocols, in order to keep glycemia constant. In this chapter, we have briefly surveyed the recent results about estimation methods for the glucose–insulin system by means of state observers for time delay systems, in perspective with the control problem.

References

1. Barbot, J.P., Monaco, S., Normand-Cyrot, D.: A sampled normal form for feedback linearization. Math. Control Signal Syst. **9**, 162–188 (1996)
2. Bellazzi, R., Nucci, G., Cobelli, C.: The subcutaneous route to insulin dependent diabetes therapy, IEEE Eng. Med. Biol. **20**, 54–64 (2001)
3. Chassin, L.J., Wilinska, M.E., Hovorka, R.: Evaluation of glucose controllers in virtual environment: methodology and sample application. Artif. Intell. Med. **32**, 171–181 (2004)
4. Chee, F., Fernando, T.: Closed loop control of blood glucose. In: Thoma, M., Morari, M. (eds.) Lecture Notes in Control and Information Sciences. Springer, Berlin (2007)
5. Chee, F., Savkin, A.V., Fernando, T.L., Nahavandi, S.: Optimal H_∞ insulin injection control for blood glucose regulation in diabetic patients. IEEE Trans. Biomed. Eng. **52**, 1625–1631 (2005)
6. Dalla Man, C., Rizza R.A., Cobelli, C.: Meal simulation model of the glucose-insulin system: IEEE Trans. Biomed. Eng. **54**, 1740–1749 (2007)
7. Dua, P., Doyle, F.J., Pistikopoulos, E.N.: Model-based blood glucose control for Type 1 diabetes via parametric programming. IEEE Trans. Biomed. Eng. **53**, 1478–1491 (2006)
8. Germani, A., Manes, C., Pepe, P.: Local asymptotic stability for nonlinear state feedback delay systems. Kybernetika **36**, 31–42 (2000)

9. Germani, A., Manes, C., Pepe, P.: An asymptotic state observer for a class of nonlinear delay systems. Kybernetika **37**(4), 459–478 (2001)
10. Germani, A., Manes, C., Pepe, P.: Input-output linearization with delay cancellation for nonlinear delay systems: the problem of the internal stability. Int. J. Robust Nonlinear Control **13**(9), 909–937 (2003)
11. Hovorka, R., Canonico, V., Chassin, L.J., Haueter, U., Massi-Benedetti, M., Federici, M.O., Pieber, T.R., Schaller, H.C., Schaupp, L., Vering, T., Wilinska, M.E.: Nonlinear model predictive control of glucose concentration in subjects with type I diabetes. Physiol Meas. **25**, 905–920 (2004)
12. Kong, J.D., Kumar, S.S., Palumbo, P.: DDE models of the glucose-insulin system: a useful tool for the artificial pancreas. In Delitala, M., Ajomne-Marsan, G. (eds.) Managing Complexity, Reducing Perplexity in Biological Systems, Springer Proceedings in Mathematics & Statistics. 67, pp. 109–117 (2014)
13. Kovàcs, L., Benyò, B., Bokor, J., Benyò, Z.: Induced L2-norm minimization of glucose–insulin system for Type I diabetic patients. Comput. Meth. Prog. Biomed. **102**, 105–118 (2011)
14. Kovàcs, L., Kulcsàr, B., Gyorgy, A., Benyò, Z.: Robust servo control of a novel type 1 diabetic model. Optim Control Appl Methods. **32**, 215–238 (2011)
15. Kovatchev, B.P., Breton, M.D., Dalla Man, C., Cobelli, C.: In silico model and computer simulation environment approximating the human glucose/insulin utilization. Food and Drug Administration Master File MAF 1521 (2008)
16. Magni, L., Raimondo, D.M.: Dalla Man, C., De Nicolao, G., Kovatchev, B., Cobelli, C.: Model predictive control of glucose concentration in type I diabetic patients: an in silico trial. Biomed Signal Process Control **4**, 338–346 (2009)
17. Makroglou, A., Li, J., Kuang, Y.: Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: an overview. Appl. Numer. Math. **56**, 559–573 (2006)
18. Marquez-Martinez, L.A., Moog, C.H.: Input-output feedback linearization of time-delay systems. IEEE Trans. Autom. Control **49**(5), 781–785 (2004)
19. Oguchi, T., Watanabe, A., Nakamizo, T.: Input-output linearization of retarded nonlinear systems by using an extension of Lie derivative. Int. J. Control **75**(8), 582–590 (2002)
20. Palumbo, P., Panunzi, S., De Gaetano, A.: Qualitative behavior of a family of delay differential models of the glucose insulin system. Discrete Contin. Dyn. Syst.-Ser. B **7**, 399–424 (2007)
21. Palumbo, P., Pepe, P., Kong, J.D., Kumar, S.S., Panunzi, S., De Gaetano, A.: Regulation of the human plasma glycemia by means of glucose measurements and subcutaneous insulin administration. 3rd IFAC International Conference on Intelligent Control and Automation Science (ICONS13), 96–101, Chengdu, China (2013)
22. Palumbo, P., Pepe, P., Panunzi, S., De Gaetano, A.: Robust closed-loop control of plasma glycemia: a discrete-delay model approach. Discrete Contin. Dyn. Syst.-Ser. B **12**(2), 455–468 (2009)
23. Palumbo, P., Pepe, P., Panunzi, S., De Gaetano, A.: Digital closed-loop control of plasma glycemia. 49th IEEE Conference on Decision and Control (CDC10), 833–838, Atlanta, Georgia (2010)
24. Palumbo, P., Pepe, P., Panunzi, S., De Gaetano, A.: Glucose control by subcutaneous insulin administration: a DDE modelling approach. 18th IFAC World Congress on Automatic Control (IFAC2011). pp. 1471–1476, Milan, Italy (2011)
25. Palumbo, P., Pepe, P., Panunzi, S., De Gaetano, A.: Time-delay model-based control of the glucose-insulin system, by means of a state observer. Eur. J. Control **6**, 591–606 (2012)
26. Palumbo, P., Pepe, P., Panunzi, S., De Gaetano, A.: Observer-based closed-loop control for the glucose-insulin system: local Input-to-State Stability with respect to unknown meal disturbances, American Control Conference (ACC13), pp.1751–1756, Washington (2013)
27. Palumbo, P., Pepe, P., Panunzi, S., De Gaetano, A.: Closed-loop glucose control: application to the Euglycemic Hyperinsulinemic Clamp. 52nd IEEE Conference on Decision and Control (CDC13). pp. 4461–4466, Florence, Italy (2013)
28. Palumbo, P., Pepe, P., Panunzi, S., De Gaetano, A.: DDE model-based control of glycemia via sub-cutaneous insulin administration. In Vyhlídal, T., Lafay, J.-F., Sipahi, R. (eds.) Delay

Systems. From Theory to Numerics and Applications. Advances in Delays and Dynamics, Springer International Publishing, 1, pp. 229–240 (2014)
29. Palumbo, P., Pizzichelli, G., Panunzi, S., Pepe, De Gaetano, A.: Closed-loop control scheme for the Euglycemic Hyperinsulinemic Clamp: validation on virtual patients. IFAC World congress on Automatic Control (IFAC2014). pp. 2088–2093, Cape Town, South Africa (2014)
30. Palumbo, P., Pizzichelli, G., Panunzi, S., Pepe, P., De Gaetano, A.: Model-based control of plasma glycemia: Tests on populations of virtual patients. Math. Biosci. **257**, 2–10 (2014)
31. Panunzi, S., De Gaetano, A., Mingrone, G.: Advantages of the single delay model for the assessment of insulin sensitivity from the intravenous glucose tolerance test. Theor Biol Med Model. **7**(9) (2010)
32. Panunzi, S., Palumbo, P., De Gaetano, A.: A discrete single delay model for the Intra-Venous Glucose Tolerance Test. Theor. Biol. Med. Model **4**(35) (2007)
33. Parker, R.S., Doyle III, F.J., Ward, J.H., Peppas, N.A.: Robust H_∞ glucose control in diabetes using a physiological model. AIChE J. **46**, 2537–2549 (2000)
34. Pepe, P.: Preservation of the full relative degree for a class of delay systems under sampling. J. Dyn. Syst. Meas. Control-Trans. ASME **125**(2), 267–270 (2003)
35. Puckett, W.R., Lightfoot, E.N.: A model for multiple subcutaneous insulin injections developed from individual diabetic patient data. Am. J. Physiol.-Endocrinol Metab. **269**, E1115–E1124 (1995)
36. RuizVelázquez, E., Femat, R., CamposDelgado, D.U.: Blood glucose control for type I diabetes mellitus: a robust H^∞ tracking problem. Control Eng. Practice **12**, 1179–1195 (2004)
37. Van den Berghe, G.: Insulin therapy for the critically ill patient. Clin. Cornerstone **5**(2), 56–63 (2003)
38. Wild, S., Roglic, G., Green, A., Sicree, R., King, H.: Global prevalence of diabetes—Estimates for the year 2000 and projections for 2030. Diabetes Care **27**(5), 1047–1053 (2004)
39. Wilinska, M.E., Chassin, L.J., Schaller, H.C., Schaupp, L., Pieber, T.R., Hovorka, R.: Insulin kinetics in Type-1 diabetes: continuous and bolus delivery of rapid acting insulin. IEEE Trans. Biomed. Eng. **52**, 3–12 (2005)

