UNIVERSITY OF PADOVA

DEPARTMENT OF INFORMATION ENGINEERING

Master's Degree in Bioengineering

Development of a measurement error model of new factory-calibrated continuous glucose monitoring sensors used in type 1 diabetes therapy

Candidate
Cristina Battocchio

Supervisor
Prof. Andrea Facchinetti

Co-Supervisor
Dr. Martina Vettoretti

Academic Year 2018/2019

Abstract

Type 1 diabetes (T1D) is a chronic autoimmune disorder caused by the destruction of β-cells in the pancreas, that leads to insulin deficiency. Thus, people with type 1 diabetes need an everyday exogenous insulin delivery to maintain the blood glucose concentration as much as possible in the euglycemic range. The tuning of insulin doses depends on many parameters; among them, one of the most significant is the glucose concentration in blood.

Therefore, frequent and accurate measurements of glucose concentration are essential in managing diabetes. Minimally-invasive continuous glucose monitoring (CGM) sensors allow to measure the glucose concentration almost continuously for several days and are becoming key instruments in the management of diabetes therapy. However, as any measurement system, CGM readings are affected by errors that can deteriorate the performance of CGM-based applications, such as bolus calculators and artificial pancreas systems. The development of a model able to describe the CGM sensor error can be very useful, e.g. to reproduce sensor behaviors in silico when designing and testing CGM-based applications. Several models to describe the CGM error are available in the literature. However, their domain of validity is limited to 12-hour windows, i.e. the time interval between two consecutive calibrations. The recent availability of factory calibrated CGM sensor is calling for CGM error models able to describe CGM inaccuracy and behaviour in longer time intervals, e.g 7-10 days. The aim of this thesis is to develop a new model of CGM sensor error for factory-calibrated CGM device.

The methodology that we propose exploits and improves the model created by Facchinetti et al. [1, 2]. Differently from that model, whos maximum domain of validity is the 12-hour window between two consecutive calibrations, the new model is able to describe the CGM behaviour on the entire

lifetime of the sensor (10 days).

The dataset used for model identification consists of 81 adults whose BG concentration has been measured in parallel by Dexcom G6 sensors (Dexcom Inc, San Diego, CA) and YSI instrument (used as gold standard). The identification is performed by using two different methods, one in two steps and one in a single step, and the results are then compared.

The thesis is organized in six chapters. In Chapter 1, after an overview of the T1D therapy, the techniques to measure the glucose concentration and a detailed description of Dexcom G6 sensor are illustrated. Then, the issue of CGM sensors inaccuracy and the principal sources of error are presented. At the end of the chapter a review on the literature of sensor error models is also reported, and the aim of the thesis is defined.

Chapter 2 describes the dataset used for the model identification and criteria used for the subjects selection, together with data pre-processed.

In Chapter 3, the new methodology for modeling the CGM sensor error is presented. As previous models, it includes three components: the BG-to-IG kinetics, the calibration error, and the measurements noise. The main innovations are two. First, the modification of the calibration error model to describe such a component in factory-calibrated sensors. In particular, several candidate models have been tested and compared. Second, the match of a single-step identification procedure, that allows overcoming the limitation of the state-of-art identification which requires two steps.

In Chapter 4, the results obtained by applying the new model with the state-of-art two-step identification procedure are reported. In the first step, we determine the optimal calibration error model, its corresponding parameters and the ones of the BG-to-IG model. Then, the parameters precision and correlations are investigated. In the second step, we determine the optimal order of the AR process that describes the measurement noise component and its corresponding parameters on the different days of monitoring.

In Chapter 5, the results obtained by applying the new model with the single-step identification procedure are reported. Differently from the two-step method, the single step performs the identification of all the parameters simultaneously. As in the previous chapter, we report the optimal AR model and the identified parameters with the corresponding precision. The results provided by the two identification methods are then compared, showing that

they are almost equivalent. However, the single-step one provides a more accurate description of the BG-to-IG kinetics and allows to estimate all the parameters simultaneously. For these advantages one may be inclined to select this identification method with respect to the two-step one.

Finally, the main findings of this work and the future developments are summarized in Chapter 6.

Contents

List of Figures

List of Tables

Chapter 1

Diabetes and Continuous Glucose Monitoring

1.1 Diabetes and its therapy

Blood glucose concentration is controlled by insulin, a hormone secreted in the pancreas by a specific type of cells, the β-cells. Insulin limits the amount of glucose in blood when it reaches too high values, keeping its level in the physiological range of 70-180 mg/dl. Diabetes is a disorder that affects the body's ability to regulate blood glucose. Diabetic people lack insulin effectiveness, so their glucose level becomes too high with respect to healthy bounds. This condition, called hyperglycemia, has no severe complications in short-term, but it may induce long-term serious effects like retinopathy, nephropathy, and neuropathy [3].

Nowadays, diabetes is considered a global health emergency: people affected by diabetes are more than 425 million worldwide, and they are expected to rapidly increase to approximately 629 million in 2045 [4]. Given such gloomy scenario, finding the best treatments is very important to guarantee a good quality of life to those affected by this disease.

There are two main types of diabetes, and each of them requires different treatments. Type 2 diabetes is a disease characterized by the inability of body tissues and organs to properly use the circulating insulin. It affects approximately the 90-95% of all diabetics; the main risk factors are genetic predispositions, obesity, and sedentary lifestyles. Common therapy for type

Figure 1.1: Insulin infusion using daily injections. (Source: Diabetes Research & Wellness Foundation, UK, https://www.drwf.org.uk)

2 diabetes includes glucose monitoring, healthy diet, physical activity, and drug administration.

Type 1 diabetes is a chronic autoimmune disorder caused by the destruction of β-cells in the pancreas. Since very little or no insulin is produced [5], people with type 1 diabetes need an everyday exogenous insulin delivery to maintain glucose concentration in the euglycemic range. This is not easy to achieve as a surplus of insulin can decrease the glucose level below the safe bound of 70 mg/dl. This condition, called hypoglycemia, can lead to severe consequences such as the loss of consciousness, seizure, coma, or even death [6]. Typical insulin doses range from 0.4 to 1.0 units/kg/day, with the right amount depending on many parameters, like patient's weight, time of meals, carbohydrate consumption and, most of all, measurements of glucose concentration in blood [7].

1.2 Glucose monitoring techniques

Precise and accurate measurements of glucose concentration have a key role in good diabetes management.

Before 1970s, the glucose in blood was estimated from the one measured in the urine. Such measurements were manual and limited, for this reason they were mainly used for diagnosis and critical care management rather than to achieve specific glycaemic goals [8]. The history of at-home diabetes

monitoring started in 1971, when Anton Clemens developed, for point-of-care use, the first blood glucose monitor based on the reflectance of the light from the surface of a glucose oxidase-based strip. In the 1980s, more convenient electrochemical techniques were developed, which brought to self-monitoring blood glucose (SMBG) devices. Since then, SMBG devices are fundamental instruments in the daily routine of diabetic patients, and they represent today the most widespread method of self glucose monitoring at home.

SMBG devices are compact and very easy to use [9, 10]: patients prick their finger with a lancet device to obtain a small blood sample that is put over a reagent strip. The strip is inserted into a reflectance photometer, and the glucose concentration value is displayed on a screen. The main drawback of these devices is the rate of sampled data: because the finger prick is intrusive and painful, a standard frequency of only 3-4 measurements per day can be collected. Because of the sparsity of these measurements, dangerous hypoglycaemic or hyperglicaemic events may be not detected by SMBG [11] [12].

The need to have frequent glucose data impels to a new generation of devices: the continuous glucose monitoring (CGM) sensors. CGM sensors are wearable devices able to take glucose measurements at regular intervals, e.g., every 1-5 minutes for several consecutive days. Continuous data provides several advantages because they can be used to implement proactive actions, like preventing potential hypoglycaemia with carbohydrates intake or performing insulin dose modifications to avoid future hyperglycaemia. Furthermore, acoustic and visual alerts can warn the careless wearer if glucose is moving out of the safe range [13].

1.3 Minimally-invasive CGM sensors

Minimally-invasive systems measure glucose concentration in the interstitial fluid between cells and capillaries, without nicking blood vessels. Depending on how the interstitial fluid is sampled these devices can be based on needle, microdialysis, or reverse iontophoresis.

Most popular minimally-invasive CGM devices consist of three main elements: a needle-based sensor, which measures an electrical signal propor-

tional to interstitial glucose concentration; a transmitter, applied over the sensor to transmit the signal; and a receiver, that displays glucose concentration to the user. More specifically, the electrical signal is generated by a glucose-oxidase electrochemical reaction, which is induced by an enzyme-coated wire inserted by the patient in the abdominal or arm subcutis. The signal is converted to a glucose concentration via a calibration process that exploits previous SMBG patient's measurements. The implemented calibration algorithm often assumes a first-order and time-independent association between the electrical signal and the glucose level. This simplistic way to model a complex relation forces the system to require periodic calibrations (every nearly 12 hours) by patients with their consequent discomfort. To solve this problem, novel calibration procedures were investigated in the last years. These new techniques allow the development of sensors that do not require in vivo calibrations: the so-called "factory-calibrated" sensors [14].

Nowadays, only few type 1 diabetes patients are using CGM sensors. The main reasons are high devices cost, perceived sensor inaccuracy, difficulties in alarms management, and aversion to wearing devices on the body. Nevertheless, the number of users is rapidly increasing thanks to new low-cost and more accurate sensors. Moreover, some CGM devices received regulatory approval for nonadjunctive use both in Europe and the United States, e.g., Dexcom G5 received both the CE marking and the FDA approval in 2016. The official permission allows patients to base their treatment decisions on CGM measurements, without the need of confirming CGM readings with SMBG measurements [15].

1.3.1 Dexcom G6 device

The Dexcom G6 System provides continuous glucose level concentration with sampling time of 5 minutes, for up to 10 consecutive days. Currently available in 16 countries, it received both the CE marking and the FDA approval for nonadjunctive use in 2018. It includes a minimally invasive CGM sensor, a transmitter, and a compatible wireless display device (Figure 1.2):

- The sensor is a sterile device made of an applicator, a transmitter holder, and a probe. The applicator is a single-use disposable unit that helps patients to insert the probe under the skin. The probe may be

Figure 1.2: Dexcom G6 components.The system includes the auto- applicator (1), the transmitter (2), and the display devices that can be a touch-screen receiver, a smartphone or a smartwatch (3).

placed in the abdomen subcutis for adults, and both in the abdomen or in the buttock subcutis for children aged 2-17 years old. Once inserted, the probe starts measuring interstitial glucose levels every 5 minutes.

- The transmitter is a Bluetooth Low Energy (BLE) device that communicates in real-time the sensor glucose readings to the receiver. It is attached to the sensor through the transmitter holder and can be used for up to 3 months.

- The display device receives glucose information from the transmitter and shows them to the user. It informs the patient about the glucose trends and gives alarms when glucose is moving out of the physiological range.

Thanks to wireless communications between the transmitter and the receiver useful apps are designed for both android and iOS. With the Dexcom Share^{TM} app, patients can share their glucose readings and trends with up to 10 people. This is very useful for parents of diabetic children that can remotely monitor their children's glucose measurements. The CLARITY app identifies clinically relevant patterns out of a huge amount of glucose read-

ings, performing statistical analysis on them: this gives patients the ability to prioritize problems and find quick solutions.

This generation of devices include new advantages and features:

- The factory calibration eliminates the need of periodically calibrating the device using SMBG. However, if sensor glucose readings do not match the patient feelings, the calibration procedure must be performed in vivo using SMBG references. An advantage of the factory calibration is the remarkable reduction of SMBG measurements and the elimination of errors related to the execution of the calibration process, which can lead to sensor inaccuracies.

- A 10-day sensor wear period, longer than the 7-day lifefime of previous generation Dexcom sensor, reduces the number of insertion and the consequent patients' discomfort.

- The "Urgent Low Soon" alert predicts hypoglycemia events within 20 minutes advance, helping to avoid severe low blood sugar episodes.

- The acetaminophen (APAP) blocking allows accurate glucose readings with no medication interference. APAP is a medicine commonly used to treat mild to moderate pain, or to reduce fever. Unfortunately, it affects the glucose readings, generating a spurious signal which interferes with the sensor signal. To minimize or prevent APAP interference a permselective membrane coating is designed for Dexcom G6 sensor.

- The transmitter is 30 percent thinner than its predecessor and its size is reduced to almost 4 centimeters length.

- The one-touch auto-applicator simplifies the sensor insertion, which becomes less painful and less intimidating.

- The interoperability allows sharing glucose information to interoperable electronic interfaces, including compatible Automated Insulin Delivery (AID) systems. The CGM sensor interacts with other devices with 3 modalities: the transmitter communicates to another device through the same protocol, the app communicates to another app on a single

mobile platform, or the app communicates through the cloud to another software device.

In order to assess Dexcom G6 reliability, the sensor performances are tested in several investigations where the outcomes always prove the device consistence and precision [16] [17] [18]. Also FDA, before the approval, have examined data from two clinical studies conducted at 11 centers across the United States. These studies evaluated the accuracy of the sensor compared to a laboratory blood glucose measurement method. Results point out a good accuracy confirmed by a 9% mean absolute relative difference (MARD) that is the main performance metrics for accuracy evaluation of CGM sensor [19]. Thanks to these outcomes, the Dexcom G6 is approved to be used nonajunctively; therefore, the CGM readings can be used by patients to make treatment decisions without confirmatory SMBG values.

1.4 The CGM sensor error

CGM devices measure glucose concentration levels almost continuously, providing enormous advantages in the diabetes management. However, as any measurement system, they are affected by unpredictable errors that cause inaccuracy in CGM readings provided in output. Dissecting the error in its different contributions and evaluating them can be the key to improve sensor performances and to reduce their inaccuracy.

Details regarding the CGM sensor error are presented in the following sections.

1.4.1 Sources of CGM sensor error

The sources that may impair the sensor accuracy are several, such as delays, interfering substances, drifts in sensitivity, and calibration error. Understanding the nature of these inaccuracies is fundamental to model the sensor error.

A first source of error is related to the body site in which the CGM devices measure the glucose level. Minimally-invasive CGM systems measure the glucose concentration in the interstitial fluid (IG), rather than in the plasma

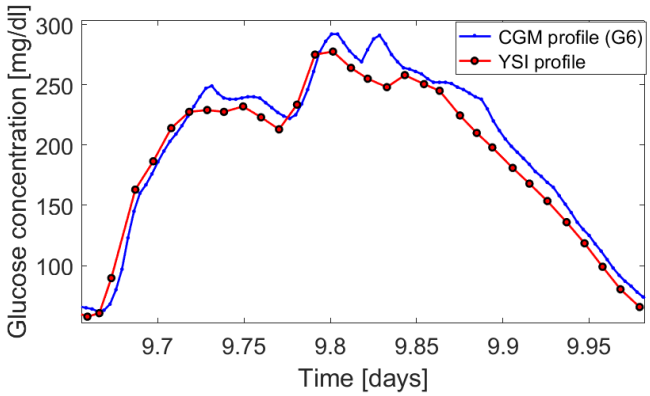

Figure 1.3: Representative Dexcom G6 sensor signal in a random time interval that exhibits a delay compared to the BG reference values (YSI values) due to IG-to-BG kinetics.

(BG) . Therefore, due to the physiological lag between IG and BG, which is in the order of minutes, the sensors glucose measurements are subjected to a delay, as reported in Figure 1.3.

A second source of error in CGM sensors is in the transduction of the electrochemical signal into an electrical signal. Some substances, like the acetaminophen (APAP) described in the previous section, can interfere with the process by generating spurious currents when they are oxidized at the sensor electrodes. This undesired effect causes an artificial raise of the measured glucose values.

A third source of error lays in the variation of the sensor sensitivity. When the sensor is inserted into the body, the immune system reacts because of the sensor membrane entering the biological environment, leading to a variation of the sensor sensitivity in time. This variation causes a non-physiological drift in time on the CGM profile, which can be observed in Figure 1.4.

Finally, a last source of error is in the sensor calibration process itself. Ideally, the calibration algorithm should perfectly match the electrical signal to the glucose concentration level of the patient by compensating the effects of the previous error sources. However, the calibration laws implemented in CGM devices are most of the times simple linear functions, which are not sufficient to completely describe the inter-subject and inter-sensor dynamics.

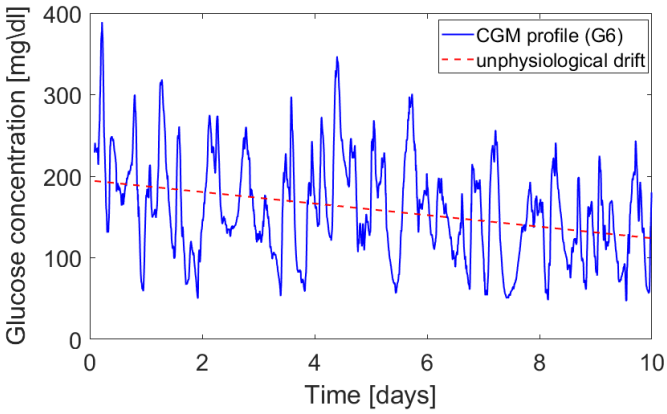

Figure 1.4: Representative Dexcom G6 sensor signal (blu continuous line) that exhibits a nonphysiological drift (red dashed line) due to time-variability of sensor sensitivity.

1.4.2 Literature models to describe the CGM sensor error

Developing a quantitative model of the CGM sensor error requires efforts from both the theoretical and the experimental point of view.

For example, the development of a mathematical model to describe the BG-to-IG kinetics is challenging because of its involved dynamic. But the effort is not only theoretic: once the model is settled, its validation requires to collect blood glucose values in parallel with CGM data, which is possible only with the hospitalization of patients and the intervention of clinics and resources. Consequently, there are only few studies on CGM sensor error modeling in literature.

In 2006, Chase and colleagues proposed a first simple sensor error model based on a random white noise process with a constant coefficient of variation [20]. Then, Breton and Kovatchev implemented a finer model based on two different datasets of the Abbott FreeStyle Navigator sensor (Chicago, IL, USA) [21]. The model included the distortion effect due to the BG-to-IG kinetics, and a linear regression model to calibrate CGM data. However, the results were impaired by some rough assumptions. First, the diffusion process was described as linear and time invariant for several days, and the inter-individual variability of the BG-to-IG kinetics was not considered. Sec-

ondly, the parameters of the calibration model were assumed to not vary in time; thus, the model was unable to describe completely the errors due to the calibration. Another study, proposed by Laguna and colleagues on the Dexcom SEVEN PLUS (San Diego, CA, USA) and the Medtronic Paradigm Veo Enlite sensors (Northridge, CA, USA), characterized several aspects of the sensor error, such as lag time, error stationary, error probability distribution, and time correlation [22].

Recently, Facchinetti and colleagues proposed a novel model [1, 2], where the error was dissected into its three main contributions: the delay due to the BG-to-IG kinetics, the calibration error, and the measurement noise. The error model was validated on CGM sensors of different generations produced by Dexcom Inc. (San Diego, CA, USA). All the literature studies in CGM sensor error have been performed using the data of past-generation CGM sensors that required periodic in vivo calibrations; thus, to the best of our knowledge, no models of the error of factory-calibrated CGM sensors are available in the literature.

1.5 Aim of the thesis

This thesis aims to develop the error model for factory-calibrated CGM sensors used in type 1 diabetes therapy. To achieve this goal, the model proposed by Facchinetti et al. in [1, 2] is modified and extended. First, several new functions for the calibration error model are tested and compared to investigate the time variability of CGM sensor sensitivity in longer time intervals, e.g. 7-10 days. Secondly, a novel single-step procedure where all the parameters are estimated simultaneously is proposed and validated to overcome the limitation of the state-of-art identification which requires two steps.

The dataset adopted to model the sensor error comes from a pivotal study on 81 patients wearing Dexcom G6 CGM sensors. A complete description of the available data and of the data pre-processing is provided in Chapter 2. In Chapter 3, first the model of Facchinetti et al. is explained in details, and then the two steps and the single step identification methods are proposed. In Chapter 4 and 5, the parameters are identified by using respectively the two-step and the single-step method. The obtained results are analyzed, each

in its corresponding chapter, and then compared in Chapter 5. Finally, the conclusions and the future works are described in Chapter 6.

Chapter 2

Dataset and Data pre-processing

In this chapter first we describe the dataset used to identify the error model of the factory-calibrated CGM sensors, and then we pre-process the data to remove saturated and spurious CGM values. Finally, we report the Bayesian smoothing procedure used in the pre-processing.

2.1 Dexcom G6 pivotal study dataset

The available data come from multi-center pivotal studies, performed in 2016-2017, on both adults (over 17 years old) and pediatric patients wearing Dexcom G6 sensors. The dataset includes 140 patients: 103 of them wearing a single sensor, and the remaining 37 wearing two sensors in parallel, placed on right and left sides of the abdominal region, respectively. Because only few patients wear two sensors, we decided to consider each of them belonging to different subjects. Consequently, the starting dataset consisted of 177 CGM profiles.

During the ten days of monitoring with the G6 sensor, patients were hospitalized on day 1 or 2, 4, 7 or 10, for a 12-hour period where blood glucose (BG) samples were collected as reference data. In particular, the BG samples were measured approximately every 15 ± 5 minutes by using a YSI (Yellow Spring, OH) glucose analyzer. For several possible reasons, this protocol was not strictly followed by all the subjects. Therefore, some of them were discarded from the dataset to avoid introducing erroneous data in the analysis. Specifically, subjects are discarded if

Table 2.1: Summary on the subject selection, according to the rules defined in Sec. 2.1.

Subjects	**Number**
total	177
discarded because without recorded YSI	-5
discarded because without YSI data on day 10	-86
discarded because with missing CGM data	-5
selected	81

- their YSI measurements are completely missing. In this case, we have no reference data to perform the analysis.
- their YSI measurements are missing on day 10. In this case, the available reference data would be not sufficient to describe the sensor error in the entire lifetime of 10 days.
- their CGM profile is not collected for the entire 10-days period and ends after few days of monitoring.

Figure 2.1 reports some examples of subjects whose YSI measurements or CGM profile were not correctly collected. After screening out the dataset, 81 subjects remain for the analysis; a summary of the selection is reported in Table 2.1.

2.2 Data pre-processing

After selecting the subjects, both their CGM and their YSI data are processed to remove not reliable values. Indeed, the accuracy of the data is fundamental to successfully identify the model of the CGM error.

The CGM sensor trace is affected by saturation to maximum and minimum displayable levels and by disconnections, i.e. missing samples. In the former, the reported level of glucose is above 400 mg/dl or below 40 mg/dl while, in the latter, the reported level assume a peculiar value below 40 mg/dl

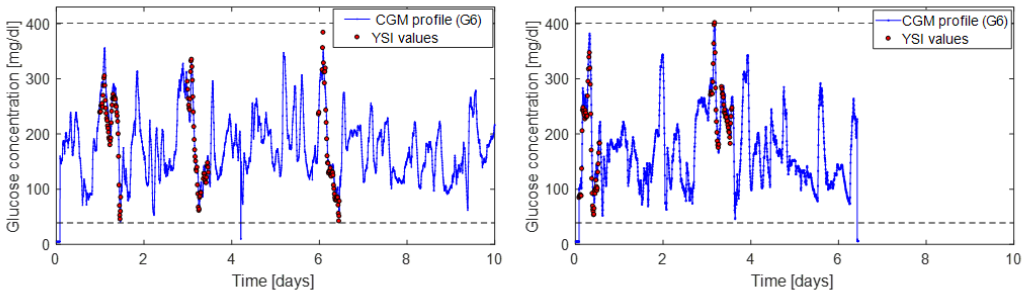

Figure 2.1: Examples of sensors profile obtained when the protocol instructions are not followed. The YSI measurements are not available in day 10 (left); the CGM signal is not collected for the entire 10-days period (right).

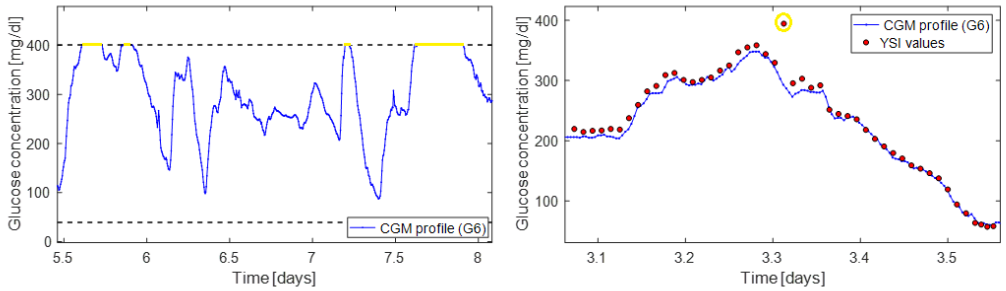

Figure 2.2: Data pre-processing on both the CGM data and YSI data. Examples reporting the elimination of the saturated CGM data (yellow lines) on the left, and of YSI considered as outliers (yellow circle) on the right.

(e.g., 5 mg/dl). In both cases, we discard the values from the analysis, as shown in Figure 2.2. Similarly, the YSI measurements can be affected by errors in their acquisition or recording. To remove the spurious data, we perform a visual inspection following two main rules. First, we eliminate the YSI values that are in contrast with the main trend of their neighbors. Second, we remove those that we consider outliers, which are the ones located too far from the others and the CGM profiles. An example of outlier value is presented in Figure 2.2.

2.2.1 Smoothing of YSI data

As explained in Sec 2.1, the reference data are collected for a 12-hour period, in different days, with a frequency of 15 ± 5 minutes. This frequency is lower than the one of the CGM samples, i.e. 5 minutes; thus, YSI profiles which will be used as a deterministic input in our identification process (see Sec. 2.1), need to be interpolated on a more dense sampling grid in order to be matched with CGM values.

For this purpose, we exploit a data approximation technique called *Bayesian smoothing*. In performing the *Bayesian smoothing*, first the uniform time grid for the smooth signal is set, then, the original signal is smoothed by following these two principles:

- Since the YSI data are affected by the measurement noise, the smoothing procedure should only *approximate* the data, without interpolate them exactly.
- Since the YSI profile is a biological signal, it must have some *regularities*. In particular, the regularity of the profile can be defined as the energy of its second derivative.

The trade-off between approximation and regularity is achieved by minimizing a target function (details available in the Appendix A). An example of Bayesian smoothing of YSI samples is reported in Figure 2.3, where the uniform grid of the smoothed profile is set to 1 minute sampling time.

The YSI data are not collected continuously but only when the patients are hospitalized. Moreover, while colleting YSI samples some measurements can miss, so the smoothed profile obtained between two adjacent samples could be unreliable. To prevent this situation, we segment the smoothed YSI profiles such that, if two adjacent YSI samples are collected more than 20 minutes away from each other, we discard the smoothed profile between the two samples. As an additional constraint, if the duration of a segment is less than one hour, we discard it. An example of the set of YSI segments obtained for a representative subject is reported in Figure 2.4.

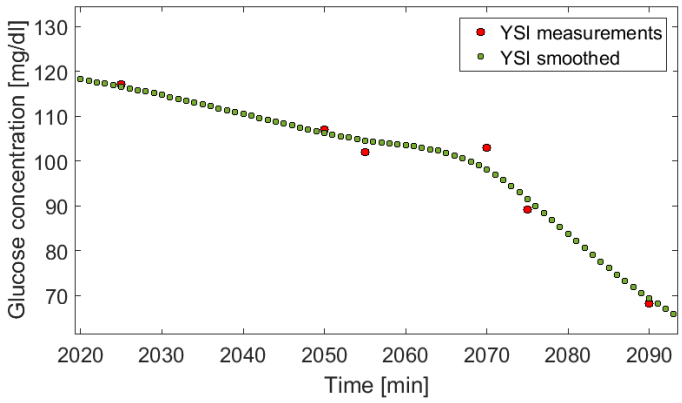

Figure 2.3: YSI values in a restricted time window before (red circles) and after (green circles) their reconstruction through the *Bayesian smoothing*. The reconstructed profile is an approximation of the original one.

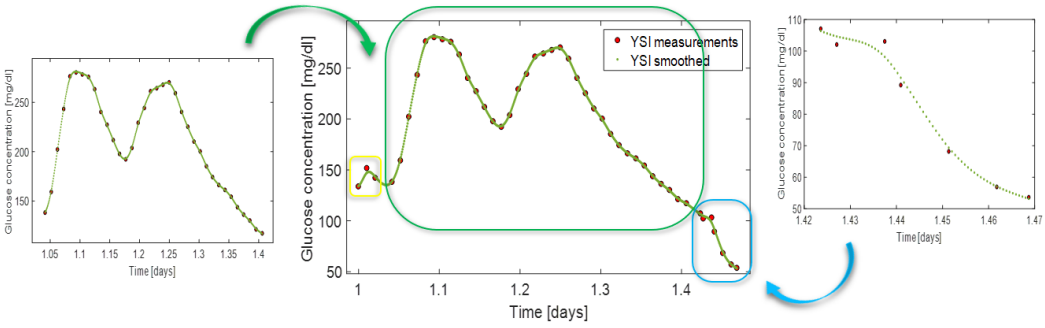

Figure 2.4: YSI profile in day 2 of a representative subject before and after its segmentation. The original signal (center) is split in three segments. Two (green square and blue square) are conserved while the remaining (yellow square) is discarded because its time period is smaller than an hour.

Chapter 3

Methodology for modeling the CGM sensor error

3.1 The model proposed by Facchinetti et al. for past generation sensors

Facchinetti et al. proposed a innovative model of sensor error based on the separation of the CGM inaccuracy components [1, 2], which allows to investigate the error of any commercial CGM sensor. A model of the different sensor error components is useful to test in simulation several applications, such as algorithms for signal processing, real-time glucose prediction, insulin dosing, and artificial pancreas (AP) (e.g., incorporating a model of the CGM error in the T1D simulator of Padua and Virginia University [23]).

According to the model, the error of the sensor has three main contributions: the BG-to-IG kinetics, the sensor calibration error, and the measurement error. While the first is related to a physiological process, the last two are specific of the sensor itself. To characterize each error component this technique exploits n simultaneous CGM sensors, as reported in Figure 3.1, and assumes the interstitial glucose (IG) concentrations underlying the CGM trace of each sensor to be the same. This hypotesis is equivalent to have no physiological variability from a sensor insertion site to another.

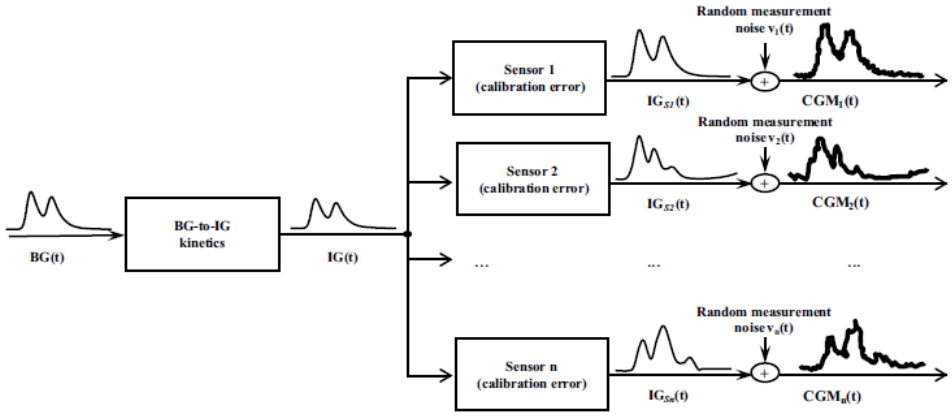

Figure 3.1: Schematic description of how simultaneus CGM data streams are modeled. From left to right: the BG(t) signal is transformed into IG(t) signal through the BG-to-IG kinetics; the IG signal is measured by each of the n CGM sensors, generating for the i-th sensor the$\mathrm{IG}_{Si}(t)$ profile; finally, the measured $\mathrm{CGM}_i(t)$ is affected by additive measurement noise $v_i(t)$.

The CGM trace of the sensor i is given by

$$\mathrm{CGM}_i(t) = \mathrm{IG}_{Si}(t) + v_i(t), \tag{3.1}$$

where, at time t, v_i is the measurement error and IG_{Si} is the value of the IG read by the sensor. If the sensors were perfectly calibrated, the $\mathrm{IG}_{Si}(t)$ signals would be equal to IG(t). In reality, however, the $\mathrm{IG}_{Si}(t)$ of each sensor deviates from the true value because of errors in the calibration process or because of drifts in time due to changes of sensor sensitivity. Such deviation, which we refer to as *calibration error*, results in different IG_{Si} signals for each sensor.

Details regarding the model components are presented in the following sections.

3.1.1 Model of BG-IG kinetics

Minimally-invasive systems are based on measurement of IG rather than BG in order to reduce invasiveness of CGM devices. Since plasma and interstitial fluid are separated by a capillary barrier, the interstitial profile is a distorted

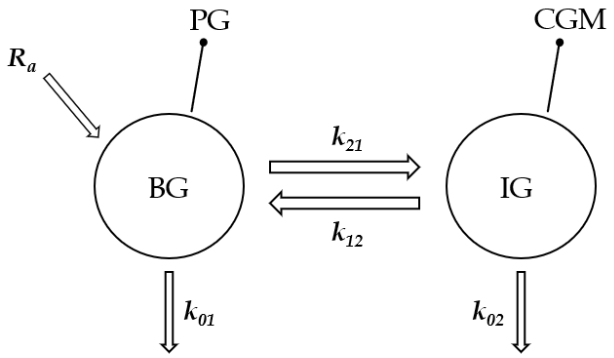

Figure 3.2: Compartmental model describing the BG-to-IG kinetics. BG is the plasma glucose concentration, IG is the interstitial glucose concentration, Ra is the rate of appearance of glucose in the blood, k_{ij} (i = 0, 1, 2 and j = 1, 2) are the diffusion constants. PG (plasma glucose) and CGM (interstitial profile) are the accessible measures of the two compartments.

and delayed version of the blood one. This physiological process is the first source of error to model in CGM sensors. A relative simple but effective model was proposed by Rebrin et al. [24], which described the BG-to-IG kinetics process as the two-compartment model represented in Figure 3.2. In such a model, BG and IG are respectively the plasma and the interstitial glucose concentration, R_a is the rate of appearance of the glucose in the blood, and k_{12}, k_{21}, k_{01}, and k_{02} are the diffusion constants. The two differential equations that describe the system are

$$\dot{\mathrm{BG}}(t) = R_a + k_{12}\mathrm{IG}(t) - (k_{01} + k_{21})\mathrm{BG}(t), \tag{3.2}$$

$$\dot{\mathrm{IG}}(t) = k_{21}\mathrm{BG}(t) - (k_{02} + k_{12})\mathrm{IG}(t). \tag{3.3}$$

The a priori non-identifiability leads to the following parametrization:

$$\dot{\mathrm{IG}}(t) = -\frac{1}{\tau}\mathrm{IG}(t) + \frac{g}{\tau}\mathrm{BG}(t), \tag{3.4}$$

where τ is the diffusion time constant and g is the gain of the system.

To find the transfer function of the system we use the Laplace transforms,

obtaining

$$\mathrm{IG}(s) = H(s)\mathrm{BG}(s) \qquad \text{and} \qquad H(s) = \frac{g}{\tau}\frac{1}{s+\frac{1}{\tau}}. \tag{3.5}$$

Therefore, in the time domain, the glucose concentration in the interstitial fluid is given by

$$\mathrm{IG}(t) = h(t) \otimes \mathrm{BG}(t) \tag{3.6}$$

where $\otimes$ is the convolution operator and $h(t)$ is the impulse response of the BG-to-IG system,

$$h(t) = \frac{g}{\tau}e^{-\frac{t}{\tau}}. \tag{3.7}$$

In steady state, $\dot{\mathrm{IG}} = 0$ and $\mathrm{IG} = \mathrm{BG}$ [25], therefore the gain g is equal to 1. Indeed, from Eq. (3.4) one gets

$$\frac{1}{\tau}\mathrm{IG} = \frac{g}{\tau}\mathrm{BG}, \tag{3.8}$$

$$\mathrm{IG} = g \cdot \mathrm{BG}, \tag{3.9}$$

$$g = 1, \tag{3.10}$$

and the IG concentration is then given by [25]

$$\begin{aligned} \mathrm{IG}(t) &= h(t) \otimes \mathrm{BG}(t) \\ &= \left(\frac{1}{\tau}e^{-\frac{t}{\tau}}\right) \otimes \mathrm{BG}(t). \end{aligned} \tag{3.11}$$

According with the model, τ is the same for every of the n sensors in a subject, and its identification is performed by nonlinear least squares.

3.1.2 Calibration error model

The IG signal is measured independently by the multiple sensors, generating the IG_{Si} profiles (S stands for the sensor, while i refers to the sensor number ($i = 1, \ldots, n$)). To define the relationship between IG_{Si} and $\mathrm{IG}(t)$ some critical aspects have to be taken into account: the calibration process of the sensor can be suboptimal, as described in section 1.4.1; and the variability of its sensitivity may produce a significant drift in time not suitably com-

pensated by the calibration process. The resulting calibration error affecting the CGM trace between two consecutive calibrations, i.e. in a time-window of 12-hour duration, is described by the following equation:

$$\begin{aligned} \mathrm{IG}_{Si}(t) &= a_i(t)\mathrm{IG}(t) + b_i(t) \\ &= a_i(t)\Big[h(t) \otimes \mathrm{BG}(t)\Big] + b_i(t), \end{aligned} \tag{3.12}$$

where $a_i(t)$ and $b_i(t)$ are the time-varying gain and offset for the i-th sensor. Ideally, sensors free of calibration error would have $a(t) = 1$ and $b(t) = 0$.