Performance Assessment of Model-Based Artificial Pancreas Control Systems

Jianyuan Feng, Kamuran Turksoy and Ali Cinar

Abstract Many artificial pancreas control systems are based on models that predict glucose concentrations. The performance of these control systems depends on the accuracy of the models and may be affected when large dynamic changes in the human body or changes in equipment performance occur and move the operating conditions away from those used in developing the models and designing the control system. A controller performance assessment (CPA) module is developed to evaluate the performance of model-based controllers and initiate controller retuning if there is significant performance deterioration. The generalized predictive control (GPC) approach that utilizes models for glucose concentration predictions is used for illustrating the performance of the CPA. The module has six indexes that capture different aspects of model and controller performance, which can be analyzed to determine the specific component of the controller that caused performance deterioration. Four different kinds of controller errors were diagnosed by indexes and used for controller retuning. Thirty subjects in the UVa/Padova metabolic simulator are used in simulations to evaluate the performance of the CPA module. The results indicate that a GPC with the proposed CPA module has a safer range of glucose concentration variation and more reasonable insulin suggestions than a GPC without controller retuning guided by the CPA module.

1 Introduction

Diabetes affects 25.8 million people in United States and about 5 % of diabetes cases are type 1 diabetes (T1D) in adults [34]. Patients with T1D rely on insulin administration from external sources to regulate their blood glucose concentration (BGC).

J. Feng · A. Cinar (✉)
Department of Chemical and Biological Engineering, Illinois Institute of Technology, Chicago, IL, USA
e-mail: cinar@iit.edu

K. Turksoy · A. Cinar
Department of Biomedical Engineering, Illinois Institute of Technology, Chicago, IL, USA

H. Kirchsteiger et al. (eds.), *Prediction Methods for Blood Glucose Concentration*, Lecture Notes in Bioengineering, DOI 10.1007/978-3-319-25913-0_13

The amount and timing of insulin administration must to be carefully selected to prevent hypoglycemia and hyperglycemia. Multiple (3–5) daily insulin injections and continuous subcutaneous insulin infusion (CSII) with an insulin pump are the most commonly used methods for administering insulin. CSII with pump has many advantages such as better insulin delivery, convenience and flexibility, greater precision, more reliable insulin action, less hypoglycemia with a lower A1C [36].

Artificial pancreas (AP) systems have been introduced to automate and improve CSII. An AP control system can automate insulin pump operation by using glucose concentration (GC) information from a continuous glucose monitor (CGM) to suggest appropriate insulin infusion rates. Significant progress has been made in AP systems [2, 4, 7, 17, 25, 30]. Various control strategies, ranging from proportional-integral-derivative (PID) control [25, 27, 29, 30] to model-based techniques, such as model predictive control [11, 17, 20] and generalized predictive control [8, 9, 33–35] and knowledge-based systems with fuzzy logic [1, 21] have been used in developing the control algorithms for the AP. All strategies except the fuzzy logic approach use mathematical models that describe the dynamics of GC and insulin in formulating the control algorithms. These models are developed by using data collected from subjects. However, many disturbances to the body or faults in the components of the AP system may reduce the accuracy of the models as the operating conditions of the body move to a different state. The effects of disturbances such as meals and physical activity and faults such as failures and drifts in sensors, and pump malfunctions have been mentioned in various studies [3, 18].

Deterioration of the AP performance due to poor representation of the dynamics of the body by the model used in the controller has not received much attention. The control system may also generate faults such as coding errors that may become active under specific conditions. Hence, controller performance assessment (CPA) modules for the AP have the potential to improve its performance by detecting the deterioration of performance and diagnosing the cause of poor performance and initiating controller retuning.

Desborough and Harris [6], Stanfelj et al. [28] initiated work on controller performance monitoring system by using autocorrelation analysis, in which the univariate control system was compared with minimum variance control standard. Huang and Shah [13] and Kendra et al. [16] extended the CPA methods to multivariable control systems. Various CPA systems have been developed based on different techniques such as frequency domain robust control techniques [15], principal component analysis (PCA) [32], Bayesian method [12] and artificial neural networks [19]. 60 % of controllers in industry have performance problems [26]. CPA systems like Loop ScoutTM are put in use in process industries [22]. A large-scale CPA system has been reported spanning over 14,000 PID controllers in 40 plants at 9 sites worldwide [23]. Many other techniques and applications can be found in review papers and books written by Qin [24], Jelali [14], Thornhill [31] and Cinar [5].

We have already developed an AP control system based on adaptive constrained weighted recursive identification methods and GPC [34, 35]. This paper focuses on CPA, detection of controller performance deterioration, diagnosis of its cause(s) and retuning of a GPC to improve its performance. Various indexes are developed to detect

and diagnose significant deterioration of controller performance and develop a CPA system. The parameters of the controller are adjusted to improve its performance. The performance of the CPA system and the effects of control parameters adjustment are illustrated with simulations by using the UVa/Padova simulator [17].

In Sect. 2, GPC and five indexes and four different controller errors are defined. In Sect. 3, retuning of the controller in response to controller faults is described. Simulations and discussion of results are presented in Sect. 4. Conclusions are provided in Sect. 5.

2 GPC and Controller Error Detection

2.1 GPC in AP System

GPC integrates model-based control and adaptive control. We have enhanced recursive time series modeling to assure the stability of every multi-input single-output model developed, used these models in a powerful GPC, and introduced rules that improve its performance in presence of physical activity [34, 35]. The controller algorithm is described in Fig. 1.

A single-variable version of this controller is used to conduct simulations with the UVa/Padova simulator. An autoregressive moving average model with exogenous inputs (ARMAX) is used for prediction of GC measured with a CGM [34]. The insulin infusion rates computed in previous steps and past and current CGM readings are the inputs to the model. Model parameters are updated recursively at each sampling time and the updated model provides GC prediction to the controller. Then, the controller computes the insulin dose to be infused.

In this chapter, we outline the components of the controller that are relevant to the CPA and focus on the objective function for model estimation and the objective function of the GPC.

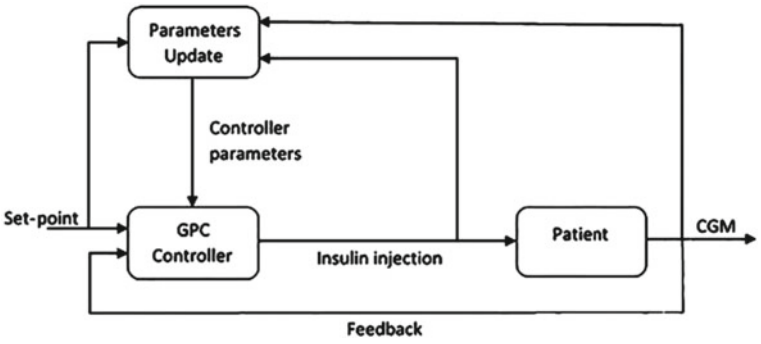

Fig. 1 AP system with GPC algorithm

The optimum coefficients of the recursive ARMAX model are obtained by minimizing the objective function:

$$V(\hat{\Theta}) = \sum_{k=1}^{N} \lambda^{N-k} e(k)^2 \tag{1}$$

where $\hat{\Theta}$ is the vector of estimated model coefficients, N is the number of samples, λ is the forgetting factor, and e denotes the modeling error (residual). The recursive model is updated by giving different weights to previous data based on the section of the data that is more important to adjust the model. λ is the critical parameter that adjusts the relative importance of GC data collected and the effective size of the moving window for the input data.

Another important component for the controller is its objective function. The constrained controller signal $I(k)$ is calculated by minimizing the objective function $J(N_1, N_2, N_u, w)$:

$$\begin{aligned} J(N_1, N_2, N_u, w) = \min_{I(k),\ldots,I(k+N_u)} & \sum_{j=N_1}^{N_2} \left[GC^{pred}(k+j|k) - GC^{ref}(k+j)\right]^2 \\ & + \sum_{j=1}^{N_u} w_j \Delta I(k+j-1)^2 \end{aligned} \tag{2}$$

where N_1 and N_2 are the first and last time instants of the modeling horizon and N_u is the control horizon. I denotes the insulin infusion rate. Only $I(k)$, the first element of the optimized future I sequence is implemented. $GC^{pred}((k+j|k))$ denotes the predicted CGM values at steps $k+j$ for $j = N_1, N_2$ in the future, based on the data collected at and before step k, GC^{ref} is the reference vector (GC^{ref} is set as a reference trajectory instead of a constant value to provide smoother insulin suggestions that reduce the potential for hypoglycemia) [34], ΔI represents the change in insulin infusion rate, and w_j is the weight vector for ΔI. Insulin delivery is constrained by both the maximum delivery rate of the insulin pump and the maximum value calculated by the model to prevent hypoglycemia, and insulin infusion rate cannot be negative. The first term of the controller objective function forces future GC to be closer to reference values, and the second term seeks to minimize the amount of insulin infused to improve the efficiency of insulin use.

2.2 Indexes Used for CPA

Six indexes are used in the controller assessment module: model prediction error index (I_{MPE}), model error elimination speed index (I_{EES}), dangerous change potential index (I_{DCP}), dangerous change index (I_{DC}), insulin constraints limitation index (I_{ICL}), and weight ratio index (I_{WR}).

2.2.1 Model Prediction Error Index

I_{MPE} indicates the absolute difference between the predicted CGM value $GC^{pred}(k|k-1)$ and the measured CGM value $GC(k)$.

$$I_{MPE}(k) = |GC^{pred}(k|k-1) - GC(k)| \tag{3}$$

where $GC^{pred}(k|k-1)$ indicates predicted GC at time k based on previous data at time $k-1$. If there is a large prediction error, the time for AP system to reduce it back to an acceptable range becomes critical.

2.2.2 Error Elimination Speed Index

I_{EES} is defined to track the time to reduce I_{MPE} to an acceptable range. If the prediction error is larger than ME^{max} we start counting the time until I_{MPE} is smaller than ME^{max}.

$$I_{EES}(k) = \begin{cases} I_{EES}(k-1) + st & \text{if } I_{MPE}(k) \geq ME^{max} \\ 0 & \text{if } I_{MPE}(k) < ME^{max} \end{cases} \tag{4}$$

where st is the sampling time that is set as 10 min in our simulator. ME^{max} (Table 3) is the maximum tolerance threshold for the difference between the predicted and measured GC.

2.2.3 Dangerous Change Potential Index and Dangerous Change Index

The main goal for AP system is to keep a patient's BGC within an acceptable range. A disturbance (meals, exercise, stress) or a fault in the AP system can force BGC to move outside this range and cause hypoglycemia or hyperglycemia. In order to signal the potential for hypo or hyperglycemia while glucose concentration GC (CGM readings) is within an acceptable range (70–180 mg/dl), dangerous change potential index (I_{DCP}) is developed:

$$I_{DCP}(k) = \begin{cases} \frac{180-GC(k)}{GC(k)-GC(k-1)} & \text{if } \begin{cases} \Delta GC(k) > 0 \\ 70 < GC(k) < 180 \end{cases} \\ \frac{GC(k)-70}{GC(k-1)-GC(k)} & \text{if } \begin{cases} \Delta GC(k) < 0 \\ 70 < GC(k) < 180 \end{cases} \end{cases} \tag{5}$$

where $\Delta GC(k) = GC(k) - GC(k-1)$. To distinguish between hypoglycemia and hyperglycemia, I_{DC} is defined as

$$I_{DC}(k) = \begin{cases} 1 & \text{if } (I_{DCP}(k) < I_{DCP}^{max} \,\&\, \Delta GC(k) > 0) \text{ or } (GC(k) \geq 180) \\ -1 & \text{if } (I_{DCP}(k) < I_{DCP}^{max} \,\&\, \Delta GC(k) < 0) \text{ or } (GC(k) \leq 70) \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

I_{DCP} signals how many time steps later GC will indicate hypoglycemia or hyperglycemia if the change in GC is maintained at the current rate, we set I_{DCP}^{max} (Table 3) as the tolerance threshold for I_{DCP}. Note that there is no value for I_{DCP} if $GC > 180$ or $GC < 70$ because I_{DCP} only shows the potential of GC to get outside the desired range.

2.2.4 Insulin Constraints Limitation Index

The controller computes an insulin dose based on the insulin on board and trends in CGM readings, constrained to calculate a safe insulin dose. But since the model may not be accurate enough at any given time, the constraints may become too strict for the controller to suggest a large enough insulin dose to lower BGC. I_{ICL} is developed to indicate whether the insulin suggestion is limited by unreasonable insulin constraints or a maximum that is medically imposed. Larger values of I_{ICL} indicate that the insulin dose computed is remaining at the maximum insulin constraints.

$$I_{ICL}(k) = \begin{cases} I_{ICL}(k-1) + st & \text{if } I(k) = I_{Max} \\ 0 & \text{if } I(k) < I_{Max} \end{cases} \tag{7}$$

2.2.5 Weight Ratio Index

The objective function of GPC is one of the key factors affecting controller performance. The weight w in the objective function (2) indicates the relative importance of its two terms in determining insulin infusion rates. If a patient is drifting to hypoglycemia or hyperglycemia, the differences between CGM predictions and reference values are much more important than reducing the amount of insulin used. I_{WR} is developed to monitor the weight ratio of the objective function between deviations from reference trajectories or ranges and insulin dose to be administrated.

$$I_{WR} = \frac{\sum_{j=N_1}^{N_2} \left[GC^{pred}(k+j|k) - GC^{ref}(k+j)\right]^2}{\sum_{j=1}^{N_u} w_j \Delta I(k+j-1)^2} \tag{8}$$

I_{WR} will become infinity if the change in insulin infusion rate or w_j become close to 0, in these case, a large enough value (e.g., 1000) is set as the maximum value.

There are many sources of error such as model inaccuracy, improper insulin constraints, and improper objective function. Binary indicators are defined to flag some of these faults or deteriorated performance in the next section.

2.3 Detection and Diagnosis of Controller Errors

Many errors can develop in the control system that would necessitate controller retuning. The CPA module currently focuses on four different kinds of errors in GPC for AP systems: model prediction error, insulin constraints error, GC deviation-insulin weight ratio error, and sensor noise-driven miscalculation error. Binary numbers 1 and 0 are used to indicate the presence or absence of such errors.

2.3.1 Model Prediction Error

CGM readings and insulin information at each new sampling time are used to recursively update the parameters in the model. Since both model and data are varying all the time, model prediction error is also changing with time. The recursive time series model can gradually eliminate model prediction errors caused by dynamic changes of the system by updating model parameters whenever new data are available. But if there are large dynamic changes and the weights for recent data that correspond to model update speed are not adequate, the model will not be accurate enough to give good predictions. If the model prediction error is too large ($I_{MPE} > I_{MPE}^{max}$) or cannot be eliminated in time ($I_{EES} \geq I_{EES}^{max}$), the controller can suggest an unreasonable insulin infusion rate by using erroneous GC predictions, and threaten the safety of patients. Thresholds I_{MPE}^{max} and I_{EES}^{max} are described in Table 3. A prediction error indicator (I_{PE}) that is a function of I_{MPE} and I_{EES} indicates the presence of prediction errors

$$I_{PE} = \begin{cases} 1 & \text{if } I_{MPE} > I_{MPE}^{max} \text{ or } I_{EES} \geq I_{EES}^{max} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

2.3.2 Maximum Insulin Constraint Error

When the insulin dose suggested is at the maximum insulin constraint value for a long time ($I_{ICL} \geq I_{ICL}^{max}$) and the CGM readings still keep on increasing rapidly, the insulin dose constraint may be too conservative. Threshold I_{ICL}^{max} is described in Table 3. Insulin constraints error indicator (I_{ICE}) is used to flag a problem with insulin constraints based on I_{DC} and I_{ICL} values.

$$I_{ICE} = \begin{cases} 1 & \text{if } \begin{cases} I_{DC} = 1 \\ I_{ICL} \geq I_{ICL}^{max} \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

2.3.3 Weight Ratio Error

When patients are drifting toward or already have hypoglycemia or hyperglycemia, the controller should give a high priority to bring GC within range and deemphasize other concerns. In the GPC, the weight ratio in the objective function implements this focus adjustment. If the weights in the optimization function necessitate a heavier weight on GC safety (large GC deviation), an objective function weight ratio error indicator (I_{WRE}) that is a function of I_{WR} and I_{DC} is defined to indicate inappropriate weights:

$$I_{WRE} = \begin{cases} 1 & \text{if} \begin{cases} I_{WR} < 1 \\ I_{DC} \neq 1 \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

2.3.4 Sensor Noise-Driven Miscalculation Error

Sensor noise is inevitable in CGM readings, and large sensor noise can temporarily change the trends in CGM readings (Fig. 2). In realistic situations, GC values should be smoother than these cases. Since the trend of GC and effects of insulin parameters in model are estimated based on the GC and insulin infusion values, erroneous GC values caused by sensor noise can cause the controller to suggest erroneous insulin infusion doses by miscalculating these model parameters (Table 1).