Several options can be taken into account to model $a_i(t)$ and $b_i(t)$. In absence of *a priori* information on their evolution in time, polynomial models are used because of their flexibility. The two functions are then given by

$$a_i(t) = \sum_{k=0}^{m} a_{ik} t^k, \tag{3.13}$$

$$b_i(t) = \sum_{k=0}^{l} b_{ik} t^k, \tag{3.14}$$

where m and l are the degrees of the polynomials, and a_{ik} and b_{ik} are the corresponding coefficients.

The selection of optimal values for m and l is performed by minimizing the Bayesian information criterion (BIC) index, while the identification of the coefficients a_{ik} and b_{ik} are estimated via nonlinear least squares.

3.1.3 Model of the measurement noise

In addition to the calibration error, CGM signals are affected by an additive noise $v_i(t)$ (see Eq. (3.1)). Thanks to the availability of multiple sensor per subject, $v_i(t)$ can be dissected into two components, one common to all the sensors and one sensor-specific. The common component, $cc(t)$, is assumed to be equal in all CGM sensors of a single subject and accounts for the possible suboptimal modeling of the previous steps, while the sensor-specific component $scc_i(t)$ is specific to the i-th device and uncorrelated with the other devices.

Since both the models of the BG-IG kinetics and the calibration error

cannot explain the sensor error completely, $v_i(t)$ contains some dynamics not considered by these models and cannot be in general considered as a zero-mean random measurement only. According to [26], the two components of $v_i(t)$ are modeled as autoregressive (AR) processes:

$$\mathrm{cc}(t) = \sum_{k=1}^{r} \beta_{ik}\mathrm{cc}(t-k) + w_1(t) \tag{3.15}$$

$$\mathrm{scc}_i(t) = \sum_{k=1}^{q} \alpha_{ik}\mathrm{scc}_i(t-k) + w_{i2}(t) \tag{3.16}$$

where α_{ik} and β_{ik} are respectively the parameters of the AR model of orders q and r, while $w_1(t) \sim \mathcal{N}(0, \sigma^2_{w_1})$ and $w_{i2}(t) \sim \mathcal{N}(0, \sigma^2_{w_{i2}})$ are white noise random processes.

The identification of the orders q and r of the two AR models is performed by minimizing the BIC criterion.

3.2 The proposed error model for factory - calibrated CGM sensors

As reported in section 1.3.1, factory-calibrated CGM devices eliminate the need of sensor calibration in vivo and associated SMBG measurements. Therefore, in factory-calibrated devices, the calibration procedure is performed by the manufacturers instead of the patients, based on the sensor sensitivity determined during the manufacturing process [27]. The only action required by the patients is to transcript the calibration code provided with the sensor during the setup phase. After that, no further calibration is needed.

The proposed model of the error of factory-calibrated CGM sensors is reported in Figure 3.3. Differently from the previous analysis, where the error was characterized in the time windows between two consecutive calibrations (12-hour time window), here we consider a time span that is the entire lifetime of the sensor (10 days). Moreover, the analysis considers one sensor, i.e., one CGM trace, for each subject instead of having multiple simultaneous sensors. The choice of having one rather than multiple CGM sensors per patient is due to the fact that CGM datasets, including the one used in this work, are

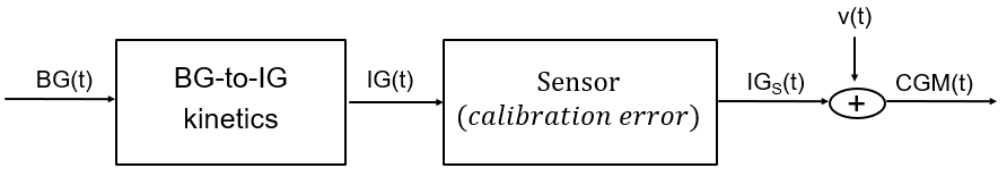

Figure 3.3: Schematic description of how a factory-calibrated CGM data stream is modeled. From left to right: the BG(t) signal is transformed into IG(t) signal through the BG-to-IG kinetics; the IG signal is measured by the CGM sensor, generating the$\text{IG}_S(t)$ profile; finally, the measured CGM(t) is affected by additive measurement noise $v(t)$.

mostly collected using a single sensor per subject.

Similarly to the previous analysis, the scheme can be decomposed in three sub-models:

1. The BG-to-IG kinetics model, which is exactly the same described in Sec 3.1.1, according to which the BG-IG relation is given by Eq. (3.11).

2. The calibration error model, which is characterized as in Eq. (3.12) where $a_i(t)$ and $b_i(t)$ are proper functions to be determined and i is the index of the subject.

3. The model of the measurement noise, which consists of a single component $v_i(t)$, modeled as the following AR process:

$$v_i(t) = \sum_{k=1}^{q} \alpha_{ik} v_i(t-k) + w_i(t) \tag{3.17}$$

   where i refers to the subject and $w_i(t)$ is a zero-mean white noise process with variance σ^2.

In conclusion, the resultant CGM_i sensor output is defined by:

$$\text{CGM}_i(t) = \Phi_i(t) + w_i(t) \ , \qquad \Phi_i(t) = \text{IG}_{Si}(t) + \sum_{k=1}^{q} \alpha_{ik} v_i(t-k) \tag{3.18}$$

where $\Phi(t)$ models the CGM sensor measurements without the zero-mean white noise component.

3.2.1 Candidate calibration error models

In order to correctly model the calibration error, we investigate several types of polynomial and exponential functions for the parameters $a_i(t)$ and $b_i(t)$.

As reported in Sec. 3.1.2, the polynomial functions in Eqs. (3.13) and (3.14) allow a wide range of behaviors by selecting the degrees m and l properly. In the model, we restrict our choice to m and l ranging from zero up to a maximum of three, corresponding to behaviors spanning from time invariant to cubic. For instance, by selecting $m = l = 2$ a quadratic time evolution is considered both for $a_i(t)$ and $b_i(t)$, and $\mathrm{IG}_{Si}(t)$ is given by

$$\mathrm{IG}_{Si}(t) = \Big[a_0 + a_1 t + a_2 t^2\Big]\mathrm{IG}(t) + \Big[b_0 + b_1 t + b_2 t^2\Big]. \tag{3.19}$$

Instead, by selecting $m = 2$ and $l = 0$ the gain obeys to a quadratic law while the offset is time invariant, $\mathrm{IG}_{Si}(t)$ is then given by

$$\mathrm{IG}_{Si}(t) = \Big[a_0 + a_1 t + a_2 t^2\Big]\mathrm{IG}(t) + b_0. \tag{3.20}$$

According to [28], and two patent applications deposited by Dexcom [29, 30], the time variability of CGM sensor sensitivity over the entire sensor lifetime can be well describe by exponential functions. Therefore, in this thesis, we considered two additional models for $a(t)$ and $b(t)$, i.e. the mono-exponential model:

$$s_1(t) = m_0 \cdot \Big[1 + \frac{m_f - m_0}{m_0} \cdot \Big(1 - e^{-\gamma t}\Big)\Big], \tag{3.21}$$

and the bi-exponential model:

$$s_2(t) = m_0 \cdot \Big\{1 + \frac{m_f - m_0}{m_0} \cdot \Big[\iota \cdot \Big(1 - e^{-\gamma t}\Big) + \Big(1 - \iota\Big) \cdot \Big(1 - e^{-\delta t}\Big)\Big]\Big\} \tag{3.22}$$

where γ, δ, and ι are the exponential decay constants that model the sensor calibration error, m_0 defines the initial sensitivity condition when the sensor is initially inserted into the tissue ($t = 0$), and m_f defines the final condition

($t = \inf$). In the following of this thesis, we will refer to $s_1(t)$ and $s_2(t)$ as exp_1 and exp_2, respectively.

An example of calibration error model is obtained by modeling $a_i(t)$ as a mono-exponential and $b_i(t)$ as a constant, thus IG_{Si} is defined as:

$$\mathrm{IG}_{Si}(t) = m_{0a} \cdot \left[1 + \frac{m_{fa} - m_{0a}}{m_{0a}} \cdot \left(1 - e^{-\gamma_a t}\right)\right] \cdot \mathrm{IG}(t) + b_0. \tag{3.23}$$

3.3 Calibration error model selection and parameter identification

After pre-processing the data, the parameters of the model can be identified using CGM data as samples of $\mathrm{CGM}_i(t)$ and pre-processed YSI references as samples of $\mathrm{BG}(t)$. The identification is performed either in two steps or in a single one, depending on the desired approach, by using least squares. Since the model is nonlinear with respect to the parameters, a closed-form solution does not exist and we must resort to numerical iterative algorithms to find the minimum of the cost function. In this work we use MATLAB built-in functions (e.g. *lsqnonlin* and *fmincon*) that require the initial values of the parameters.

The model parameters are identified for each subject and for all the candidate calibration error models. To select the optimal calibration error model, the Bayesian Information Criterion (BIC) index of the same subject but obtained with different models are compared, and the model allowing for the lowest BIC values is chosen. The details regarding the calibration error model selection and parameter identification are presented in the following sections.

3.3.1 Least-squares parameters estimation in two steps

The model parameters can be estimated in two consecutive steps, as in [1, 2]. The first step estimates the time constant τ of Eq. (3.7) and the calibration error model parameters by using nonlinear least squares; while the second step estimates the AR model of the measurement error $v_i(t)$ from the residuals of the first step.

Initial parameters' values

As explained previously, the identification of the model parameters of each subject is performed via nonlinear least squares. To find a solution to the least square problem, iterative minimization algorithms are used. Such algorithms require to set some initial values for the parameters as starting point. Sometimes, in order to avoid local minima, the initial values have to be "sufficiently near" to the global minimum.

To make an educated guess, we draw the initial values of the parameters from probability density functions (pdf), taken from literature. If the pdfs are accurate, the least square solution provided by the iterative algorithm will converge independently from the realization of the initial values; otherwise, the algorithm falls in different local minima. When the latter situation occurs, we test 10 combinations of initial values and select the one yielding the lowest residual sum of squares (RSS).

The probability density functions of τ and the polynomials coefficients $a_0,\ a_1,\ b_0,\ b_1$ are chosen in a way to reflect the distributions of such parameters obtained in a work with previous generation Dexcom sensor [1]. Consequently, as showed in Figure 3.4, they are given by

$$\tau \sim \Gamma(k, \theta) \qquad (k = 3,\ \theta = 2.5) \tag{3.24}$$

$$a_0 \sim Log\mathcal{N}(\mu, \sigma^2) \qquad (\mu = 1.3,\ \sigma = 0.5) \tag{3.25}$$

$$a_1 \sim \mathcal{N}(\mu, \sigma^2) \qquad (\mu = 0,\ \sigma = 0.0015) \tag{3.26}$$

$$b_0 \sim \mathcal{N}(\mu, \sigma^2) \qquad (\mu = -9,\ \sigma = 55) \tag{3.27}$$

$$b_1 \sim \mathcal{N}(\mu, \sigma^2) \qquad (\mu = 0.003,\ \sigma = 0.17) \tag{3.28}$$

where Γ, $Log\mathcal{N}$ and $\mathcal{N}$ are the gamma, the log-normal and the normal density functions respectively, k and θ are the shape and the scale parameters, while μ and σ are the mean and the standard deviation.

With respect to the polynomial coefficients $a_2,\ a_3,\ b_2$ and b_3, we have no available data from the literature to model their initial distributions. However, this is not an issue since we tested different pdf models and we verified *a posteriori* that they are not required to obtain the convergence of the estimates.

Regarding the exponential parameters, we do not have any data to ex-

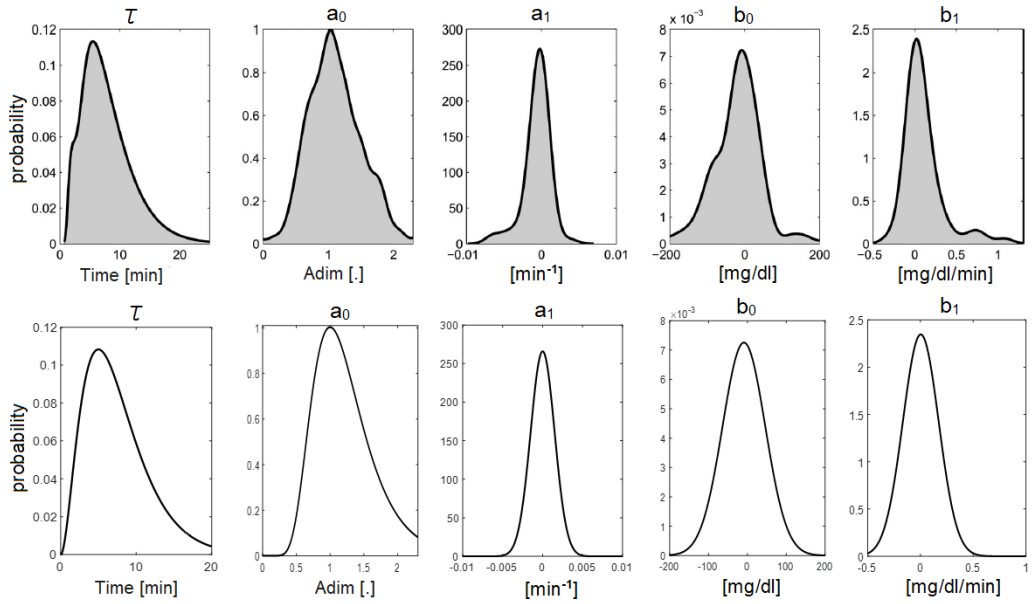

Figure 3.4: Comparison of the probability density functions of parameters $\hat{\tau}$, $\hat{a}_0$, $\hat{a}_1$, $\hat{b}_0$, and $\hat{b}_1$ obtained in a work with previous generation Dexcom sensor [1] (top) and modeled as initial distributions in this thesis (bottom).

tract their pdfs since the use of the exponential functions in calibration error models is a novelty of this thesis. Therefore, whenever possible, we model the pdfs similarly to those of the polynomial case, where the correspondences between the parameters are given by qualitatively reasoning on the asymptotic behaviors of the exponential functions. Otherwise, we base our choice on the role of the parameters in the functions. For instance, we model the pdf of γ and δ to guarantee a reasonable evolution in time of the exponential, avoiding the quasi-constant or the linear behavior.

The initial pdfs of the parameters are

$$m_0 \sim Log\mathcal{N}(\mu, \sigma^2) \qquad (\mu = 1.3,\ \sigma = 0.5) \tag{3.29}$$

$$m_f \sim \mathcal{N}(\mu, \sigma^2) \qquad (\mu = -9,\ \sigma = 55) \tag{3.30}$$

$$\gamma \sim \mathcal{U}(a, b) \qquad (a = 0.4,\ b = 0.6) \tag{3.31}$$

$$\delta \sim \mathcal{U}(a, b) \qquad (a = 0.4,\ b = 0.6) \tag{3.32}$$

$$h \sim \mathcal{U}(a, b) \qquad (a = 0.4,\ b = 0.6) \tag{3.33}$$

where $\mathcal{U}$ is the uniform density function and (a, b) is the support of the uniform distribution.

Unfortunately, the nonlinear least square algorithm provides different local minima by using these pdfs, and the solution diverge. Therefore, we operate as described above.

Identification of the BG-to-IG and calibration error model parameters

The model parameters p of the i-th subject are identified by minimizing the sum of the squared residuals (RSS) between the vector of the CGM sensor measurements Y_i and the $\mathrm{IG}_{Si}(p)$ one, which is computed from the selected model. Specifically, according to Eq. (3.1), we can write

$$Y_i = \mathrm{IG}_{Si}(\mathrm{BG}_i, p) + v_i. \tag{3.34}$$

Therefore, the estimated parameters are computed as

$$\hat{p} = \mathrm{argmin}\|Y_i - \mathrm{IG}_{Si}(\mathrm{BG}_i, p)\|^2, \tag{3.35}$$

and the residuals are

$$\hat{v}_i = Y_i - \mathrm{IG}_{Si}(\hat{\mathrm{BG}}_i, \hat{p}). \tag{3.36}$$

The minimization is computed by using the MATLAB built-in function *lsqnonlin*.

Precision of parameters' estimates

Once we have the estimated parameters, their precision can be assessed by evaluating the coefficient of variation (CV). The CV measures the dispersion of the estimates around the mean value, and it is usually expressed in percentage as

$$\mathrm{CV}(\hat{p}) = 100 \cdot \frac{sd(\hat{p})}{\hat{p}} \tag{3.37}$$

where $sd(\hat{p})$ is the standard deviation of the estimated parameters. As a consequence, a low CV corresponds to a low dispersion and a good precision of the estimated parameters.

Selection of calibration error model by Bayesian Information Criterion

When the parameters are identified for each subject and for all the candidate calibration error models, we have to choose the optimal model.

The number of parameters of the calibration error model depends on the selected function $a_i(t)$ and $b_i(t)$. In the simplest case, when $a_i(t)$ and $b_i(t)$ are both constants, the number of parameters to identify is three ($\tau,\ a_0,\ b_0$); while in the most complex one, when $a_i(t)$ and $b_i(t)$ are both bi-exponential, the number of parameters is eleven. As we consider calibration error models with different number of parameters, to select the best model we need an indicator able to take into account both the model fit and the complexity of the model. Indeed, the increase of the number of parameters improves the model fit goodness but it also increases the chances of overfitting.

To tackle the trade-off between goodness of fit and complexity we resort to the Bayesian Information Criterion (BIC) index. The BIC is given by

$$\mathrm{BIC}_{(mod,i)} = d_i \ln(\mathrm{RSS}_i) + p_{mod} \ln(d_i), \tag{3.38}$$

where d_i is the number of CGM data available for the i-th subject, p_{mod} is the number of parameters of the model mod, and RSS_i is the residual sum of squares computed as

$$\mathrm{RSS}_i = \sum_{j=1}^{d_i} \eta_{ij}^2, \tag{3.39}$$

with η_{ij} being the uncorrelated version of v_{ij}, $j \in (1, \ldots, d_i)$, samples of the measurement noise $v_i(t)$. As defined in Eq. (3.38), the BIC index considers both the goodness of the fit (first term) and the parsimony of the model (second term).

When facing the choice between two models, the one with the lower BIC should be preferred. Specifically, to compare two models for the i-th subject we define the ΔBIC as

$$\Delta\mathrm{BIC}_{(i)} = \mathrm{BIC}_{(mod_1,i)} - \mathrm{BIC}_{(mod_2,i)}, \tag{3.40}$$

where mod_2 has more parameters than mod_1. If $\Delta\mathrm{BIC}_{(i)}$ is positive, then

$\text{BIC}_{(mod_2,i)} < \text{BIC}_{(mod_1,i)}$ and we prefer mod_2, otherwise we prefer mod_1.