Consider the sampling time is 10 min and dynamic changes of BGC related to disturbances such as meal or exercise are observed for a 30-min period. There should not be two inflection points within 4–5 samples (three different signs of CGM slope) in CGM trajectory. Figure 2 illustrates four typical patterns of noisy CGM readings.

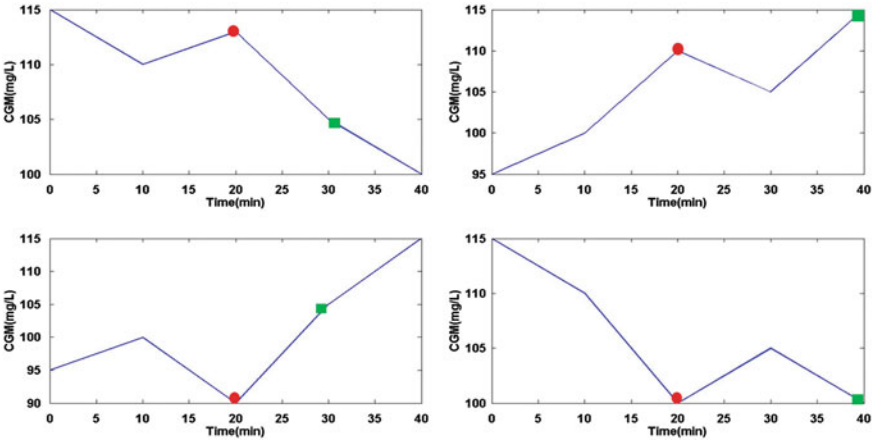

Fig. 2 Patterns of abnormal readings in CGM caused by sensor noise

Table 1 Summary of all indexes and indicators in CPA

Indexes and indicators	Definition	Equation number for definition
I_{MPE}	Model prediction error index	3
I_{EES}	Error elimination speed index	4
I_{DCP}	Dangerous change potential index	5
I_{DC}	Dangerous change index	6
I_{ICL}	Insulin constraints limitation index	7
I_{WR}	Weight ratio index	8
I_{PE}	Prediction error indicator	9
I_{ICE}	Insulin constraints error indicator	10
I_{WRE}	Weight ratio error indicator	11
I_{SME}	Sensor-noise-driven miscalculation error indicator	12
I_{SMEA}	Auxiliary SME indicator	13

On the left side of Fig. 2, the two points (green square) where sensor noise and miscalculation error (SME) is detected are at the next sampling time to the abnormal CGM reading (red circle). On the right side, there is one additional sampling time between abnormal CGM reading and the detection time. All these four types of abnormal data patterns have two inflection points. Consequently, SME can be detected by SME indicator (I_{SME}) with the following expressions:

$$I_{SME} = \begin{cases} 1 & \text{if } \begin{cases} \Delta GC(k)\Delta GC(k-1) < 0 \\ \Delta GC(k-1)\Delta GC(k-2) < 0 \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

For the controller retuning, it is important to know which CGM value is erroneous. An auxiliary binary indicator I_{SMEA} is defined to distinguish between the two cases (fault detected in the first or second sampling time after SME) as:

$$I_{SMEA} = \begin{cases} 1 & \text{if } (\Delta GC(k) - \Delta GC(k-2))\Delta GC(k) > 0 \\ 0 & \text{if } (\Delta GC(k) - \Delta GC(k-2))\Delta GC(k) < 0 \end{cases} \tag{13}$$

when $I_{SMEA} = 1$, it indicates fault is detected in the first sampling time after SME, and $I_{SMEA} = 0$ if fault is detected in the second sampling time after SME.

CGM accuracy and reliability are among the important limiting factors for insulin delivery automation [18], mainly due to the fact that glucose is measured in the interstitial region rather than directly from the blood. Different sensor errors like outliers, drifts, and missing values affect the reliability of the AP system as well. CGM sensor fault detection and analysis is still an active research area. The method proposed in here is focused on detecting single discontinuous noise to illustrate

incorporation of sensor errors in data on controller performance. Since this chapter is mainly focused on the errors from controller itself. More sophisticated sensor fault detection strategies are beyond the scope of this paper.

3 Controller Retuning

In order to retune the controller, the specific reasons that cause poor performance, and the consequences of various errors that cause poor performance should be considered. This interaction of detection, diagnosis, and adjustment is illustrated with a number of frequent errors. Guidance for retuning the controller in response to four kinds of error is described in this section.

3.1 Controller Retuning for Model Prediction Error

Since there are large dynamic changes in GC caused by various disturbances such as meals, a fixed model is not appropriate for computing the insulin infusion rates. The need for an adaptive model becomes more evident when no manual announcements are made. The parameters of a recursive ARMAX model in GPC are updated at every sampling time. The accuracy of model prediction depends on the selected weights for the inputs used in updating the model parameters. New data after large dynamic changes in GC cannot cause appropriate changes in the model if the weight for recent data is not large enough. Otherwise, most of the data representing the previous state of the body would continue to influence the model parameter values computed. A forgetting factor λ is used to adjust the weights for the inputs. The range of λ is from 1 to 0. When $\lambda = 1$, same weight is given to all of the data (infinite window size), when $\lambda < 1$, most of the weight is given to recent data. The window size N can be estimated by [10]:

$$N \approx \frac{1}{1-\lambda} \tag{14}$$

Obviously, when large dynamic changes occur, smaller value of λ are used to give more weight to recent data so that the model parameters can be adjusted faster toward the new state of the body.

Larger values of MPE are caused by the model not recognizing the changes in the dynamics and adjusting the parameters fast enough to represent the new changes in GC dynamics. In this case, λ is modified to λ' if the prediction error is large or the error reduction speed is too slow:

$$\lambda'(k|k-1) = \begin{cases} \left(1 - \frac{I_{MPE}}{den}\right)\lambda(k|k-1) & \text{if } \begin{cases} I_{PE} = 1 \\ I_{MPE} < I_{MPE}^{SE} \\ \lambda'(k|k-1) > 0.5 \end{cases} \\ 0.5 & \text{if } \begin{cases} I_{PE} = 1 \\ I_{MPE} > I_{MPE}^{SE} \text{ or } \lambda'(k|k-1) < 0.5 \end{cases} \end{cases} \tag{15}$$

λ' is the modified forgetting factor, *den* is a fixed parameter (60 for adults or adolescents and 90 for children), and I_{MPE}^{SE} is the threshold to determine extreme prediction error (Table 3). According to (1), when $\lambda < 0.5$, put most of the weight only on latest data will make the model become too sensitive to data changes. Errors in prediction may happen if the value of λ is too small. When prediction error is detected, the minimum value for λ is set as 0.5 to reduce the possibility of overturning.