For choosing the best performing calibration error model we compare them pairwise. The main selection criterion between two models is to choose the model that allows a lower BIC value in the majority of subjects. When the rate of positive and negative ΔBIC values is close to 50%, we work out the dilemma by examining the outliers in the boxplot of the ΔBIC: we may prefer one model if it performs particularly better than the other for a set of subjects. In case neither the outliers are relevant to identify the best model, we follow the parsimony principle and we choose the simplest calibration error model.

Noise model parameters identification

In the second step, we use the optimal error calibration model to identify the AR model of the residuals $\hat{v}$. We expect the residuals not to be generated by a zero-mean white noise process; thus, we call them *colored residuals*. To determine the optimal order of the AR model, we explore for each residuals segment (where the segments are those described in Sec. 2.2.1), different possible orders, ranging from 1 to 10, and we compute their associated BIC. We choose the best order in each segment as the one with the lowest BIC and we aggregate the results in a histogram. The final optimal order is determined by selecting the most frequent one in the histogram.

Once we select the optimal AR model, we identify its parameters and we whiten the residuals as

$$\hat{w}(t) = \hat{v}(t) - \sum_{k=1}^{q} \hat{\alpha}_k \hat{v}(t-k), \tag{3.41}$$

where the $\hat{\alpha}_k$ and q are the AR model coefficients and order respectively.

3.3.2 Least-squares parameter estimation in a single step

In addition to the two-step method, in this thesis we implement a single step identification method, in which the parameters p of the BG-to-IG model, the error calibration model, and the measurement noise model are estimated simultaneously.

While in the two-step identification the MATLAB function *ar* always provided a stable AR model, in the single-step we follow three different procedures to deal with the issue of the stability. The first does not consider any stability constraint during the identification process, which is performed with *lsqnonlin*, but verifies *a posteriori* whether the AR model is stable. The second procedure, instead, implements the constraints by using the *fmincon* function to guarantee the stability. The last procedure, includes in the identification process *a priori* information on the parameters by using the Bayesian approach. The priors are obtained resorting to the information on the error model of the Dexcom G4AP sensor [2]. Specifically, we use the mean and standard deviation of τ, a_0 and b_0 in the fourth day of the study, when we expect the results to be more accurate. Accordingly, the *a priori* information are given by

$$\mu_\tau = 7.7, \qquad \sigma_\tau = 3, \tag{3.42}$$

$$\mu_{a_0} = 1.05, \qquad \sigma_{a_0} = 0.15, \tag{3.43}$$

$$\mu_{b_0} = -2.6, \qquad \sigma_{b_0} = 14.9, \tag{3.44}$$

where μ and σ are the mean and the standard deviation of the corresponding parameters.

To simplify the analysis, we consider only the optimal calibration error model obtained in the two-step identification and, for the first two procedures, we fix the order of the AR model to $q = 2$. Then, in the third procedure we verify the choice of the AR order exploring different possibilities ($q \in \{1, 2, 3\}$) and selecting the optimal using the Bayesian Information Criterion (BIC) similarly to Sec. 3.3.1.

Initial parameters' values

As in the two-step identification, we draw the initial values of the parameters from their pdfs. While the probability density functions for the BG-to-IG model and the calibration error model are the same of the previous analysis, we need to provide those for the AR parameters α_1, α_2 and α_3. Since we have no available data from literature, we exploit the residuals obtained from the two-step identification. In particular, we identify the AR parameters for

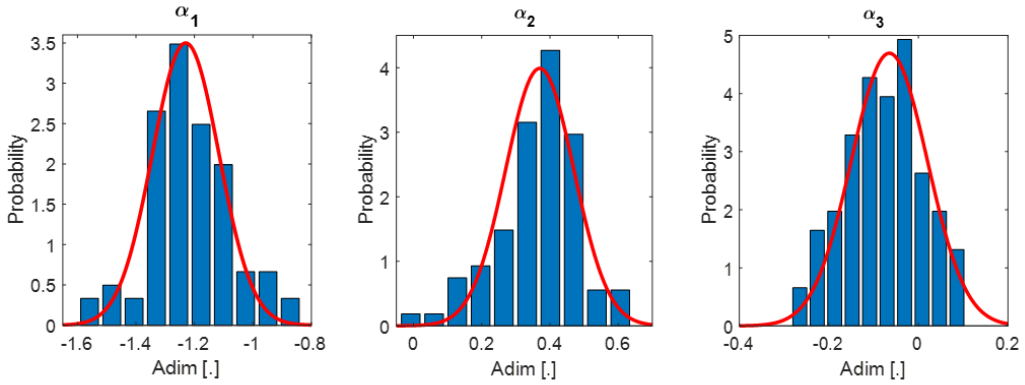

Figure 3.5: Histograms of α_1, α_2 and α_3 identified for each subject by aggregating all the segments of the residuals which refer to that subject, and relative initial probability density functions from which draw the corresponding initial values.

each subject by aggregating all the segments of the residuals which refer to that subject. As reported in Figure 3.5, the initial pdfs of α_1, α_2 and α_3 are then modeled as normal random variables, where their shapes are chosen to be similar to the corresponding histograms obtained from the AR parameters identification on subjects. Thus, the pdfs are defined as

$$\alpha_1 \sim \mathcal{N}(\mu, \sigma^2), \qquad (\mu = -1.23,\ \sigma = 0.1), \tag{3.45}$$

$$\alpha_2 \sim \mathcal{N}(\mu, \sigma^2), \qquad (\mu = 0.37,\ \sigma = 0.1), \tag{3.46}$$

$$\alpha_3 \sim \mathcal{N}(\mu, \sigma^2), \qquad (\mu = -0.047,\ \sigma = 0.09), \tag{3.47}$$

where $\mathcal{N}$ is the normal density function, while μ and σ are the mean and the standard deviation. To model the initial parameters we also have to take into account the stability of the AR model. Given the characteristic polynomial of the AR process of order q

$$\mathcal{C}(z) = 1 + \sum_{j=1}^{q} \alpha_j z^j, \tag{3.48}$$

the process is stable when the absolute values of its roots ζ are lower than one ($|\zeta| < 1$). Therefore, we draw the initial values of the AR parameters

from their pdfs and we check whether they respect the stability constraints. If the stability is verified, we use α_1, α_2 and α_3 in the identification step; otherwise, we draw other initial values until we find a stable combination of the AR parameters. Unfortunately, when we use these initial pdfs to identify the model with the single step procedure, the solution is not unique and depends on the particular realization of the initial parameters. To overcome the problem, we operate as described in Sec 3.3.1.

Identification of the model parameters

Again, the estimation of the model parameters is performed by using the nonlinear least square method but the residuals to minimize are now the differences between the CGM sensor values Y_i and the ones provided by the whole model, which includes also the AR model of the measurement noise. As a result, the residuals are expected to be a zero-mean white noise process.

According to the whole model defined in Eq. (3.1), we can write

$$Y_i = \Phi_i(p) + w_i. \tag{3.49}$$

Therefore, the estimated parameters are computed as

$$\hat{p} = \text{argmin}\|Y_i - \Phi(p)\|^2, \tag{3.50}$$

and the residuals are

$$\hat{w}_i = Y_i - \Phi(\hat{p}). \tag{3.51}$$

The minimization of (3.50) is performed using the MATLAB built-in functions *lsqnonlin* and *fmincon* and the precision of the estimates are computed exactly as in Sec. 3.3.1.

Chapter 4

Results of the identification of the new model with the two-step procedure

The identification of the new model with the two-step procedure has been carried out in two different phases. In the first phase, we develop the strategies to identify the optimal model and the corresponding parameters, and then we test them on 10 random subjects. This approach has many advantages: we have a preliminary view on how the algorithms work, we can solve more easily the problems thanks to the limited number of data, and above all, we can carry out an initial selection process, where we can discard the calibration models that perform much worse than the others. In the second phase, we extend the analysis to the entire dataset.

4.1 Calibration error model selection

To select the optimal calibration error model, we split the comparisons among models into nine groups: three comprising comparisons between polynomial models, three comprising comparisons between mono-exponential hybrid models, and three comprising comparisons between bi-exponential hybrid models. Starting from Group **I** and proceeding iteratively, we select the best model in each group and we include it in the next one.

The model resulting the best one in Group **IX** is then selected as the

Table 4.1: Composition of the polynomial groups for the calibration error model selection where m and l are the degree of the polynomials functions $a(t)$ and $b(t)$ respectively. For each group, the pairs of models are compared by using BIC criterion. Then, the models selected as best models are compared each other. The procedure is repeated until a single best model for the group is obtained. The best model of the i-th group is the initial reference model for the $(i+1)$-th group.

	Groups		
	I	**II**	**III**
Polynomial models	(m=l=0)	(m=1, l=0)	(m=2, l=0)
	vs	vs	vs
	(m=1, l=0)	(m=2, l=0)	(m=3, l=0)
	(m=l=0)	(m=2, l=0)	(m=3, l=0)
	vs	vs	vs
	(m=0, l=1)	(m=2, l=1)	(m=3, l=1)
	(m=1, l=0)	(m=2, l=0)	(m=3, l=1)
	vs	vs	vs
	(m=l=1)	(m=0, l=2)	(m=3, l=2)
	(m=0, l=1)	(m=2, l=1)	(m=3, l=2)
	vs	vs	vs
	(m=l=1)	(m=l=2)	(m=l=3)

Table 4.2: Composition of the mono-exponential groups for the calibration error model selection. $a(t)$ and $b(t)$ are polynomial or exponential functions. In case of polynomials m and l are their degrees, in case of mono-exponential functions we refer to them as exp_1. For each group, the pairs of models are compared by using BIC criterion. Then, the models selected as best models are compared each other. The procedure is repeated until a single best model for the group is obtained. The best model of the i-th group is the initial reference model for the $(i+1)$-th group.

	Groups		
	IV	**V**	**VI**
Mono-exponential hybrid models	(m=2, l=0) vs ($\exp_1$, l=0)	(m=2, l=0) vs ($\exp_1$, l=1)	(m=2, l=0) vs ($\exp_1$, l=3)
	(m=2, l=0) vs (m=0, $\exp_1$)	($\exp_1$, l=1) vs ($\exp_1$, l=2)	($\exp_1$, l=3) vs (m=3,$\exp_1$)
	($\exp_1$, l=0) vs ($\exp_1$, $\exp_1$)	($\exp_1$, l=1) vs (m=1, $\exp_1$)	/
	(m=0, $\exp_1$) vs ($\exp_1$, $\exp_1$)	($\exp_1$, l=2) vs (m=2, $\exp_1$)	/

Table 4.3: Composition of the bi-exponential groups for the calibration error model selection. $a(t)$ and $b(t)$ are polynomial or exponential functions. In case of polynomials m and l are their degrees, in case of bi-exponential functions we refer to them as exp_2. For each group, the pairs of models are compared by using BIC criterion. Then, the models selected as best models are compared each other. The procedure is repeated until a single best model for the group is obtained. The best model of the i-th group is the initial reference model for the $(i+1)$-th group.

	Groups		
	VII	**VIII**	**IX**
Bi-exponential hybrid models	(m=2, l=0) vs ($\exp_2$, l=0)	(m=2, l=0) vs ($\exp_2$, l=1)	(m=2, l=0) vs ($\exp_2$, l=3)
	(m=2, l=0) vs (m=0, $\exp_2$)	($\exp_2$, l=1) vs ($\exp_2$, l=2)	($\exp_2$, l=3) vs (m=3,$\exp_2$)
	($\exp_2$, l=0) vs ($\exp_2$, $\exp_2$)	($\exp_2$, l=1) vs (m=1, $\exp_2$)	/
	(m=0, $\exp_2$) vs ($\exp_2$, $\exp_2$)	($\exp_2$, l=2) vs (m=2, $\exp_2$)	/

optimal. To select the best model in a group, the models are sorted according to their increasing number of parameters and then arranged in adjacent pairs. Each pair is then compared by using BIC criterion, as described in Sec. 3.3.1. This procedure is repeated for the models that were selected as best models in the previous round of comparisons, until a single best model is obtained in the group. The composition of the nine groups are resumed in Tables 4.1, 4.2, 4.3.

From the preliminary analysis on the 10 subjects, we find out that the best performing calibration error model is the ($m = 2,\ l = 0$), then we expect the gain contribution $a(t)$ to be more involved than the offset contribution $b(t)$. Moreover, calibration error models including cubic order polynomials or bi-exponential functions never outperform the others in terms of BIC. This is likely because their higher complexity is not rewarded with a better fit; taking into account also the parsimony principle, we decide to remove such models from the remainder of the study, i.e., in the following analysis we are not considering Groups **III**, **VI**, **VII**, **VIII**, and **IX**.

The main results of the analysis on the entire dataset are reported in the boxplots of Figures 4.1, 4.2, 4.3, and 4.4. Regarding Group **I**, we discard ($m = l = 0$) because $\text{BIC}_{(m=1,l=0)} < \text{BIC}_{(m=l=0)}$ and $\text{BIC}_{(m=0,l=1)} < \text{BIC}_{(m=l=0)}$ in more than 60% of the cases and there are many outliers with positive ΔBIC. Similarly, we eliminate the model ($m = l = 1$) because $\Delta\text{BIC}_{((m=1,l=0),\ m=l=1)} > 0$ and $\Delta\text{BIC}_{((m=0,l=1),\ m=l=1)} > 0$ in the 31% and in the 39% of the subjects, respectively. Therefore, the final comparison for this group is between the ($m = 1,\ l = 0$) model and the ($m = 0,\ l = 1$) one. As reported in the last boxplot of Figure 4.1, $\Delta\text{BIC}_{((m=1,l=0),\ (m=0,l=1))} > 0$ in the 36% of the subjects; thus we select the polynomial model ($m = 1, l = 0$) as the optimal for the group and we insert it into the next one.

Regarding Group **II**, the BIC value of the ($m = 2, l = 0$) is lower than that of the ($m = 1, l = 0$) in the 55% of the cases, and the ΔBIC presents many outliers with positive values. For these reasons we discard the ($m = 1, l = 0$) model. Incrementing the number of parameters used to describe the offset does not lead to a better description of the data. Indeed, $\Delta\text{BIC}_{((m=2,l=0),\ (m=2,l=1))} > 0$ in the 33% of the subjects. Also considering ($m = 0, l = 2$) does not significantly improve the results, with $\text{BIC}_{(m=2,l=0)} < \text{BIC}_{(m=0,l=2)}$ in the 71% of the subjects.

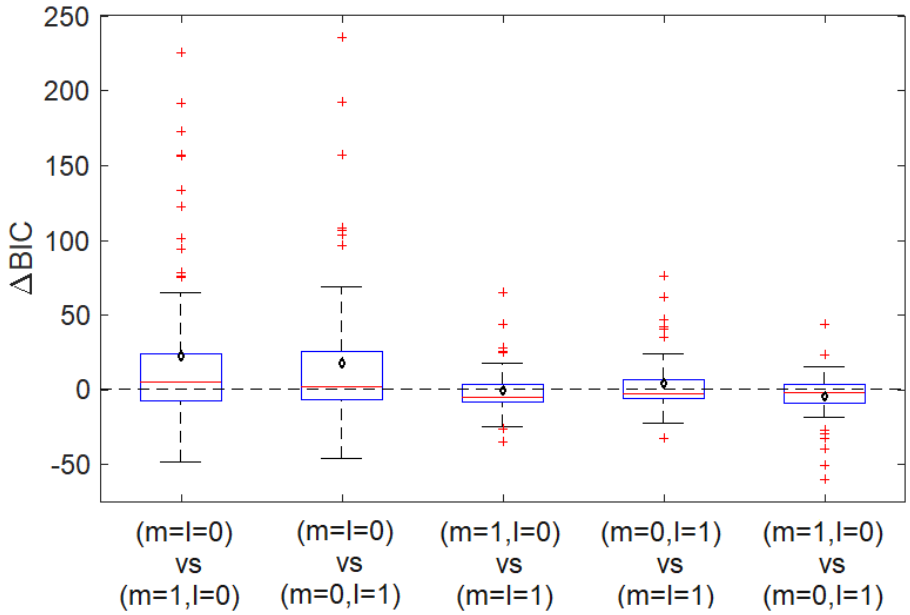

Figure 4.1: Boxplot of ΔBIC values obtained in the group **I** of polynomial models while selecting the calibration error function. The red line indicates the median value, the diamond shows the mean value, the plus are the outliers values. Note that a positive value of ΔBIC means that the model with more parameters is preferable.

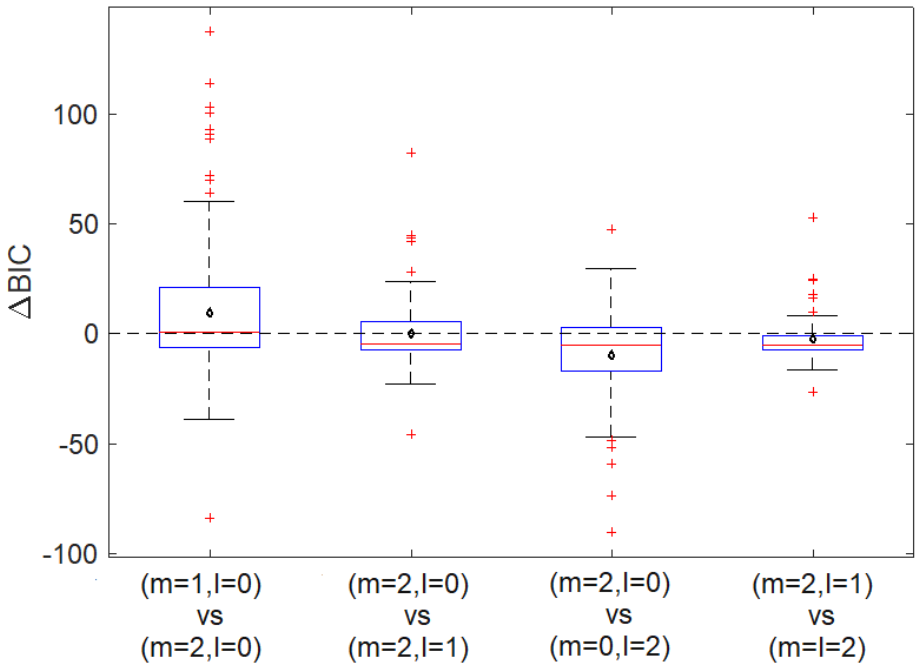

Figure 4.2: Boxplot of ΔBIC values obtained in the group **II** of polynomial models while selecting the calibration error function. The red line indicates the median value, the diamond shows the mean value, the plus are the outliers values. Note that a positive value of ΔBIC means that the model with more parameters is preferable.