Model prediction error can be reduced further by using a filter such as a Kalman filter. A simpler approach is to use an adjustment with a proportional feedback constant β (Table 3), so that MPE can be eliminated faster:

$$GC^{pred'}(k+1|k) = GC^{pred}(k+1|k) - \left(GC^{pred}(k|k-1) - GC(k)\right)\beta \quad \text{if } I_{PE} = 1 \tag{16}$$

If the predicted CGM reading $GC^{pred}(k|k-1)$ is higher than the actual value $GC(k)$, the controller will recommend a higher insulin to balance the miscalculated future GC. If the prediction is lower than actual value, controller will recommend less insulin than necessary. As a solution, Insulin infusion rate is adjusted based on the difference between predicted and measured glucose concentration:

$$I'(k) = \begin{cases} I(k)\left(1 + \frac{\log_{10}(GC(k) - GC^{pred}(k|k-1))BW}{\alpha}\right)ISF \\ \qquad \text{if } \begin{cases} GC(k) - GC^{pred}(k|k-1) > 0 \\ I_{PE} = 1 \\ I_{DC} \neq -1 \end{cases} \\ I(k)\left(1 - \frac{\log_{10}(GC^{pred}(k|k-1) - GC(k))BW}{\alpha}\right)ISF \\ \qquad \text{if } \begin{cases} GC(k) - GC^{pred}(k|k-1) < 0 \\ I_{PE} = 1 \\ I_{DC} \neq 1 \end{cases} \end{cases} \tag{17}$$

where log function is used since it is more sensitive when prediction error is considerably small and prohibit overtuning when prediction error is too large. α is a fixed parameter (Table 3), *BW* is the body weight of the patient and *ISF* is the insulin sensitivity factor which is related to a patient's insulin sensitivity level (ISL) (Table 2).

Table 2 Value for parameter ISF

	Adult	Adolescent	Child
Sensitive	0.8	0.8	0.4
Normal	1	1	0.5
Resistant	1.2	1.2	0.6

Table 3 Summary of parameters and their values in CPA

Parameters	Definition	Equation number	Value
ME^{max}	Threshold of I_{MPE} to determine I_{EES}	4	10
$I_{DCP}^{\max}$	Threshold of I_{DCP} to determine I_{DC}	6	4
$I_{MPE}^{\max}$	Threshold of I_{MPE} to determine I_{PE}	9	20
$I_{EES}^{\max}$	Threshold of I_{EES} to determine I_{PE}	9	30
$I_{ICL}^{\max}$	Threshold of I_{ICL} to determine I_{ICE}	10	40
I_{MPE}^{SE}	Threshold of I_{MPE} to tuning λ'	15	30
α	Parameter indicating the sensitivity of insulin in responding to *MPE* and *SME*	17, 21	270
β	Proportional feedback constant for $GC^{pred'}$	16	0.35
γ	Filter constant determines how fast the insulin constraint is increasing	18	1.1

3.2 Controller Retuning for Insulin Dose Constraint Error

For insulin constraints error (ICE), maximum constraint of insulin suggestion is relaxed so that controller can give adequate insulin. However, the maximum constraint is also limited by the capacity of the insulin pump. For example, the maximum insulin delivery capacity for Medtronic pump is 35 U/h. Considering such physical limitations, the insulin constraint is modified as:

$$I'_{Max} = \begin{cases} \gamma I_{Max} & \text{if } I'_{Max} < 35(\text{U/h}) \ \& \ I_{ICE} = 1 \\ 35(\text{U/h}) & \text{if } I'_{Max} \geq 35(\text{U/h}) \ \& \ I_{ICE} = 1 \end{cases} \tag{18}$$

where γ (Table 3) is a filter constant that determines how fast the insulin constraint is increasing when ICE occurs.

3.3 Controller Retuning for Objective Function Weight Ratio Error

When a patient's GC has a critical trend of rapid change ($I_{DC} = 1$ or $I_{DC} = -1$), the controller should calculate the insulin suggestion mainly to guarantee the safety of the patient, focusing on the difference between the GC predictions and reference

trajectory. In processing of controller retuning for WRE, the weight vector w_j should be reduced and in the limit only the difference between predicted CGM values and reference value is considered in the objective function, which reduces the optimization function to:

$$J(N_1, N_2) = \min_{I(k)} \sum_{j=N_1}^{N_2} \left[GC^{pred}(k+j|k) - GC^{ref}(k+j) \right]^2 \quad \text{if} \quad I_{WRE} = 1 \tag{19}$$

3.4 Controller Retuning for Sensor-Noise-Driven Miscalculation Error

Similar to MPE, excessive CGM noise can make the controller suggest erroneous insulin doses to the pump. The insulin dose suggested by the controller can be further modified based on SCF values:

$$SCF(k) = \begin{cases} GC(k-1) - \frac{GC(k)+GC(k-2)}{2} & \text{if} \begin{cases} I_{SME} = 1 \\ I_{SMEA} = 1 \end{cases} \\ GC(k-2) - \frac{GC(k-3)+GC(k-1)}{2} & \text{if} \begin{cases} I_{SME} = 1 \\ I_{SMEA} = 0 \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

$$I'(k) = \begin{cases} I(k)\left(1 + \frac{\log_{10}(GC(k) - GC^{pred}(k|k-1) + SCF(k))BW}{\alpha}\right) ISF \\ \qquad \text{if} \begin{cases} GC(k) - GC^{pred}(k|k-1) + SCF > 1 \\ I_{DC} \neq -1 \\ I_{SME} = 1 \text{ or } I_{PE} = 1 \end{cases} \\ I(k)\left(1 - \frac{\log_{10}(GC^{pred}(k|k-1) - GC(k) - SCF)BW}{\alpha}\right) ISF \\ \qquad \text{if} \begin{cases} GC(k) - GC^{pred}(k|k-1) + SCF < -1 \\ I_{DC} \neq -1 \\ I_{SME} = 1 \text{ or } I_{PE} = 1 \end{cases} \end{cases} \tag{21}$$

SCF is the correcting factor of *SME*, which uses the average value of two nearest sets of data to estimate the real value and *SCF* is the difference between the estimated value and the abnormal value caused by sensor noise. The parameter α indicates the sensitivity of insulin in responding to *MPE* and *SME*. In (12), *SME* is detected if there are continuously two inflection points. If excessive sensor noise is detected at time k, (20) and (21) can be used to distinguish and retune the controller depending on the specific noise pattern matching the cases in Fig. 2. Note in (15), the retuning is triggered only when $I_{PE} = 1$ to guarantee an appropriate value in the argument of

the log function. However, if we consider both *MPE* and *SCF*, the effects of these two errors may interfere with each other. In this case, a new if condition is added to (21) to make sure the argument of log function is satisfied.

The thresholds for fault detection were selected experimentally by trying different values and selecting values that minimize false alarms (there is no extreme cases like hypoglycemia and hyperglycemia after fault is detected) and missed alarms (there exist extreme cases but no fault is detected). The parameters for fault detection were selected by trying different values and taking the best values that minimize the extreme cases (hypoglycemia and hyperglycemia). For example, γ values tried include [1.05, 1.1, 1.15, 1.2], and the best outcome was obtained when $\gamma = 1.1$. If γ is larger or smaller than 1.1, either insulin infusion rate becomes too aggressive and may cause cause hypoglycemia ($\gamma > 1.1$), or not enough increase of maximum insulin constraint is made which will cause hyperglycemia ($\gamma < 1.1$).

The predesigned parameters and thresholds are critical for the CPA performance. Thresholds for indexes decide how sensitive CPA responds to controller errors. The goal for CPA is not to make the controller perfectly regulate GC but to keep the controller at its optimal condition. If the threshold is too aggressive, the controller may always have a weak performance since the high standard is seldom reached, on the other hand if the threshold is too conservative, the weak performance of controllers cannot be captured. The tuning parameter in CPA module decide how can the controller compensates for these errors. Retuning process is just like another feedback controller, if controller errors are determined, the CPA makes changes in the opposite direction that the errors would drive the system in order to compensate for the effects of the diagnosed error. Most of the parameters and thresholds are often decided by trial and error and experience. It is important to adjust these parameters and thresholds for different AP controllers.

4 Results

Thirty patients (10 adults, 10 adolescents, and 10 children) in UVa/Padova simulator were tested by using both AP control systems with and without the CPA module. Each patient tested though 10 times of simulations with 3 days' scenario each time. The results are summarized in Tables 4 and 5. All CGM readings are clustered into five different ranges. AP control system with CPA module in most cases has a higher percentage of GC inside the desired GC range (70–180 mg/dl). The controller with the CPA module successfully modified its parameters to reduce serious hypo- and hyperglycemia (CGM > 300 or CGM < 50) (Table 5). To keep GC in target range, GPC with CPA module had better insulin efficiency. For simulations lasting 3 days, the total insulin use is reduced by about 8 % on average.

The effects of CPA module on GC are illustrated in Figs. 3 and 4. Figure 3 displays the average GC trajectories and standard deviations for adults, adolescents and

Table 4 Comparison of performance between controller with and without CPA module (percent time in each concentration range)

Patient type	ISL	BW (kg)	GC with CPA module (mg/dl)					GC without CPA module (mg/dl)				
			<50	50–70	70–180	180–300	>300	<50	50–70	70–180	180–300	>300
Ado 1	N	68.7	0.00	0.00	100.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00
Ado 2	R	51.0	0.00	0.83	82.56	16.60	0.00	0.00	0.32	82.10	17.47	0.11
Ado 3	S	44.8	0.00	0.21	90.60	9.19	0.00	0.00	0.08	87.65	12.27	0.00
Ado 4	N	49.6	0.40	1.13	90.33	8.15	0.00	0.43	0.63	89.00	9.93	0.00
Ado 5	S	47.1	0.00	1.00	76.91	19.59	2.50	0.00	0.00	75.50	20.31	4.20
Ado 6	N	45.4	0.00	0.09	87.44	12.41	0.06	0.00	0.00	86.39	13.41	0.19
Ado 7	N	37.9	0.44	2.06	74.42	19.03	4.05	1.18	2.11	72.59	17.81	6.31
Ado 8	R	41.2	0.00	0.17	78.60	19.82	1.40	0.00	0.00	78.36	19.91	1.72
Ado 9	S	43.9	0.00	0.11	88.80	11.10	0.00	0.00	0.06	87.85	12.09	0.00
Ado 10	N	47.4	0.00	1.24	92.20	6.55	0.00	0.00	0.66	91.62	7.71	0.00
Mean		*47.7*	*0.08*	*0.68*	*86.19*	*12.24*	*0.80*	*0.16*	*0.39*	*85.11*	*13.09*	*1.25*
Adu 1	N	102.3	0.00	0.74	98.91	0.35	0.00	0.00	0.01	99.23	0.75	0.00
Adu 2	N	111.1	0.00	0.00	100.00	0.00	0.00	0.00	0.00	99.88	0.12	0.00
Adu 3	N	81.6	0.00	0.68	98.65	0.67	0.00	0.00	0.00	99.16	0.84	0.00
Adu 4	S	63.0	0.47	2.07	85.23	12.20	0.03	0.81	3.64	82.80	12.57	0.19
Adu 5	R	94.1	0.00	0.03	96.48	3.49	0.00	0.00	0.00	96.34	3.66	0.00
Adu 6	N	66.1	0.00	1.01	91.28	7.71	0.00	0.00	0.00	90.25	9.75	0.00
Adu 7	S	91.2	0.26	0.88	94.20	4.66	0.00	0.00	0.00	82.28	17.72	0.00
Adu 8	N	102.8	0.00	0.00	100.00	0.00	0.00	0.00	0.16	99.84	0.00	0.00
Adu 9	R	74.6	1.53	3.88	93.07	1.52	0.00	3.03	11.11	84.21	1.65	0.00
Adu 10	R	73.9	0.00	0.06	98.34	1.60	0.00	0.10	2.03	92.46	5.41	0.00
Mean		*86.1*	*0.23*	*0.93*	*95.62*	*3.22*	*0.00*	*0.39*	*1.69*	*92.65*	*5.25*	*0.02*

(continued)

Table 4 (continued)

Patient type	ISL	BW (kg)	GC with CPA module (mg/dl)					GC without CPA module (mg/dl)				
			<50	50–70	70–180	180–300	>300	<50	50–70	70–180	180–300	>300
Chi 1	S	34.6	12.71	12.90	73.18	1.15	0.06	18.89	13.13	62.08	5.64	0.27
Chi 2	R	28.5	0.00	0.23	99.53	0.24	0.00	0.00	0.00	99.33	0.67	0.00
Chi 3	S	41.2	0.00	1.46	88.30	10.24	0.00	0.00	0.56	87.04	12.27	0.13
Chi 4	S	35.5	5.23	10.55	83.11	2.11	0.00	8.73	15.09	71.78	4.40	0.00
Chi 5	R	37.8	0.00	1.56	97.22	1.22	0.00	0.00	0.20	97.75	2.05	0.00
Chi 6	N	41.0	0.00	1.11	91.19	7.70	0.00	0.15	0.67	88.12	11.02	0.03
Chi 7	R	45.5	0.00	0.75	98.93	0.32	0.00	0.00	0.19	99.35	0.46	0.00
Chi 8	R	23.7	10.78	2.55	86.96	8.72	1.00	19.35	2.23	62.38	14.06	1.98
Chi 9	N	35.5	0.00	1.32	95.47	4.53	0.00	0.00	0.13	94.57	5.31	0.00
Chi 10	N	35.2	0.00	0.98	84.42	14.41	0.19	0.00	0.00	81.40	17.96	0.64
Mean		*35.9*	*2.87*	*3.34*	*89.83*	*5.06*	*0.12*	*4.71*	*3.22*	*84.38*	*7.38*	*0.31*
Total		*56.6*	*1.06*	*1.65*	*90.54*	*6.84*	*0.31*	*1.76*	*1.77*	*87.38*	*8.57*	*0.53*

Table 5 Comparison of performance between controller with and without CPA module

Patient type	GC with CPA module					GC without CPA module				
	Number of severe hyperglycemia episodes (>300 mg/dl)	Number of severe hyperglycemia episodes (<50 mg/dl)	Max GC (mg/dl)	Min GC (mg/dl)	Total insulin use (U/3 days)	Number of severe hyperglycemia episodes (>300 mg/dl)	Number of severe hyperglycemia episodes (<50 mg/dl)	Max GC (mg/dl)	Min GC (mg/dl)	Total insulin use (U/3 days)
Ado 1	0	0	179.77	70.99	104.00	0	0	179.63	93.45	101.52
Ado 2	0	0	286.86	66.37	210.42	1	0	323.96	64.27	221.16
Ado 3	0	0	258.26	62.06	65.93	0	0	258.10	66.69	70.32
Ado 4	0	1	248.94	46.94	106.07	0	2	249.27	42.98	110.08
Ado 5	27	0	350.71	62.67	94.80	42	0	360.60	85.16	100.37
Ado 6	1	0	310.39	65.05	122.83	3	0	338.82	71.19	133.25
Ado 7	51	2	89.91	41.89	138.49	56	5	458.78	25.37	149.18
Ado 8	16	0	337.76	68.10	194.15	16	0	351.12	71.92	202.06
Ado 9	0	0	242.73	69.29	67.93	0	0	251.76	67.97	68.42
Ado 10	1	0	254.66	64.50	96.07	0	0	267.32	59.12	100.82
Mean	*9.60*	*0.30*	*286.00*	*61.79*	*120.07*	*11.80*	*0.70*	*303.94*	*64.81*	*125.72*
Adu 1	0	0	236.73	62.41	154.90	0	0	240.34	69.74	155.26
Adu 2	0	0	178.41	76.25	163.55	0	0	208.32	72.04	164.03
Adu 3	0	0	210.81	62.66	168.65	0	0	214.08	70.35	169.21
Adu 4	0	3	304.46	46.42	100.71	5	5	310.28	37.41	115.24
Adu 5	0	0	255.71	68.86	197.93	0	0	260.08	74.64	201.66
Adu 6	0	0	231.24	52.77	182.11	0	0	239.14	73.85	176.93
Adu 7	0	1	266.39	49.32	121.41	0	0	261.47	74.34	168.12
Adu 8	0	0	174.76	70.07	126.78	0	0	187.63	60.92	127.28
Adu 9	0	4	220.11	47.21	141.88	0	23	240.87	29.72	148.91
Adu 10	0	0	225.19	69.04	199.39	0	1	271.69	42.54	199.63
Mean	*0.00*	*0.80*	*230.38*	*60.50*	*155.73*	*0.50*	*2.90*	*243.39*	*60.55*	*162.63*

(continued)

Table 5 (continued)