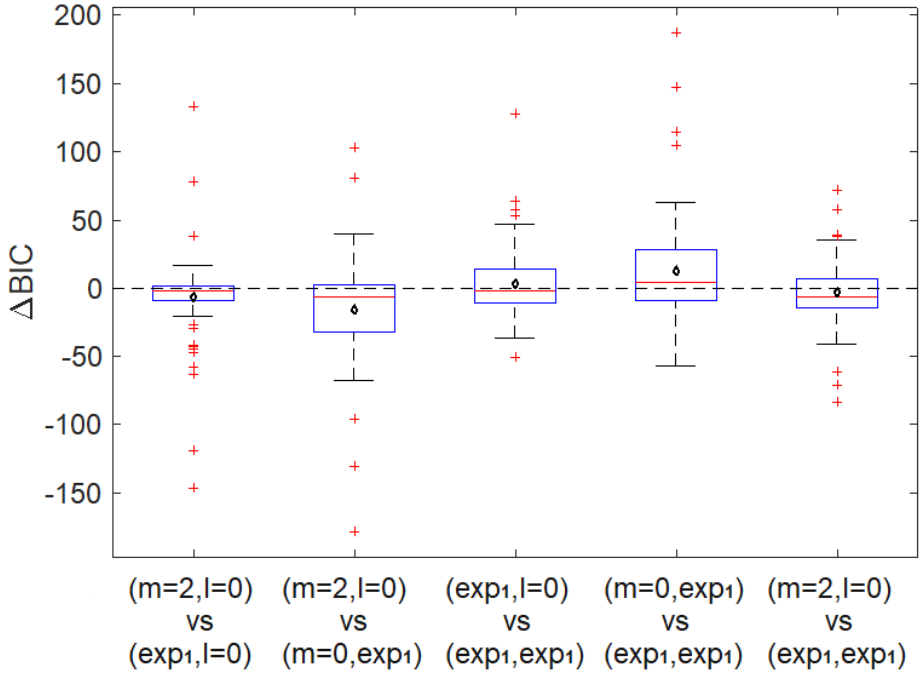

Figure 4.3: Boxplot of ΔBIC values obtained in the group **IV** of mono-exponential hybrid models while selecting the calibration error function. The red line indicates the median value, the diamond shows the mean value, the plus are the outliers values. Note that a positive value of ΔBIC means that the model with more parameters is preferable.

Finally, comparing $(m = 2, l = 1)$ with $(m = l = 2)$ yields to $\text{BIC}_{(m=l=2)} < \text{BIC}_{(m=2,l=1)}$ in the 21% of the cases; thus we discard also the model $(m = l = 2)$. Owing to these results, we select the $(m = 2, l = 0)$ model as the best model of Group **II** and we insert it into the next group.

Considering Group **IV**, where the two functions $a(t)$ and $b(t)$ can be either a polynomial or a mono-exponential, we first compare the $(m = 2, l = 0)$ with the $(\exp_1, l = 0)$ and the $(m = 0, \exp_1)$. As represented in Figure 4.3, $\text{BIC}_{(\exp_1,l=0)} < \text{BIC}_{(m=2,l=0)}$ and $\text{BIC}_{(m=0,\exp_1)} < \text{BIC}_{(m=2,l=0)}$ in the 33% and the 30% of the cases, respectively. Thus, the model $(m = 2, l = 0)$ is the one that perform better. Instead, when we compare both the $(\exp_1, l =$

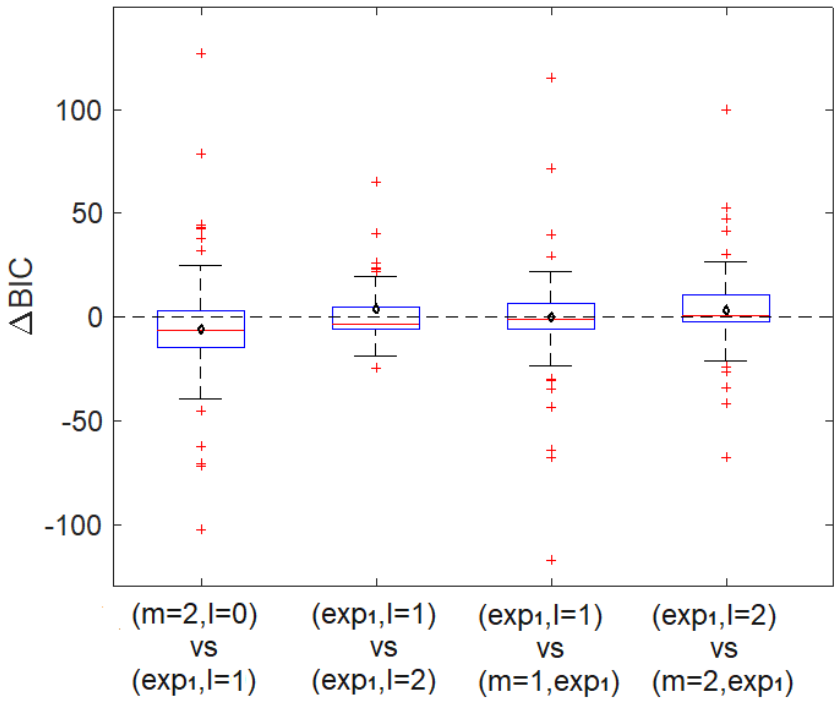

Figure 4.4: Boxplot of ΔBIC values obtained in the group **V** of mono-exponential hybrid models while selecting the calibration error function. The red line indicates the median value, the diamond shows the mean value, the plus are the outliers values. Note that a positive value of ΔBIC means that the model with more parameters is preferable.

0) and the $(m = 0, \exp_1)$ with the $(\exp_1, \exp_1)$, we obtain $\text{BIC}_{(\exp_1,\exp_1)} < \text{BIC}_{(\exp_1,l=0)}$ in the 48% of the cases and $\text{BIC}_{(\exp_1,\exp_1)} < \text{BIC}_{(m=0,\exp_1)}$ in the 54%. In both cases there are few outliers that performs better with one of the two model analysed, so we follow the parsimony principle and we discard the model $(\exp_1, \exp_1)$. From these comparisons, the polynomial model $(m = 2, l = 0)$ is the best model of Group **IV**. To validate the choice, we also directly compare the $(m = 2, l = 0)$ with the $(\exp_1, \exp_1)$, resulting in $\Delta\text{BIC}_{((m=2,l=0),\ (\exp_1,\exp_1))} > 0$ in the 38% of the subjects.

Finally, we analyze the models in Group **V**. First, we compare the model $(m = 2, l = 0)$ with the $(\exp_1, l = 1)$ one and we obtain $\text{BIC}_{(\exp_1,l=1)} < \text{BIC}_{(m=2,l=0)}$ in the 28% of the cases. Therefore, the $(m = 2, l = 0)$ is preferable. Then, as reported in Figure 4.4, we find out that increasing the

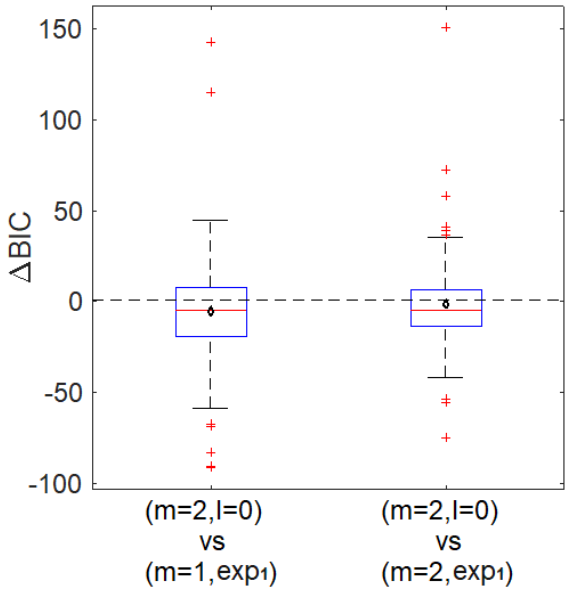

Figure 4.5: Boxplot of ΔBIC values obtained in the second round of comparisons in the group **V** of mono-exponential hybrid models while selecting the calibration error function. The red line indicates the median value, the diamond shows the mean value, the plus are the outliers values. Note that a positive value of ΔBIC means that the model with more parameters is preferable.

number of parameters is not convenient. Indeed, $\Delta\text{BIC}_{((\exp_1,l=1),\ (\exp_1,l=2))} > 0$ in the 41% of the subjects and there are few outliers with positive ΔBIC. Regarding the model $(m = 1, \exp_1)$, it is equivalent to the $(\exp_1, l = 1)$ one because $\text{BIC}_{(\exp_1,l=1)} < \text{BIC}_{(m=1,\exp_1)}$ in the 53% of the cases, the outliers are not relevant and the complexity of the model is the same (they both involve 5 parameters). Thus, the model $(m = 1, \exp_1)$ is set apart. Instead, when we compare the models $(\exp_1, l = 2)$ and $(m = 2, \exp_1)$, we obtain $\text{BIC}_{(m=2,\exp_1)} < \text{BIC}_{(\exp_1,l=2)}$ in the 59% of the cases. For this reason we discard the $(\exp_1, l = 2)$.

From these comparisons, we obtain three possible optimal models: $(m = 2, l = 0)$, $(m = 1, \exp_1)$ and $(m = 2, \exp_1)$. We sorted these models according to their increasing number of parameters and then we arranged in adjacent pairs. As represented in Figure 4.5, $\text{BIC}_{(m=1,\exp_1)} < \text{BIC}_{(m=2,l=0)}$

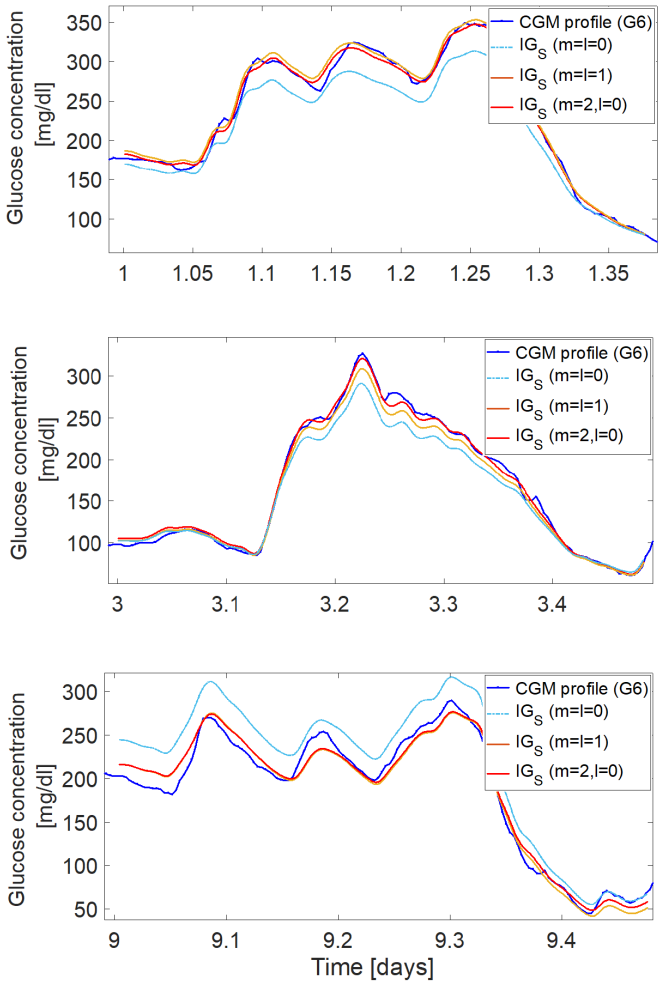

Figure 4.6: IG_{Si} curves comparison on a representative subject i with different calibration error models during day 2 (top), day 4 (center) and day 10 (bottom). The IG_{Si} curves are obtained with the model $(m = l = 0)$ (light blue lines), the model $(m = l = 1)$ (yellow lines) and the model $(m = 2, l = 0)$ (red lines).

and $\text{BIC}_{(m=2,\exp_1)} < \text{BIC}_{(m=2,l=0)}$ in the 34% and the 36% of the cases, respectively. Therefore, we discard both the models $(m = 1, \exp_1)$ and $(m = 2, \exp_1)$ and we select the polynomial model $(m = 2,\ l = 0)$ as the optimal one.

In Figure 4.6, we reported an example of how different calibration error models perform. The use of the optimal model $(m = 2, l = 0)$ (red line) allows to describe the CGM sensor profile much better than the model $(m = l = 0)$ (light blue line), which is not sufficient to explain the time-variant behavior of the calibration error. Instead, the model $(m = l = 1)$ gives a satisfactory description of the calibration error on day 2 and 10 but it worsens on day 4, meaning that a better characterization of the time variation of $a(t)$ and $b(t)$ is needed.

4.2 Parameter estimation

Once the calibration error functions have been set, we have to identify its parameters $a_0,\ a_1,\ a_2,\ b_0$ and the time constant τ for each subject. We can then compute the IG_{Si} profile with the estimated parameters.

The values and the precision of the estimates are shown in Table 4.4, where the median, the 5th and the 95th percentiles, and the percentage of the estimates with $CV < 5\%$, $CV < 10\%$, and $CV < 30\%$ are reported. Results

Table 4.4: Median, 5th and 95th percentiles values for model parameters $\hat{\tau},\ \hat{a}_0,\ \hat{a}_1,\ \hat{a}_2,\ \hat{b}_0$, and percentage of values estimated with $CV < 5\%$, $CV < 10\%$, $CV < 30\%$ with the selected calibration error model $(m = 2, l = 0)$.

Parameter	**Percentile**			**% of values estimated with**		
	50^{th}	5^{th}	95^{th}	$CV < 5\%$	$CV < 10\%$	$CV < 30\%$
$\hat{\tau}$	3.7675	0.0323	11.6446	85.2%	88.8%	90.1%
$\hat{a}_0$	0.9405	0.6398	1.1171	100%	100%	100%
$\hat{a}_1$	0.0068	-0.0657	0.1196	79%	93.8%	98.8%
$\hat{a}_2$	-0.0007	-0.0093	0.0054	79%	93.8%	95.1%
$\hat{b}_0$	7.5857	-1.2057	18.5540	81.5%	87.7%	91.4%

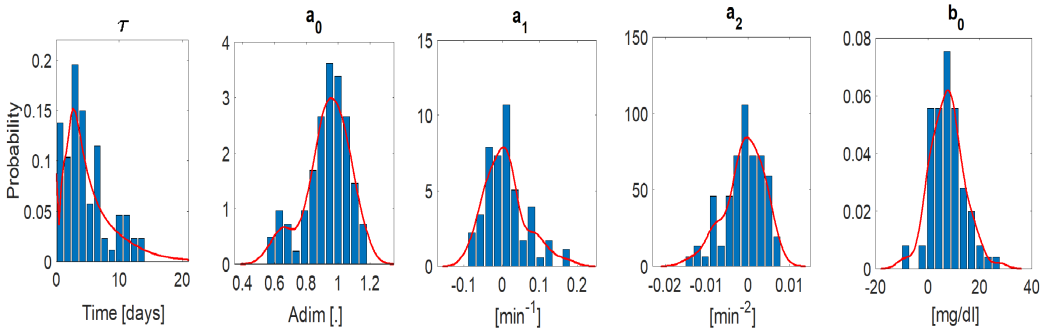

Figure 4.7: Histograms and relative probability density functions of parameters $\hat{\tau}$, $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$, and $\hat{b}_0$ obtained from estimated values and kernel density estimation procedure with the selected calibration error model ($m = 2, l = 0$).

show that the parameters are identified with great precision: the percentage of estimates with $CV < 30\%$ is always greater than 90%, while the one with $CV < 5\%$ never drops below 79%. We can make some considerations on the distribution of $\hat{\tau}$: its median value of $\simeq 3.8$ *min* is surprisingly low from a physiological point of view; indeed, from previous studies, τ is expected to span from 3 to 12 minutes, with a median around 7-8 minutes. However, the inter-individual variability of $\hat{\tau}$ is large, with values spanning from almost 0 to 14.1, confirming that it is subject specific. Regarding the parameters of the calibration error model, the values of $\hat{a}_0$, $\hat{a}_1$, and $\hat{a}_2$ are much variable among subjects. We remark that non null values of $\hat{a}_1$ and $\hat{a}_2$ correspond to a drift in time of the sensor sensitivity.

To better appreciate the parameter estimates' distributions, we assess both their histograms and their probability density functions that are obtained through a kernel estimation procedure. Similar to the histogram, the kernel estimation process builds a function to represent the probability distribution of the parameters. But unlike the histogram, which places the values of parameters into discrete bins, the kernel distribution sums the component smoothing functions for each value to produce a smooth, continuous probability curve. Having continuous pdfs can be useful in several applications, like generating random values for the parameters in a simulation environment. From Figure 4.7, we verify that the resulting distributions are consistent with the ones observed in [1, 2], apart from the distribution of $\hat{\tau}$. Specifically, in the analysis we obtain $\hat{\tau} \simeq 0$ for a certain number of subjects; this value is

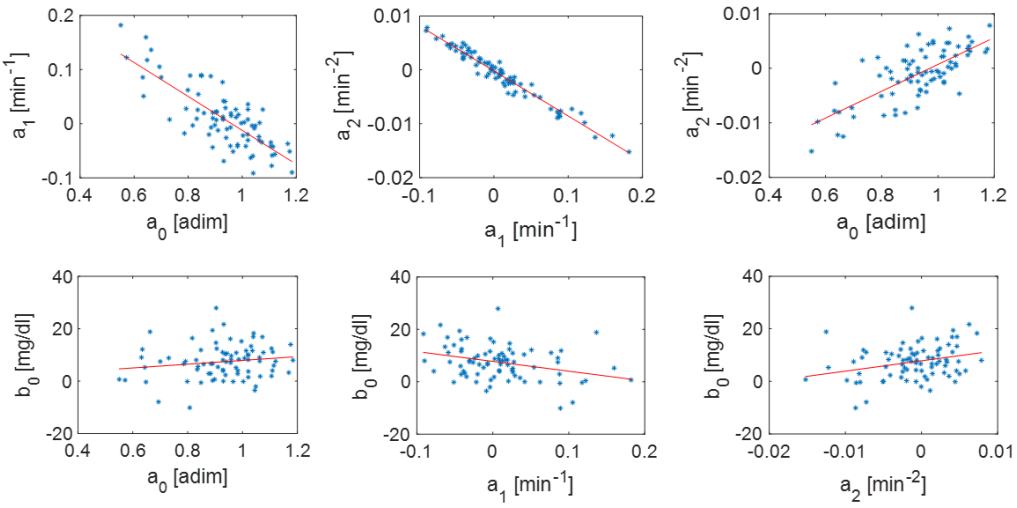

Figure 4.8: Scatterplots between pairs of parameters of the selected calibration error model. The boxes on the top represent the scatterplots between $a(t)$ parameters a_0, a_1 and a_2, while the boxes on the bottom represent the scatterplots between a(t) parameters and b_0.

not realistic because the glucose diffusion from blood to interstitial fluid is not instantaneous. A deeper insight of the problem is described in the next section.

After assessing the parameters values, we plot their scatterplots to study their correlations. In Figure 4.8 we study the correlation among the calibration error model parameters. In particular, to verify whether a simple linear relationship is sufficient to describe their relation, we introduce the coefficient of Pearson ρ. Such coefficient measures the strength and the direction of the linear relationship between two variables $\hat{x}$ and $\hat{y}$, and it is defined as

$$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} \qquad \rho_{x,y} \in [-1,1] \tag{4.1}$$

where $cov(x,y)$ is the covariance, σ_x is the standard deviation of x, and σ_y is the standard deviation of y.

When $|\rho_{x,y}| = 1$ the relation between x and y is perfectly described by a linear equation, and we refer to it as a "perfect correlation"; whereas, when $\rho_{x,y} = 0$ there is no linear correlation between the variables. Moreover,

$\rho_{x,y} > 0$ indicates a positive correlation, while $\rho_{x,y} < 0$ indicates a negative one. Observing Figure 4.8 (top), we can see strong negative correlations between $\hat{a}_1$ and $\hat{a}_2$ ($\rho_{\hat{a}_1\hat{a}_2} = -0.98$) and between $\hat{a}_0$ and $\hat{a}_1$ ($\rho_{\hat{a}_0\hat{a}_1} = -0.79$), while there is a strong positive correlation between $\hat{a}_0$ and $\hat{a}_2$ ($\rho_{\hat{a}_0\hat{a}_2} = 0.73$). On the contrary, there is not a linear relationships between the coefficients of the sensor gain and the one of the sensor offset (Figure 4.8 (bottom)). This result is confirmed by the low values of the Pearson's correlation coefficients: $\rho_{\hat{a}_0\hat{b}_0} = 0.16$, $\rho_{\hat{a}_1\hat{b}_0} = -0.32$, and $\rho_{\hat{a}_2\hat{b}_0} = 0.29$.

4.2.1 Problem in τ estimation

In this section we discuss about the problem in estimating the diffusion time constant we encountered in the previous section for few subjects. In particular, eight subjects have an estimated time constant almost equal to zero, and this is not consistent with the physiology of the BG-to-IG kinetic. The CVs obtained for these subjects are generally low but except the CV values referring to the diffusion time constant, which have very high values. This indicates a low accuracy in the estimation of τ only. Identifying the causes of this problem is not trivial: we can speculate there may be some errors in the recording of the temporal CGM/YSI data for these subjects. The presence of possible errors in the data acquisition suggests to discard *a posteriori* these subjects from the analysis.

After the elimination, the obtained histogram (right panel in Figure 4.9) has no spurious peak around zero, and it is now in line with the one reported in [2]. The new estimated parameters and their CVs are reported in Table 4.5. The results are improved: the median value of $\hat{\tau}$ increases from $\simeq 3.77$ to $\simeq 3.98$, while the percentage of subjects with the CV referred to $\hat{\tau}$ below the 30% changes from 90.1% to 100%. Finally, regarding the other parameters, there are not significant changes both in their values and in their distributions (Figure 4.10).

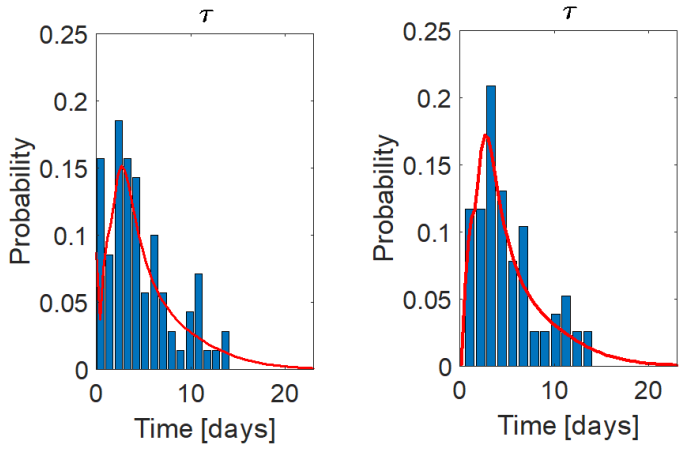

Figure 4.9: Histogram and relative probability density function of $\hat{\tau}$ obtained from estimated values and kernel density estimation procedure before (left) and after (right) the elimination of subjects with $\hat{\tau} \simeq 0$, with the selected calibration error model $(m = 2, l = 0)$.

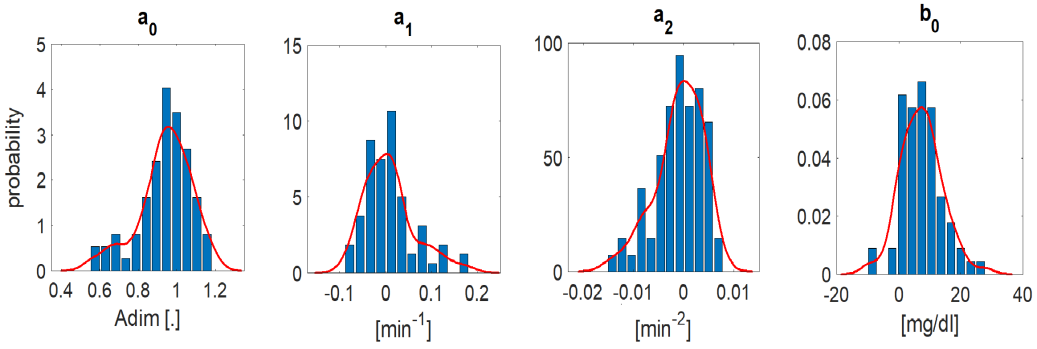

Figure 4.10: Histograms and relative probability density functions of parameters $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$, and $\hat{b}_0$ obtained from estimated values and kernel density estimation procedure after the elimination of subjects with $\hat{\tau} \simeq 0$ with the selected calibration error model $(m = 2, l = 0)$.

Table 4.5: Median, 5th and 95th percentiles values for model parameters $\hat{\tau}$, $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$, $\hat{b}_0$, and percentage of values estimated with $CV < 5\%$, $CV < 10\%$, $CV < 30\%$ after the elimination of subjects with $\hat{\tau} \simeq 0$ with the selected calibration error model ($m = 2, l = 0$).

Parameter	Percentile			% of values estimated with		
	50^{th}	5^{th}	95^{th}	$CV < 5\%$	$CV < 10\%$	$CV < 30\%$
$\hat{\tau}$	3.9755	1.1289	11.9837	94.5%	98.6%	100%
$\hat{a}_0$	0.9559	0.6504	1.1198	100%	100%	100%
$\hat{a}_1$	0.0017	-0.0624	0.1215	83.6%	93.1%	98.6%
$\hat{a}_2$	-0.0004	-0.0097	0.0050	82.2%	93.1%	94.5%
$\hat{b}_0$	7.4000	-1.7903	18.5788	79.4%	86.3%	90.4%

4.2.2 Trends of the sensors gain

The optimal calibration error model describes the gain via a quadratic function. Therefore, depending on the values assumed by the coefficients a_0, a_1 and a_2, $a(t)$ evolves differently in time. It is interesting to understand how $a(t)$ varies among subjects to include this information in CGM sensors so to improve their calibration algorithms. In particular, we can distinguish the subjects on the basis of the concavity of $a(t)$, which can be upward or downward depending on the sign of a_2.

We find that 32 subjects out of 73 have an upward concavity ($a_2 > 0$), while the remaining 41 subjects have a downward concavity ($a_2 < 0$). Thus, we conclude that there is no prevalence between the two forms. In Figure 4.11 we report the spaghetti plot (to the left) and variability bands (to the right) for the two concavity groups separately. The spaghetti plots highlight two main behaviours: quasi-linear, either increasing or decreasing, which has a_2 almost equal to zero, and quadratic. Finally, we compute the variability bands at 90%. The results show that they are similarly wide, about 0.4, and the median curve is in both cases quasi-constant.

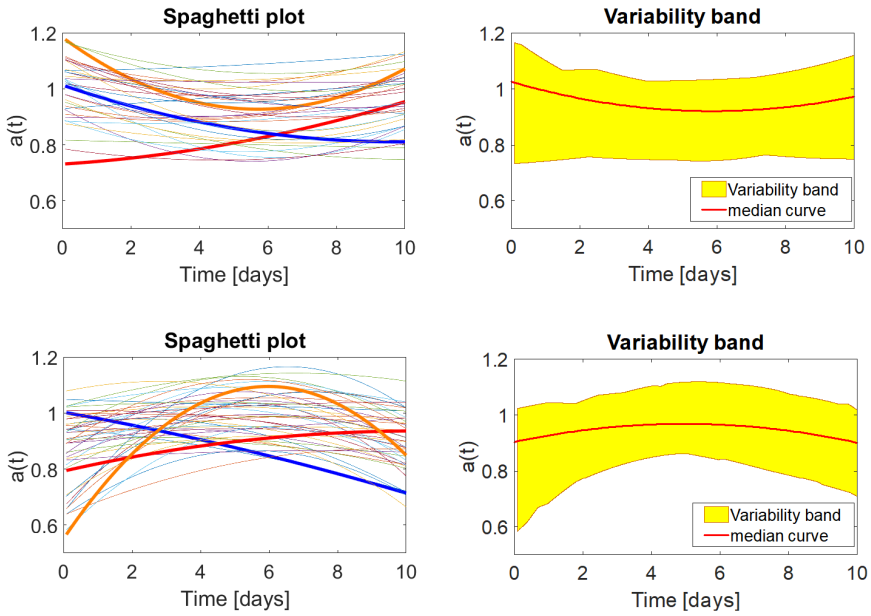

Figure 4.11: Spaghetti plots and variability bands of $a(t)$ among subjects with concavity upward $a_2 > 0$ (top), and with concavity downward $a_2 < 0$ with the selected calibration error model ($m = 2, l = 0$). The orange lines represent the quadratic behaviour of $a(t)$, while the red lines and the blue lines represent the quasi-linear behaviour, respectively increasing and decreasing.

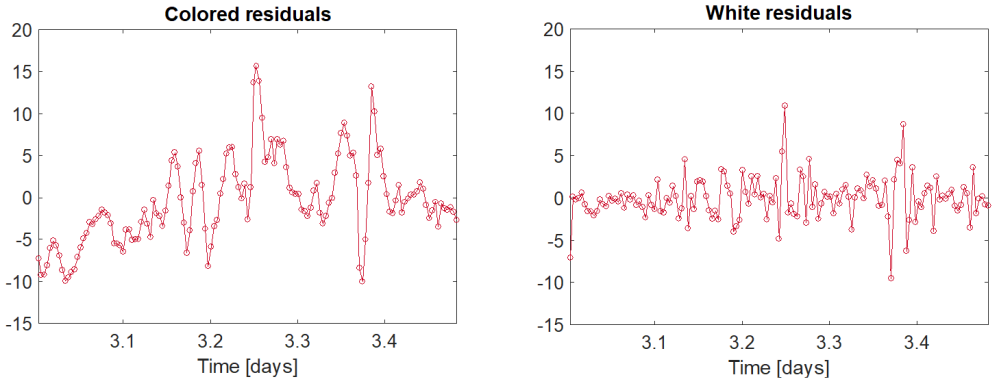

Figure 4.12: Colored residual profile (left) and corresponding white residual profile (right) of a representative subject obtained with the selected calibration error model ($m = 2, l = 0$).

4.3 Auto-regressive model identification

Once the diffusion constant τ and the parameters of the calibration error models are identified, we can obtain the IG_{Si} profile and compute the residuals with respect to the CGM sensor data as explain in Sec. 3.3.1. As an example, Figure 4.12 (left) represents the colored residuals $\hat{v}$, in a specific time window for a random subject, where we can clearly distinguish the auto-regressive component, while Figure 4.12 (right) shows the corresponding whitened residuals obtained as in Eq. (3.41) in Sec. 3.3.1.

It is important to note that while the different coefficients of the AR model are estimated for each specific residuals segment, the order q used is the same for all the segments. To find the optimal order we follow the procedure specified in Sec. 3.3.1.

4.3.1 Optimal order selection

Figure 4.13 represents the histogram with the counts of segments in which the specific order has been selected as the best. As we can see, the most frequent order is 2. This result is consistent with the one in [2] and suggests that a low order is sufficient to describe the auto-regressive component of the measurement noise.

Table 4.6: Median, 5th and 95th percentiles values for the auto-regressive model parameters $\hat{\alpha_1}$ and $\hat{\alpha_2}$ obtained with the selected order 2 of the AR model .

Parameter	**Percentile**		
	50^{th}	5^{th}	95^{th}
$\hat{\alpha}_1$	-1.2766	-1.5762	-0.8441
$\hat{\alpha}_2$	0.4339	0.0387	0.6905

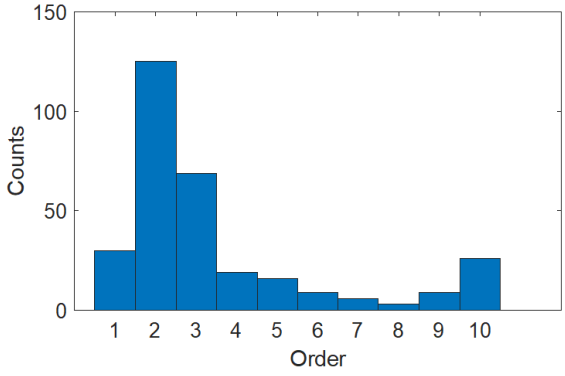

Figure 4.13: Histogram of AR orders. It reports the counts of segments in which the specific order has been selected as optimal.

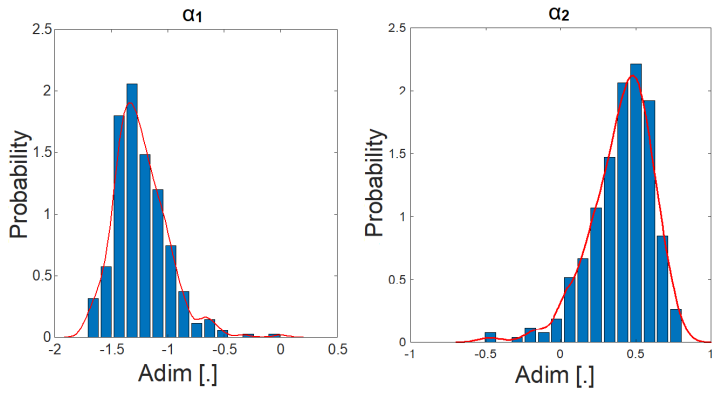

Figure 4.14: Histograms and relative probability density functions obtained from estimated values and kernel density estimation procedure for the auto-regressive model parameters $\hat{\alpha_1}$ and $\hat{\alpha_2}$.

The output of the AR model is then

$$v_{ji}(t) = \hat{\alpha}_{1ji}\ v_{ji}(t-1) + \hat{\alpha}_{2ji}\ v_{ji}(t-2) + w_{ji}(t) \tag{4.2}$$

where $\hat{\alpha}_{1ji}$ and $\hat{\alpha}_{2ji}$ are the model parameters for the j-th segment of the i-th subject and $w_{ji}(t)$ is the zero-mean white noise process.

Table 4.6 shows the results of the AR model parameters identification, where the median, the 5th, the 95th percentiles values are reported. Furthermore, both the histogram and the relative probability density function derived from the kernel density estimation procedure are represented in Figure 4.14.

4.3.2 Analysis of the auto-regressive model parameters for different days of monitoring

To better understand the evolution of the sensor error in time, we analyse the auto-regressive component in the three time windows of YSI recording, namely day 1 or 2, 4 and 10. To select the optimal order we operate using the same procedure of the previous section but arranging the segments according to the time window they belong. The corresponding histograms are reported in Figure 4.15 for the three different time windows. We can see that the order 2 is the optimal for all days; however, the distributions vary among the time windows: in the first days order 2 clearly dominates the others, while in the last day the gap between order 2 and 3 is smaller. A possible explanation is that a degradation in time of sensor performances leads to the need of a more complex model near the end of the sensor lifetime.

Once the optimal order as been selected for each time window, we estimate the AR parameters as in previous section. Table 4.7 summarizes the results of the parameters identification, reporting the median, the 5th and the 95th percentiles values of $\hat{\alpha_1}$, $\hat{\alpha_2}$ and the variance of the estimated random noise $\hat{w}(t)$ $\hat{\sigma}^2$. Regarding $\hat{\alpha_1}$ and $\hat{\alpha_2}$, there are no evident differences among the time windows. Conversely, $\hat{\sigma}^2$ is greater in the first and in the last day because of the uncertainty on the sensor measurements after its insertion and close to the end of its lifetime. This confirms our expectations since in the first period, after the insertion, a greater immune system response can lead to

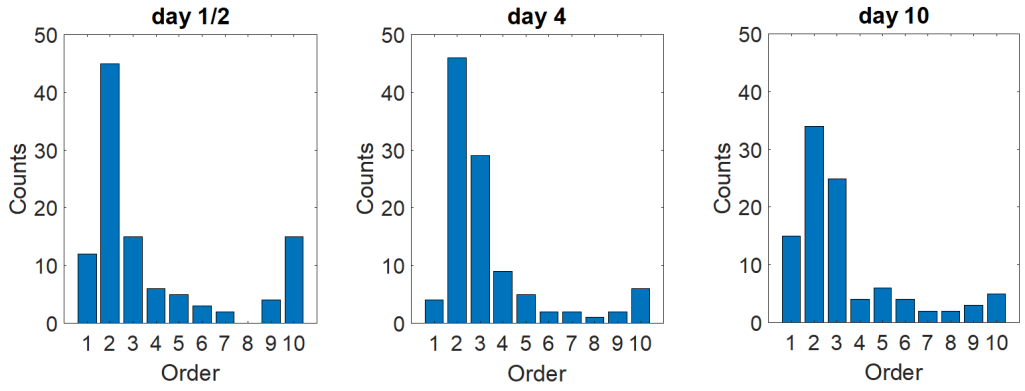

Figure 4.15: Histogram of AR orders in in days 1 or 2 (left), 4 (center), 10 (right). Each histogram reports the counts of segments in which the specific order has been selected as optimal.

instabilities in the measurements, while in the last period a degradation of the sensor performances can occur due to the closeness of the sensor end of life. To complete the analysis we also plot the histograms of $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\sigma}^2$ in Figures 4.16, 4.17, 4.18 respectively. The same considerations made for the values reported in Table 4.7 still hold for the histograms of Figures 4.16, 4.17, 4.18.

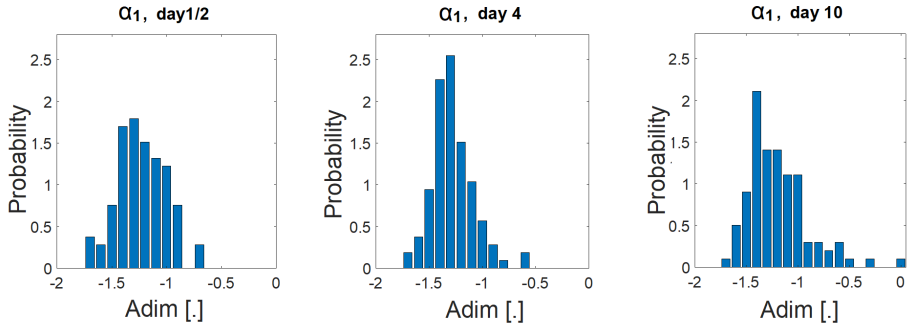

Figure 4.16: Histograms of parameter $\hat{\alpha_1}$ obtained from estimated values in days 1 or 2 (left), 4 (center), 10 (right).