Patient type	GC with CPA module					GC without CPA module				
	Number of severe hyperglycemia episodes (>300 mg/dl)	Number of severe hyperglycemia episodes (<50 mg/dl)	Max GC (mg/dl)	Min GC (mg/dl)	Total insulin use (U/3 days)	Number of severe hyperglycemia episodes (>300 mg/dl)	Number of severe hyperglycemia episodes (<50 mg/dl)	Max GC (mg/dl)	Min GC (mg/dl)	Total insulin use (U/3 days)
Chi 1	1	22	317.85	41.64	12.18	7	63	319.87	1.13	59.24
Chi 2	0	0	208.14	68.04	37.95	0	0	213.59	77.80	40.50
Chi 3	0	0	287.52	55.47	32.46	3	0	318.18	54.17	43.15
Chi 4	0	31	246.67	31.45	23.97	0	49	254.70	15.07	54.53
Chi 5	0	0	229.58	60.15	52.06	0	0	239.16	62.06	58.22
Chi 6	0	0	280.83	63.13	42.45	1	2	304.93	46.97	51.95
Chi 7	0	0	213.41	59.84	53.80	0	0	216.01	53.61	53.84
Chi 8	9	8	372.44	31.46	32.29	25	12	373.33	5.40	58.55
Chi 9	0	0	228.23	64.15	32.52	0	0	243.08	61.80	35.90
Chi 10	4	0	308.39	66.02	39.43	10	0	322.62	83.71	49.09
Mean	*1.40*	*6.10*	*269.30*	*54.13*	*35.91*	*4.60*	*12.60*	*280.55*	*46.17*	*50.50*
Total	*3.67*	*2.40*	*261.89*	*58.81*	*103.90*	*5.63*	*5.40*	*275.96*	*57.18*	*112.95*

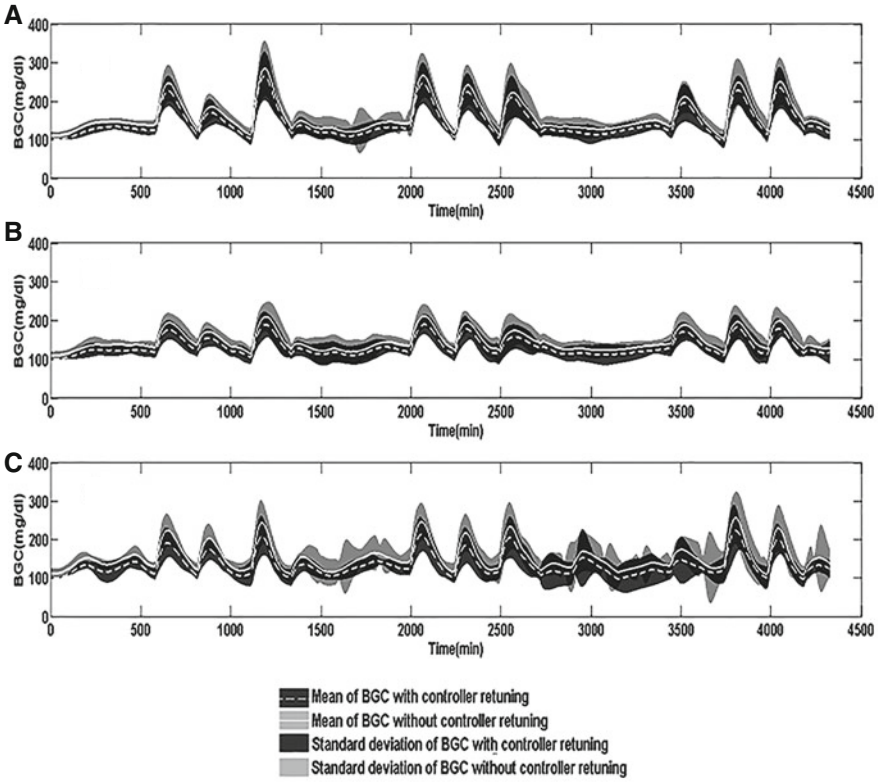

Fig. 3 Average GC trajectories and standard deviations for all patients (**A** Adolescents, **B** Adults, **C** Children)

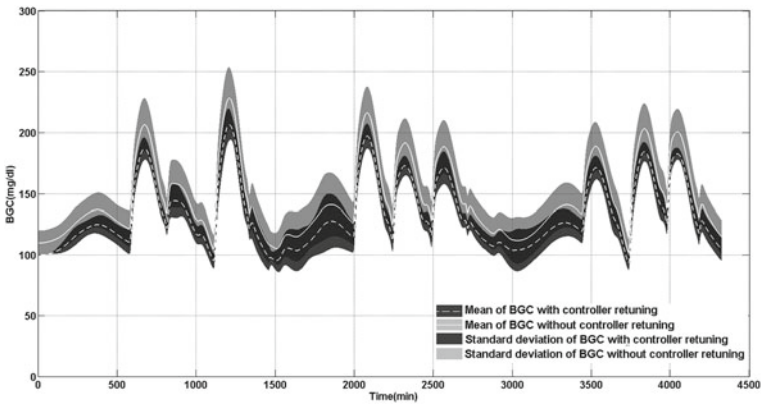

Fig. 4 Average GC trajectories and standard deviations for adult #10

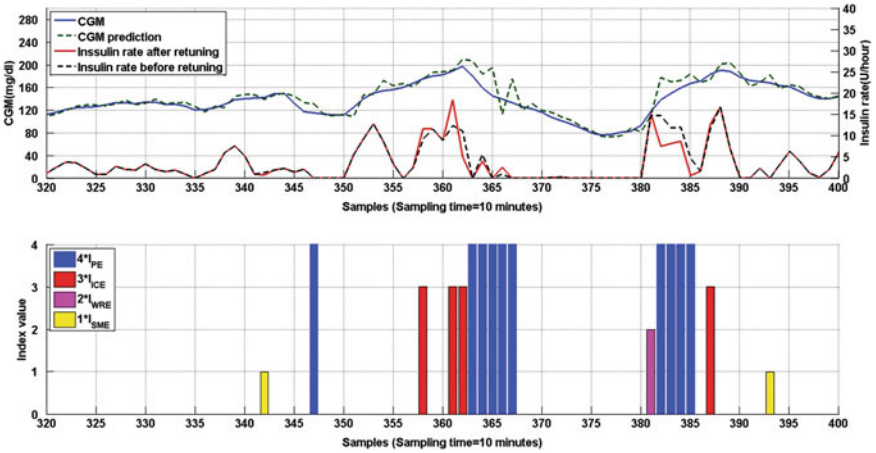

Fig. 5 CPA performance and effects of controller adjustment during part of the simulation of adult #10

children in the UVa/Padova simulator. Figure 4 illustrates the GC values for one adult subject. The variations in GC are reduced and the maximum values at hyperglycemia are reduced significantly. Hypoglycemia and the drops in GC are reduced as well. The effects are less pronounced since the original GPC algorithm was designed to be aggressive in minimizing hypoglycemia. Figure 5 illustrates the performance of CPA during part of a simulation of adult #10. CGM data, GC prediction, insulin rate before and after retuning are displayed on the top frame of Fig. 5. Four different kinds of controller errors are displayed by their indicators in the lower frame. Figures 3 and 4 illustrate how the presence of the CPA module improves the control of GC with a GPC system. The mean GC trajectories are lower and within a narrower range when CPA is used (dashed line) and the standard deviation of the trajectories are smaller (dark gray band).

If no controller error is detected in Fig. 5, the insulin infusion rate remain the same, with or without CPA. Otherwise, different detections of CPA and retuning are displayed. For example, at samples 361 and 362 the ICE were detected and by increasing the insulin constraints, more insulin can be given earlier. For samples 363–365, MPE shows the prediction is higher than real CGM, insulin rate then decreased after returning to prevent future hypoglycemia. For samples 382–385, WRE and MPE are detected so that the controller is retuned to lower the insulin infusion rates to prevent future hypoglycemia. In sample 342 the SME is detected so that insulin rate is changed accordingly to make the CGM more stable. For samples like 365 and 393, although controller errors are detected, the insulin infusion rate cannot be reduced because it is already at 0, so the insulin rate is kept the same before and after retuning.

5 Conclusions

A new module for controller performance assessment in model-based AP control systems is developed and tested. Six indexes are developed for controller error detection. Four kinds of controller errors have been defined, detected, and used for controller retuning. The CPA module was tested with simulation studies. The results indicate that a controller performance assessment module with controller fault detection and controller retuning can improve the performance of the AP. Controller faults analysis and controller retuning are part of the fault detection and diagnosis in AP systems to assure that the controller itself is maintained at peak performance.

Acknowledgments This work is supported by the National Institutes of Health (NIH) under grants 1DP3DK101077-01 and 1DP3DK101075-01 and the Juvenile Diabetes Research Foundation International (JDRF) under grant 17-2013-472.