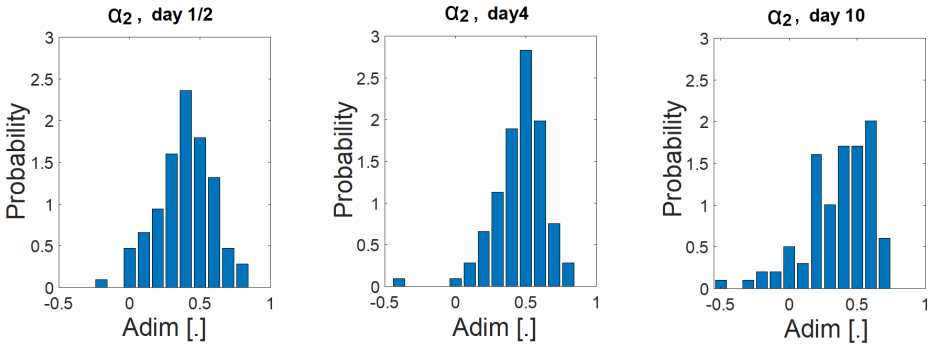

Figure 4.17: Histograms of parameter $\hat{\alpha_2}$ obtained from estimated values in days 1 or 2 (left), 4 (center), 10 (right).

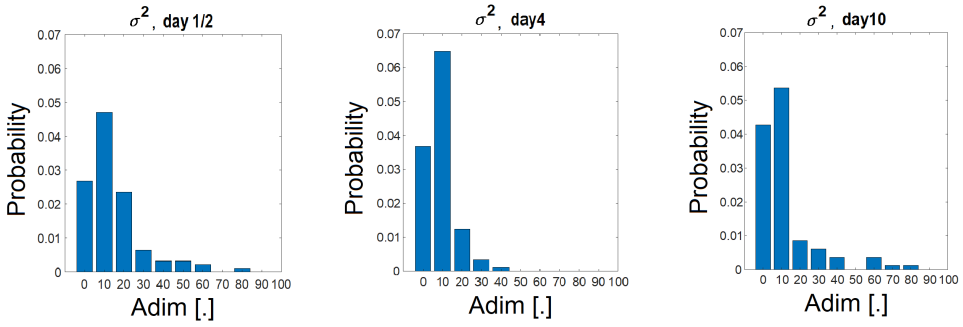

Figure 4.18: Histograms of parameter $\hat{\sigma}^2$ obtained from estimated values in days 1 or 2 (left), 4 (center), 10 (right).

Table 4.7: Median, 5th and 95th percentiles values for model parameters $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\sigma}^2$ in days 1 or 2 , 4 , 10 obtained with the selected order 2 of the AR model.

Parameter	**Days**	**Percentile**		
		50^{th}	5^{th}	95^{th}
$\hat{\alpha}_1$	1/2	-1.2405	-1.5797	-0.8718
	4	-1.3031	-1.5693	-0.9349
	10	-1.2519	-1.5811	-0.6317
$\hat{\alpha}_2$	1/2	0.4038	0.0375	0.6862
	4	0.4802	0.1524	0.7207
	10	0.4062	-0.0992	0.6669
$\hat{\sigma}^2$	1/2	10.2654	1.6031	46.6246
	4	7.0339	2.4440	21.8484
	10	7.5690	2.1049	63.7126

Chapter 5

Results of the identification of the new model with the single-step procedure

The single-step identification analysis has been accomplished with three alternative procedures considering only the optimal calibration error model ($m = 2, l = 0$) and the optimal AR model of order 2 obtained in the two-step analysis. The first method identifies the model parameters without considering any stability constraint for the AR model, and verifies *a posteriori* whether the model is stable. In the second procedure, we include the constraints in the identification process to guarantee the AR model stability. Finally, we repeat the parameters identification by including some priors on the parameters distributions and we verify the choice of the AR order $q = 2$ exploring different orders ($q \in (1, 2, 3)$).

5.1 Parameter estimation

We identify for each subject the parameters τ, a_0, a_1, a_2, b_0, α_1 and α_2 by using the three alternative procedures, as described in Sec. 3.3.2.

The values and the precision of the estimates achieved with the first procedure are shown in Table 5.1, where the median, the 5th and the 95th percentiles, and the percentage of the estimates with $CV < 5\%$, $CV < 10\%$, and $CV < 30\%$ are reported. Not all the parameters have a great precision:

Table 5.1: Median, 5th and 95th percentiles values for model parameters $\hat{\tau}$, $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$, $\hat{b}_0$, $\hat{\alpha}_1$, $\hat{\alpha}_2$, and percentage of values estimated with $CV < 5\%$, $CV < 10\%$, $CV < 30\%$ obtained in the single-step analysis without constraints.

Parameter	Percentile			% of values estimated with		
	50^{th}	5^{th}	95^{th}	$CV < 5\%$	$CV < 10\%$	$CV < 30\%$
$\hat{\tau}$	4.7897	0.9366	12.2352	54.3%	96.3%	100%
$\hat{a}_0$	0.9442	0.6047	1.1155	97.5%	97.5%	100%
$\hat{a}_1$	0.0060	-0.0679	0.1156	16%	34.5%	74%
$\hat{a}_2$	-0.0006	-0.0108	0.0053	14.8%	39.5%	75.3%
$\hat{b}_0$	6.4406	-4.0155	22.55	11.1%	39.5%	72.8%
$\hat{\alpha}_1$	-1.2770	-1.5252	-0.9517	100%	100%	100%
$\hat{\alpha}_2$	0.4049	0.1005	0.5816	63%	92.6%	100%

in particular, $\hat{\tau}$, $\hat{a}_0$, $\hat{\alpha}_1$ and $\hat{\alpha}_2$ have a very good accuracy with $CV < 30\%$ for all the subjects, while $\hat{a}_1$, $\hat{a}_2$ and $\hat{b}_0$ are not so accurate. Indeed, the percentage of the estimates with $CV < 30\%$ is greater than the 70% and the one with $CV < 10\%$ assumes values around the 30-40%. Regarding the values of the parameters, we see that $\hat{\tau}$ preserves the inter-individual variability spanning from almost 1 to 12.2 minutes. Moreover, its median value is $\simeq 4.8$ *min*, almost a minute greater than the one of the two steps analysis and consequently more physiologically meaningful. In this approach, we verify the stability of the AR model *a posteriori*, finding out that the identified model is stable for each subject.

Even if the first procedure always provides stable models, it is convenient to guarantee the stability in the estimation process without checking it *a posteriori*. Details about such procedure have been described in Sec. 3.3.2. The obtained values and the precision of the estimates are shown in Table 5.2 where the median, the 5th and the 95th percentiles, and the percentage of the estimates with $CV < 5\%$, $CV < 10\%$, and $CV < 30\%$ are reported. We can see that the CVs are the same of the previous procedure and also the estimates are very similar; thus, the introduction of stability constraints does not affect significantly the model parameters estimation.

Table 5.2: Median, 5th and 95th percentiles values for model parameters $\hat{\tau}$, $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$, $\hat{b}_0$, $\hat{\alpha}_1$, $\hat{\alpha}_2$, and percentage of values estimated with $CV < 5\%$, $CV < 10\%$, $CV < 30\%$ obtained in the single-step analysis with constraints.

Parameter	Percentile			% of values estimated with		
	50^{th}	5^{th}	95^{th}	$CV < 5\%$	$CV < 10\%$	$CV < 30\%$
$\hat{\tau}$	4.8113	1.2530	12.6564	54.3%	96.3%	100%
$\hat{a}_0$	0.9442	0.6337	1.1425	97.5%	97.5%	100%
$\hat{a}_1$	0.0060	-0.0679	0.1080	16%	34.5%	74%
$\hat{a}_2$	-0.0006	-0.0099	0.0057	14.8%	39.5%	75.3%
$\hat{b}_0$	6.2777	-5.6133	22.5517	11.1%	39.5%	72.8%
$\hat{\alpha}_1$	-1.2770	-1.5465	-0.9517	100%	100%	100%
$\hat{\alpha}_2$	0.4049	0.1005	0.6038	63%	92.6%	100%

Therefore, to improve the precision of the estimates, we include in the identification process *a priori* information by using the Bayesian approach as defined in Sec 3.3.2. The values and the precision obtained with the Bayesian estimation are shown in Table 5.3, where the median, the 5th and the 95th percentiles, and the percentage of the estimates with $CV < 5\%$, $CV < 10\%$, and $CV < 30\%$ are reported. The inclusion of priors leads to slight changes in the parameters' accuracy. The CVs of τ are improved; indeed, the percentage of the estimates with $CV < 10\%$ changes from the 96.3% to the 97.5%, and the one with $CV < 5\%$ increases from the 54.3% to the 56.8%. However, the accuracy of b_0 worsens, suggesting that the inclusion of its prior does not give the expected results. By repeating the analysis without adding the prior in b_0, the percentage of the estimates with $CV < 30\%$ slightly increases from the 71.6% to the 74%, as reported in Table 5.4 but, the overall improvement with the introduction of priors is small, suggesting that the priors are not so effective. Actually, we could expect these results because including priors is especially useful when the data are "poor" or "very noisy", and our dataset does not belong to neither of the two cases.

Table 5.3: Median, 5th and 95th percentiles values for model parameters $\hat{\tau}$, $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$, $\hat{b}_0$, $\hat{\alpha}_1$, $\hat{\alpha}_2$, and percentage of values estimated with $CV < 5\%$, $CV < 10\%$, $CV < 30\%$ obtained in the single-step analysis with constraints and prior on τ, a_0, b_0.

Parameter	**Percentile**			**% of values estimated with**		
	50^{th}	5^{th}	95^{th}	$CV < 5\%$	$CV < 10\%$	$CV < 30\%$
$\hat{\tau}$	4.8645	1.2815	12.5189	56.8%	97.5%	100%
$\hat{a}_0$	0.9448	0.6368	1.1424	97.5%	97.5%	100%
$\hat{a}_1$	0.0060	-0.0679	0.1068	16%	33.3%	72.8%
$\hat{a}_2$	-0.0006	-0.0099	0.0057	14.8%	39.5%	75.3%
$\hat{b}_0$	6.1974	-5.6324	22.4371	11.1%	39.5%	71.6%
$\hat{\alpha}_1$	-1.2770	-1.5472	-0.9517	100%	100%	100%
$\hat{\alpha}_2$	0.4051	0.1005	0.6040	63%	92.6%	100%

Table 5.4: Median, 5th and 95th percentiles values for model parameters $\hat{\tau}$, $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$, $\hat{b}_0$, $\hat{\alpha}_1$, $\hat{\alpha}_2$, and percentage of values estimated with $CV < 5\%$, $CV < 10\%$, $CV < 30\%$ obtained in the single-step analysis with constraints and prior on τ and a_0.

Parameter	**Percentile**			**% of values estimated with**		
	50^{th}	5^{th}	95^{th}	$CV < 5\%$	$CV < 10\%$	$CV < 30\%$
$\hat{\tau}$	4.8442	1.2818	12.5131	56.8%	97.5%	100%
$\hat{a}_0$	0.9445	0.6359	1.1424	97.5%	97.5%	100%
$\hat{a}_1$	0.0060	-0.0677	0.1067	16%	33.3%	72.8%
$\hat{a}_2$	-0.0006	-0.0099	0.0057	14.8%	39.5%	75.3%
$\hat{b}_0$	6.2187	-5.6863	22.5278	11.1%	39.5%	74%
$\hat{\alpha}_1$	-1.2770	-1.5470	-0.9517	100%	100%	100%
$\hat{\alpha}_2$	0.4051	0.1004	0.6042	63%	92.6%	100%

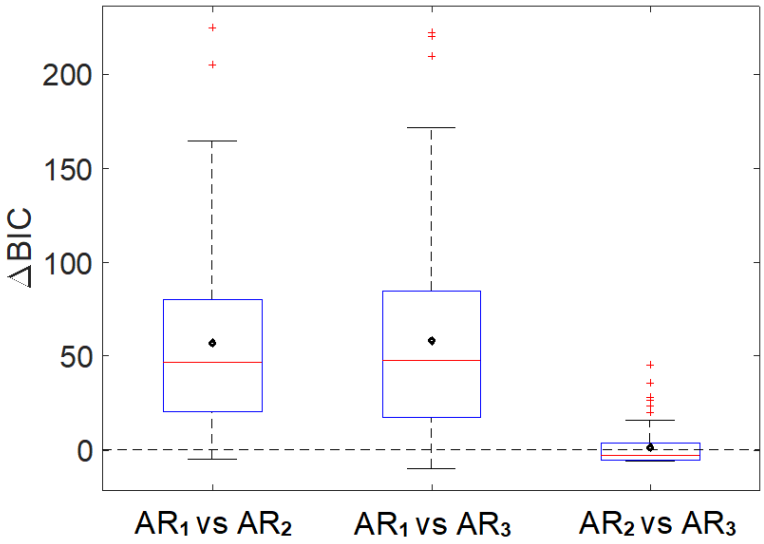

Figure 5.1: Boxplots of comparisons of the AR models obtained in the single-step analysis with constraints and prior on τ and a_0. The red line indicates the median value, the diamond shows the mean value, the plus are the outliers values.

5.2 Auto-regressive model selection

As explained above, we explore different orders for the AR model of the measurement noise. To select the optimal one we apply the procedure described in Sec. 3.3.1, where the ΔBIC are defined as

$$\Delta\text{BIC}_{(\text{AR}_1,\text{AR}_2)} = \text{BIC}_{(\text{AR}_1)} - \text{BIC}_{(\text{AR}_2)}, \tag{5.1}$$

$$\Delta\text{BIC}_{(\text{AR}_1,\text{AR}_3)} = \text{BIC}_{(\text{AR}_1)} - \text{BIC}_{(\text{AR}_3)}, \tag{5.2}$$

$$\Delta\text{BIC}_{(\text{AR}_2,\text{AR}_3)} = \text{BIC}_{(\text{AR}_2)} - \text{BIC}_{(\text{AR}_3)}, \tag{5.3}$$

where AR_1, AR_2 and AR_3 refer to the auto-regressive model of order 1, 2 and 3 respectively. Results show that $\Delta\text{BIC}_{(AR_1,AR_2)} > 0$ in the 77% of subjects and $\Delta\text{BIC}_{(AR_1,AR_3)} > 0$ in the 71% ; thus we discard the order 1. The final selection is between the orders 2 and 3, where $\Delta\text{BIC}_{(AR_2,AR_3)} > 0$ in the 30% of subjects. Due to this percentage, we are tempted to select the order $q = 2$ but, looking at the boxplots reported in Figure 5.1, we can see that

Table 5.5: Median values of error model parameters and percentage of values estimated with $CV < 30\%$, obtained with the two-step identification and the single-step identification (with constraints and prior on τ and a_0).

Parameter	50^{th} **Percentile**		**% of values estimated with CV<30%**	
	Two steps identification	***Single step identification***	***Two steps identification***	***Single step identification***
$\hat{\tau}$	3.9755	4.8442	100%	100%
$\hat{a}_0$	0.9559	0.9445	100%	100%
$\hat{a}_1$	0,0017	0,0060	98.6%	72.8%
$\hat{a}_2$	-0.0004	-0.0006	94.5%	75.3%
$\hat{b}_0$	7.4000	6.2187	90.4%	74%
$\hat{\alpha}_1$	-1.2234	-1.2770	/	100%
$\hat{\alpha}_2$	0.3827	0.4051	/	100%

neither of the two model is dominant. Taking into account both these results and following the parsimony principle, we select 2 as the final order of the model. Thus, the optimal order obtained with the single step and the two steps procedure is the same, highlighting the consistency of the results.

5.3 Comparison to two-steps parameters estimation

The two-step identification allows to separate the complex identification problem into two simpler problems with a lower number of parameters estimated. However, as the parameters are estimated sequentially and not simultaneously, this may introduce bias in the parameters' estimates. The problem should be overcome by the single-step identification method, although the estimation of a large number of parameters simultaneously can be difficult for some sensors. Therefore, we compare the results of the two methods to verify the soundness of the results obtained with them.

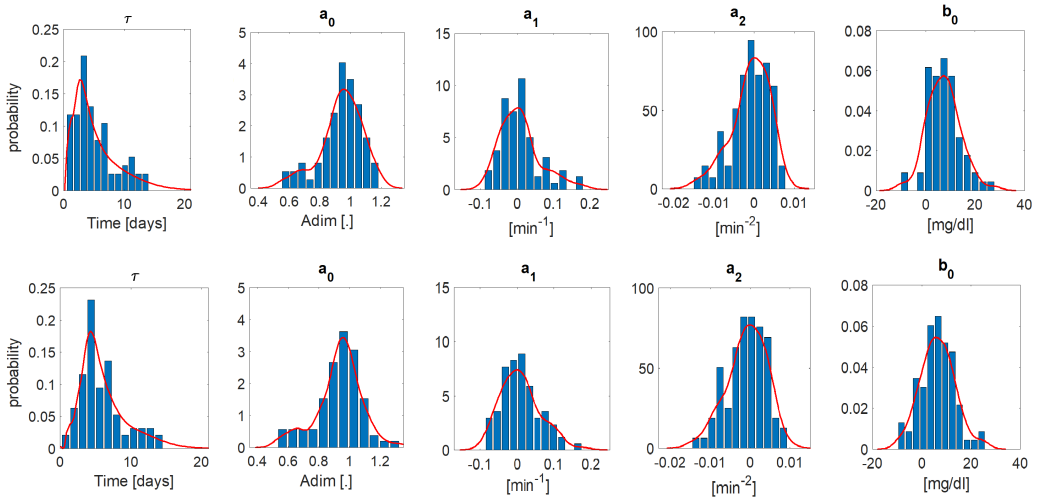

Figure 5.2: Histograms and relative probability density functions of parameters $\hat{\tau}$, $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$, and $\hat{b}_0$ obtained from estimated values and kernel density estimation procedure in the two-step identification (above) and in the single-step identification (below).

First, we compare the median values of the model parameters and the percentage of the values estimated with $CV < 30\%$. They are reported in Table 5.5, where the $\hat{\alpha}_1$ and the $\hat{\alpha}_2$ referring to the two-step identification are obtained by aggregating all the segments of each subject. Regarding the median values, the estimates given by the two methods are very close to each other; thus, the single-step identification represents a reliable alternative for estimating the sensor error model parameters. As explained in the previous section, the median value of $\hat{\tau}$ in the single-step analysis is almost a minute greater than the one of the two-step analysis, suggesting an improvement in the estimation of the time constant. This is confirmed by the parameters distributions, reported in Figure 5.2, where the distribution of $\hat{\tau}$, identified with the single-step analysis, is comparable to the one obtained with the two steps analysis but shifted to the right. Consequently, there are no values with τ close to zero and the problem of its identification, described in Sec. 4.2.1, is absent with the single-step identification. Regarding the precision of the estimates, we can see that the two-step identification yields better results, even if, also the ones obtained with the single-step are acceptable.