References

1. Atlas, E., Nimri, R., Miller, S., Grunberg, E.A., Phillip, M.: MD-logic artificial pancreas system: a pilot study in adults with type 1 diabetes. Diabetes Care **33**(5), 1072–1076 (2010)
2. Bequette, B.W.: Challenges and recent progress in the development of a closed-loop artificial pancreas. Annu. Rev. Control **36**(2), 255–266 (2012)
3. Bequette, B.W.: Fault detection and safety in closed-loop artificial pancreas systems. J. Diabetes Sci. Technol. **8**(6), 1204–1214 (2014)
4. Breton, M., Farret, A., Bruttomesso, D., Anderson, S., Magni, L., Patek, S., Dalla Man, C., Place, J., Demartini, S., Del Favero, S., Toffanin, C., Hughes-Karvetski, C., Dassau, E., Zisser, H., Doyle, F.J., De Nicolao, G., Avogaro, A., Cobelli, C., Renard, E., Kovatchev, B.: Fully integrated artificial pancreas in type 1 diabetes: modular closed-loop glucose control maintains near normoglycemia. Diabetes **61**(9), 2230–2237 (2012)
5. Cinar, A., Palazoglu, A., Kayihan, F.: Chemical Process Performance Evaluation. CRC Press, Boca Raton (2007)
6. Desborough, L., Harris, T.: Performance assessment measures for univariate feedback control. Can. J. Chem. Eng. **6**(70), 1186–1197 (1992)
7. Doyle, F.J., Huyett, L.M., Lee, J.B., Zisser, H.C., Dassau, E.: Closed-loop artificial pancreas systems: engineering the algorithms. Diabetes Care **37**(5), 1191–1197 (2014)
8. El-Khatib, F.H., Russell, S.J., Magyar, K.L., Sinha, M., McKeon, K., Nathan, D.M., Damiano, E.R.: Autonomous and continuous adaptation of a bihormonal bionic pancreas in adults and adolescents with type 1 diabetes. J. Clin. Endocrinol. Metab. **99**(5), 1701–1711 (2014)
9. Eren-Oruklu, M., Cinar, A., Quinn, L., Smith, D.: Adaptive control strategy for regulation of blood glucose levels in patients with type 1 diabetes. J. Process control **19**(8), 1333–1346 (2009)
10. Gorton, I., Gracio, D.: Data-Intensive Computing: Architectures, Algorithms, and Applications. Cambridge University Press, New York (2012)
11. Harvey, R.A., Dassau, E., Bevier, W.C., Seborg, D.E., Jovanovič, L., Doyle, F.J., Zisser, H.C.: Clinical evaluation of an automated artificial pancreas using zone-model predictive control and health monitoring system. Diabetes Technol. Ther. **16**(6), 348–357 (2014)
12. Huang, B.: Bayesian methods for control loop monitoring and diagnosis. J. Process Control **9**(18), 829–838 (2008)
13. Huang, B., Shah, S.: Performance Assessment of Control Loops: Theory and Applications. Springer Science & Business Media, London (1999)

14. Jelali, M.: An overview of control performance assessment technology and industrial applications. Control Eng. Pract. **5**(14), 441–466 (2006)
15. Kendra, S., Cinar, A.: Controller performance assessment by frequency domain techniques. J. Process Control **3**(7), 181–194 (1997)
16. Kendra, S.J., Basila, M.R., Cinar, A.: A supervisory KBS for real—time monitoring and modification of multivariable controllers for continuous processes. Methods and Applications of Intelligent Control, pp. 139–171. Kluwer academic publishers, Norwell (1997)
17. Kovatchev, B.P., Breton, M., Man, C.D., Cobelli, C.: In silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes. J. Diabetes Sci. Technol. **3**(1), 44–55 (2009)
18. Leal, Y., Ruiz, M., Lorencio, C., Bondia, J., Mujica, L., Vehi, J.: Principal component analysis in combination with case-based reasoning for detecting therapeutically correct and incorrect measurements in continuous glucose monitoring systems. Biomed. Signal Process. Control **6**(8), 603–614 (2013)
19. Loquasto, F., Seborg, D.: Monitoring model predictive control systems using pattern classification and neural networks. Ind. Eng. Chem. Res. **20**(42), 4689–4701 (2003)
20. Luijf, Y.M., DeVries, J.H., Zwinderman, K., Leelarathna, L., Nodale, M., Caldwell, K., Kumareswaran, K., Elleri, D., Allen, J.M., Wilinska, M.E., Evans, M.L., Hovorka, R., Doll, W., Ellmerer, M., Mader, J.K., Renard, E., Place, J., Farret, A., Cobelli, C., Del Favero, S., Dalla Man, C., Avogaro, A., Bruttomesso, D., Filippi, A., Scotton, R., Magni, L., Lanzola, G., Di Palma, F., Soru, P., Toffanin, C., De Nicolao, G., Arnolds, S., Benesch, C., Heinemann, L.: Day and night closed-loop control in adults with type 1 diabetes: a comparison of two closed-loop algorithms driving continuous subcutaneous insulin infusion versus patient self-management. Diabetes Care **36**(12), 3882–3887 (2013)
21. Mauseth, R., Wang, Y., Dassau, E., Kircher, R., Matheson, D., Zisser, H., Jovanovic, L., Doyle, F.J.: Proposed clinical application for tuning fuzzy logic controller of artificial pancreas utilizing a personalization factor. J. Diabetes Sci. Technol. **4**(4), 913–922 (2010)
22. Miller, R., Desborough, L., Timmons, C.: Citgo's experience with controller performance assessment. In: Proceedings of the NPRA 1998 Computer Conference (San Antonio,TX, USA 1998)
23. Paulonis, M., Cox, J.: A practical approach for large-scale controller performance assessment, diagnosis, and improvement. J. Process Control **2**(13), 155–168 (2003)
24. Qin, S., Yu, J.: Recent developments in multivariable controller performance monitoring. J. Process Control **3**(17), 221–227 (2007)
25. Renard, E., Place, J., Cantwell, M., Chevassus, H., Palerm, C.C.: Closed-loop insulin delivery using a subcutaneous glucose sensor and intraperitoneal insulin delivery: feasibility study testing a new model for the artificial pancreas. Diabetes Care **33**(1), 121–127 (2010)
26. Schäfer, J., Cinar, A.: Multivariable MPC system performance assessment, monitoring, and diagnosis. J. Process Control **2**(14), 113–129 (2004)
27. Sherr, J.L., Cengiz, E., Palerm, C.C., Clark, B., Kurtz, N., Roy, A., Carria, L., Cantwell, M., Tamborlane, W.V., Weinzimer, S.A.: Reduced hypoglycemia and increased time in target using closed-loop insulin delivery during nights with or without antecedent afternoon exercise in type 1 diabetes. Diabetes Care **36**(10), 2909–2914 (2013)
28. Stanfelj, N., Marlin, T., MacGregor, J.: Monitoring and diagnosing process control performance: the single-loop case. Ind. Eng. Chem. Res. **2**(32), 301–314 (1993)
29. Steil, G., Rebrin, K., Mastrototaro, J.J.: Metabolic modelling and the closed-loop insulin delivery problem. Diabetes Res. Clin. Prac. **74**(Suppl 2), S183–186 (2006)
30. Steil, G.M., Palerm, C.C., Kurtz, N., Voskanyan, G., Roy, A., Paz, S., Kandeel, F.R.: The effect of insulin feedback on closed loop glucose control. J. Clin. Endocrinol. Metab. **96**(5), 1402–1408 (2011)
31. Thornhill, N., Horch, A.: Advances and new directions in plant-wide disturbance detection and diagnosis. Control Eng. Prac. **10**(15), 1196–1206 (2007)
32. Tian, X., Chen, G., Cao, Y., Chen, S.: Performance monitoring of mpc based on dynamic principal component analysis. In: Preprints of the 18th IFAC World Congress (2011)

33. Turksoy, K., Cinar, A.: Adaptive control of artificial pancreas systems—a review. J. Healthc. Eng. **5**(1), 1–22 (2014)
34. Turksoy, K., Quinn, L., Littlejohn, E., Cinar, A.: Multivariable adaptive identification and control for artificial pancreas systems. IEEE Trans. Biomed. Eng. **61**(3), 883–891 (2014)
35. Turksoy, K., Quinn, L.T., Littlejohn, E., Cinar, A.: An integrated multivariable artificial pancreas control system. J. Diabetes Sci. Technol. **8**(3), 498–507 (2014)
36. Walsh, J., Roberts, R.: Pumping Insulin. Torrey Pines Press, San Diego (2006)