Then, we compare both the *colored residuals* and the *white residuals*

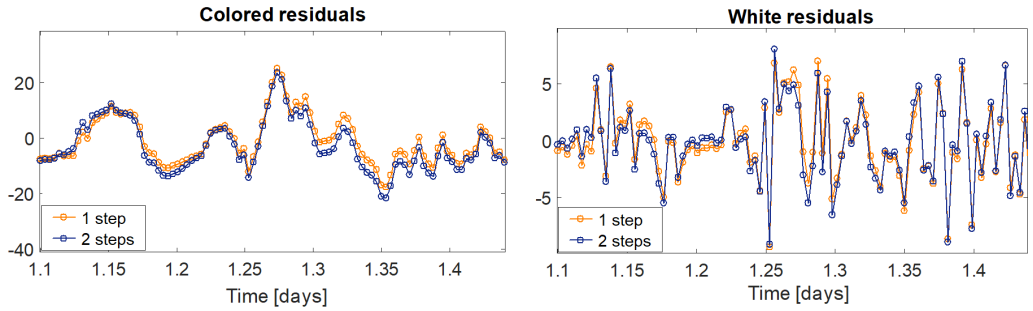

Figure 5.3: Comparisons between the colored residuals (left) and the white residuals (right) in the single step identification (orange signal) and in the two steps identification (blue signal).

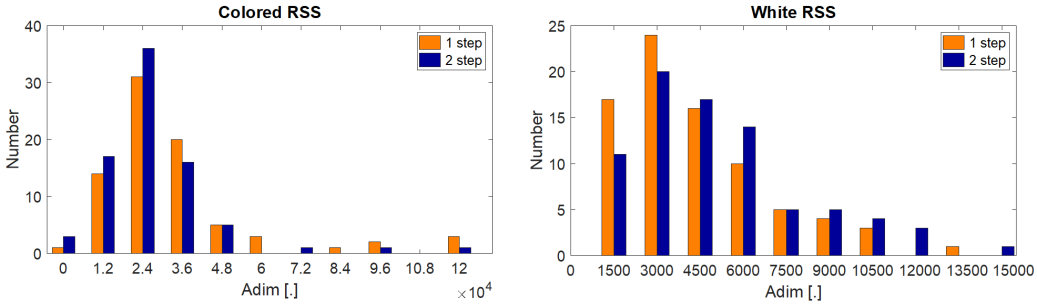

Figure 5.4: Comparisons between the RSS distributions of the colored residuals (left) and the white residuals (right) in the single step identification (orange bars) and in the two steps identification (blue bars).

computed with the two methods. It is important to remark that these methods minimize different objective functions. In the two-step analysis, the *colored residuals* are those to be minimized, while the *white residuals* are obtained afterwards, as described in Sec. 4.3. On the contrary, in the single-step analysis, the *white residuals* are those to be minimized, while the colored ones are saved during the least square computation. Ideally, we would like the residuals, both colored and white, to be the same for the two methods, independently from the objective function. Figure 5.3 shows that the residuals obtained by the two-step analysis and the single-step analysis are not equal but they are very similar, confirming the goodness of the results.

Finally, Figure 5.4 compares the distributions of the sum of the residuals

squares (RSS), colored and white, obtained with the two methods. The results achieved are similar but the single-step method provides white residuals with RSS lower than those of the two-step method. Since the white residuals represent the dynamics that the model is not able to explain, the single-step identification should be preferred.

Chapter 6

Conclusion and future developments

CGM sensors are becoming essential monitoring systems in diabetes therapy, but as all the measurement systems, they are affected by errors that weaken their accuracy and limit in practice the use of CGM-based applications. Developing a model of the error can be extremely useful not only to better understand the sources of error, but also to create more realistic simulations that can help the design and test of CGM-based applications. CGM error models developed in the literature were derived for sensor requiring twice calibrations per day, and thus their domain of validity is limited to 12 hours. These models are not suitable to describe the error of new generation sensors, which are factory calibrated and lost for 10 days. The aim of this thesis was to propose a new model for the error in factory-calibrated devices.

The proposed model of the sensor error is based on the previous works of Facchinetti et al. [1, 2]. Similarly, the error arises from three different sources: the BG-to-IG kinetics, the sensor calibration, and a random noise in the measurement. While in [2] the model parameters change at each calibration, in this work they are identified for the whole sensor lifetime of 10 days. Moreover, new calibration error models are introduced to better investigate the time variability of CGM sensor sensitivity. Specifically, two procedures have been proposed for the identification: a two-step method [2]; and a single-step method, where all the parameters are estimated simultaneously.

The analysis has been conducted on a dataset of 81 subjects wearing

factory calibrated Dexcom G6 sensors.

Regarding the BG-to-IG kinetics, the results obtained with the two methods confirm that the time constant τ is subject specific. Both the methods identify the time constant with great precision ($CV < 30\%$) for all the subjects. However, τ values provided by the single-step method seems to be more reliable in terms of physiological values, because its median value is increased from 3.8 to 4.8 minutes, and no estimates with $\hat{\tau} \simeq 0$, which is it not meaningful from the physiological point of view, result from the analysis. Nevertheless, as remarked in [2], the BG-to-IG kinetics is the part of the model which is most difficult to identify; in fact, we expect the median value of $\hat{\tau}$ to be around the physiological value of $7 - 8$ minutes.

Concerning the calibration error model, the results have shown that the error is time-variant, with the gain $a(t)$ and the offset $b(t)$ polynomials of order 2 and 0, respectively. Therefore, while a constant function is sufficient to describe the offset, the gain is better described by a quadratic function. The parameters are estimated with varying precision, where the two-step method achieves better results in general. Despite a great precision in estimating $\hat{a}_0$ for both the methods, $\hat{a}_1$, $\hat{a}_2$, $\hat{b}_0$ have $CV < 30\%$ for more than 90% of the subjects by using the two-step method, while the precision of the estimates for these parameters falls to 70% with the single-step method. Finally, from the analysis of the spaghetti plots, we were able to distinguish two main behaviours of the sensors gain $a(t)$: quasi-linear and quadratic.

Regarding the measurement noise, the results of both the analysis have shown that an AR model of order 2 is sufficient to effectively describe the error contribution. The variance of the estimated random noise change according to the day of monitoring, confirming an evolution of the sensor behavior in time. In particular, it is greater during the first days and the last ones. This may be explained by the immune system response after the sensor insertion and by the degradation of the performances at the sensor end of life.

Other remarks can be drawn by analyzing the relations among the model parameters. We found out that the parameters of the calibration error model are strongly correlated together by a linear relationship, with absolute values of the coefficients of Pearson around $0.7 - 0.9$. Unfortunately, such a simple relation does not exist between the calibration error parameters and τ.

Overall, the two identification methods provide similar results. However, the single-step method results preferable because it guarantees lower values of the residual sum of squares, i.e. a better overall fit of the data, and it improves the estimation of τ, providing estimates of the time constant which are more physiologically meaningful.

Future developments

A first future development is to include the error model of factory-calibrated CGM sensors in the T1D patient decision Simulator [31], to generate realistic scenarios and re-creating reliable glucose sensors profiles. This would be particularly useful to develop and test CGM-based applications and to evaluate the impact of the error on glycemic control algorithms. The implementation of the model in the Simulator should take into account the provided trends of the sensors gain and linear relations among the parameters of the calibration error model.

The work assumed that at each CGM profile of the dataset corresponds a different subject. While considering a single sensor per subject helped to simplify the analysis, it also hindered the possibility to dissect the error into a common contribution and a sensor specific one. In this sense, the model can be further improved by using a dataset with patients wearing more than one sensor simultaneously.

The obtained results provide acceptable precision of the estimates. Further improvements in the precision are likely to be achieved by developing *ad hoc* iterative estimation techniques. Moreover, a deeper knowledge of the underling physiological phenomena could help in better investigating the error model. For instance, new studies and experiments employing tracers can give new insights on the glucose diffusion from blood to interstitial fluid, potentially leading to a finer model of the BG-to-IG kinetics.

Another future development could be to integrate the sensor error model with the one of the transient errors. Indeed, CGM sensor can be occasionally affected by transient faults like disconnections or artifacts due to compression. Disconnections are caused by the interruption of the communication between the transmitter and the receiver, and they cause the loss of one or

more consecutive samples of CGM data. Compression artifacts are caused by mechanical pressure applied to the sensors by the patients, e.g., while sleeping prone, and they can induce a temporary loss of sensitivity with a consequent distortion of the CGM trace. The model to describe these kinds of transient errors have been already proposed by Facchinetti et. al for sensors requiring multiple calibrations (every 12 hours) in [32]. The integration of such model in factory-calibrated sensors could be useful to assess the possible interplay with the sensor errors.

Finally, additional studies on the errors of factory-calibrated sensors, other than Dexcom G6, would allow to confirm the soundness of the proposed model.

Appendix A

Bayesian smoothing

The Bayesian smoothing is a non-parametric technique that allows to turn a discrete signal y with n sparse values sampled from a non-uniform time grid Ω_s, into a signal $\hat{u}$ with N dense values sampled from a uniform time "virtual" grid Ω_{v}, where $\Omega_s \subset \Omega_{\mathrm{v}}$ and $n \ll N$.

The general measurement model of the signal y is given by

$$y = Gu + v_y, \tag{A.1}$$

where G is the $n \times N$ transfer matrix, and v_y is an additive measurement noise with covariance matrix $\Sigma_v = {\sigma_v}^2 B$. Specifically, the matrix G is obtained by eliminating the rows of G_{v} that do not correspond to the time samples of Ω_s, where G_{v} is the $N \times N$ matrix that relates y_{v} to u in the domain Ω_{v}. Without considering the measurement noise, we want $u = y_{\mathrm{v}}$, i.e., $G_{\mathrm{v}} = I_N$, where I_N is the $N \times N$ identity matrix.

In order to obtain the smoothed signal $\hat{u}$, we observe that

- the y data are affected by the measurement noise; thus, the smoothing procedure should only *approximate* the data, without interpolate them exactly;

- u is a biological signal; thus, it must have some *regularities*.

The trade-off between these two aspects is modeled by the problem

$$\begin{cases} \hat{u} = \arg\min_u \|Fu\|^2 \\ \text{subject to :} \\ \|y - G\hat{u}\|^2 = \varepsilon \end{cases},$$

where F is a difference operator and ϵ is the quadratic norm of the residuals $y - G\hat{u}$. In particular, the quality of the approximation is given by ϵ, while the regularity is measured by $\|Fu\|^2$, which is defined as the discrete form of the energy of the m-th derivative of u. A closed form solution of the problem is given by exploiting the method of Lagrange multipliers and it is computed as

$$\hat{u} = (G^T B^{-1} G + \varphi^o F^T F)^{-1} G^T B^{-1} y \tag{A.2}$$

where φ^o is the optimal value of the smoothing parameter φ, the index that determine the trade-off between approximation and regularity of $\hat{u}$.

Bibliography

[1] A. Facchinetti, S. Del Favero, G. Sparacino, J. R. Castle, W. K. Ward, and C. Cobelli, "Modeling the glucose sensor error," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 3, pp. 620–629, 2013.

[2] A. Facchinetti, S. Del Favero, G. Sparacino, and C. Cobelli, "Model of glucose sensor error components: identification and assessment for new dexcom g4 generation devices," *Medical & biological engineering & computing*, vol. 53, no. 12, pp. 1259–1269, 2015.

[3] W. Definition, "diagnosis and classification of diabetes mellitus and its complications part 1: diagnosis and classification of diabetes mellitus. geneva: Who," *World Health Organization, Geneva, Switzerland*, 1999.

[4] I. D. Federation, "Idf diabetes atlas, 8th edn.," *Brussels, Belgium:International Diabetes Federation,*, vol. 3, pp. 40–65, 2017.

[5] A. D. Association, "Diagnosis and classification of diabetes mellitus," *Diabetes Care*, vol. 33, no. 1, pp. S62–S69, 2010.

[6] A. D. Association, "Classification and diagnosis of diabetes: Standards of medical care in diabetes-2018," *Diabetes Care*, vol. 41, pp. S55–S64, 2018.

[7] A. D. Association, "Classification and diagnosis of diabetes: Standards of medical care in diabetes-2018," *Diabetes Care*, vol. 42, pp. S1–S194, 2018.

[8] P. Rossetti, J. Bondia, J. Vehí, and C. G. Fanelli, "Estimating plasma glucose from interstitial glucose: the issue of calibration algorithms in

commercial continuous glucose monitoring devices," *Sensors*, vol. 10, no. 12, pp. 10936–10952, 2010.

[9] S. Clarke and J. Foster, "A history of blood glucose meters and their role in self-monitoring of diabetes mellitus," *British journal of biomedical science*, vol. 69, no. 2, pp. 83–93, 2012.

[10] J. S. Skyler, "Continuous glucose monitoring: an overview of its development," *Diabetes Technology & Therapeutics*, vol. 11, no. S1, pp. S–5, 2009.

[11] E. M. Benjamin, "Self-monitoring of blood glucose: the basics," *Clinical diabetes*, vol. 20, no. 1, pp. 45–47, 2002.

[12] W. L. Clarke, D. Cox, L. A. Gonder-Frederick, W. Carter, and S. L. Pohl, "Evaluating clinical accuracy of systems for self-monitoring of blood glucose," *Diabetes care*, vol. 10, no. 5, pp. 622–628, 1987.

[13] J. R. Castle and P. G. Jacobs, "Nonadjunctive use of continuous glucose monitoring for diabetes treatment decisions," *Journal of diabetes science and technology*, vol. 10, no. 5, pp. 1169–1173, 2016.

[14] G. Acciaroli, M. Vettoretti, A. Facchinetti, and G. Sparacino, "Calibration of minimally invasive continuous glucose monitoring sensors: state-of-the-art and current perspectives," *Biosensors*, vol. 8, no. 1, p. 24, 2018.

[15] M. Vettoretti, G. Cappon, G. Acciaroli, A. Facchinetti, and G. Sparacino, "Continuous glucose monitoring: current use in diabetes management and possible future applications," *Journal of diabetes science and technology*, vol. 12, no. 5, pp. 1064–1071, 2018.

[16] R. P. Wadwa, L. M. Laffel, V. N. Shah, and S. K. Garg, "Accuracy of a factory-calibrated, real-time continuous glucose monitoring system during 10 days of use in youth and adults with diabetes," *Diabetes technology & therapeutics*, vol. 20, no. 6, pp. 395–402, 2018.

[17] V. N. Shah, L. M. Laffel, R. P. Wadwa, and S. K. Garg, "Performance of a factory-calibrated real-time continuous glucose monitoring system

utilizing an automated sensor applicator," *Diabetes technology & therapeutics*, vol. 20, no. 6, pp. 428–433, 2018.

[18] J. B. Welsh, X. Zhang, S. A. Puhr, T. K. Johnson, T. C. Walker, A. K. Balo, and D. Price, "Performance of a factory-calibrated, real-time continuous glucose monitoring system in pediatric participants with type 1 diabetes," *Journal of diabetes science and technology*, vol. 13, no. 2, pp. 254–258, 2019.

[19] F. Reiterer, P. Polterauer, M. Schoemaker, G. Schmelzeisen-Redecker, G. Freckmann, L. Heinemann, and L. del Re, "Significance and reliability of mard for the accuracy of cgm systems," *Journal of diabetes science and technology*, vol. 11, no. 1, pp. 59–67, 2017.

[20] J. G. Chase, C. E. Hann, M. Jackson, J. Lin, T. Lotz, X.-W. Wong, and G. M. Shaw, "Integral-based filtering of continuous glucose sensor measurements for glycaemic control in critical care," *Computer methods and programs in biomedicine*, vol. 82, no. 3, pp. 238–247, 2006.

[21] C. King, S. M. Anderson, M. Breton, W. L. Clarke, and B. P. Kovatchev, "Modeling of calibration effectiveness and blood-to-interstitial glucose dynamics as potential confounders of the accuracy of continuous glucose sensors during hyperinsulinemic clamp," 2007.

[22] A. J. Laguna, P. Rossetti, F. J. Ampudia-Blasco, J. Vehí, and J. Bondia, "Postprandial performance of dexcom® seven® plus and medtronic® paradigm® veo™: modeling and statistical analysis," *Biomedical Signal Processing and Control*, vol. 10, pp. 322–331, 2014.

[23] C. D. Man, F. Micheletto, D. Lv, M. Breton, B. Kovatchev, and C. Cobelli, "The uva/padova type 1 diabetes simulator: new features," *Journal of diabetes science and technology*, vol. 8, no. 1, pp. 26–34, 2014.

[24] K. Rebrin, G. M. Steil, W. P. Van Antwerp, and J. J. Mastrototaro, "Subcutaneous glucose predicts plasma glucose independent of insulin: implications for continuous monitoring," *American Journal of Physiology-Endocrinology and Metabolism*, vol. 277, no. 3, pp. E561–E571, 1999.

[25] A. Facchinetti, G. Sparacino, F. Zanderigo, and C. Cobelli, "Reconstructing by deconvolution plasma glucose from continuous glucose monitoring sensor data," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 55–58, IEEE, 2006.

[26] M. Breton and B. Kovatchev, "Analysis, modeling, and simulation of the accuracy of continuous glucose sensors," *Journal of diabetes science and technology*, vol. 2, no. 5, pp. 853–862, 2008.

[27] U. Hoss and E. S. Budiman, "Factory-calibrated continuous glucose sensors: the science behind the technology," *Diabetes technology & therapeutics*, vol. 19, no. S2, pp. S–44, 2017.

[28] G. Acciaroli, M. Vettoretti, A. Facchinetti, G. Sparacino, and C. Cobelli, "Reduction of blood glucose measurements to calibrate subcutaneous glucose sensors: a bayesian multiday framework," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 3, pp. 587–595, 2017.

[29] M. J. E. et al., "Advanced analyte sensor calibration and error detection," PCT/US2012/033645.

[30] S. J. V. et al., "Advanced calibration for analyte sensors," PCT/US2014/013146.

[31] M. Vettoretti, A. Facchinetti, G. Sparacino, and C. Cobelli, "Type-1 diabetes patient decision simulator for in silico testing safety and effectiveness of insulin treatments," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 6, pp. 1281–1290, 2017.

[32] A. Facchinetti, S. Del Favero, G. Sparacino, and C. Cobelli, "Modeling transient disconnections and compression artifacts of continuous glucose sensors," *Diabetes technology & therapeutics*, vol. 18, no. 4, pp. 264–272, 2016.