Istituto di Analisi dei Sistemi ed Informatica

Consiglio Nazionale delle Ricerche

U. Picchini, S. Ditlevsen, A. De Gaetano

SYSTEM NOISE MODELIZATION IN GLUCOSE/INSULIN DYNAMICS

R. 630   February 2005

Umberto Picchini – CNR-IASI BioMatLab, Largo A. Gemelli 8 - 00168 Rome, Italy. Email: umberto.picchini@biomatematica.it.

Susanne Ditlevsen – Department of Biostatistics, University of Copenhagen, Denmark. Email: sudi@pubhealth.ku.dk.

Andrea De Gaetano – CNR-IASI BioMatLab, Largo A. Gemelli 8 - 00168 Rome, Italy. Email: andrea.degaetano@biomatematica.it.

Abstract

The Euglycemic Hyperinsulinemic Clamp (EHC) is the most widely used experimental procedure for the determination of insulin sensitivity. In the present study, sixteen subjects with BMI between 18.5 and 63.6 kg/m^2 have been studied with a long-duration (five hours) EHC. In order to explain the oscillations of glycemia occurring in response to the hyperinsulinization and to the continuous glucose infusion at varying speeds, we first hypothesized a system of ordinary differential equations (ODEs), with limited success. We then extended the model and represented the experiment using a system of stochastic differential equations (SDEs). The latter allow for distinction between (i) random variation imputable to observation error and (ii) system noise (intrinsic variability of the metabolic system), due to a variety of influences which change over time. The stochastic model of the EHC was fitted to data and the system noise was estimated by means of a (simulated) maximum likelihood procedure, for a series of different hypothetical measurement error values. We showed that, for the whole range of reasonable measurement error values: (i) the system noise estimates are non-negligible; and (ii) these estimates are robust to changes in the likely value of the measurement error. Explicit expression of system noise is physiologically relevant in this case, since glucose uptake rate is known to be affected by a host of additive influences, usually neglected when modelling metabolism. While in some of the studied subjects system noise appeared to only marginally affect the dynamics, in others the system appeared to be driven more by the erratic oscillations in tissue glucose transport rather than by the overall glucose-insulin control system. It is possible that the quantitative relevance of the unexpressed effects (system noise) should be considered in other physiological situations, represented so far only with deterministic models.

Key words: mathematical models, dynamical systems, euglycemic hyperinsulinemic clamp, stochastic differential equations, parameter estimation, Monte Carlo methods, simulated maximum likelihood

1. Introduction

With the growing epidemiological importance of insulin resistance states (like obesity and Type 2 Diabetes Mellitus, T2DM) and with the increasing clinical recognition of the impact of the so-called metabolic syndrome, the assessment of insulin sensitivity has become a very relevant issue in metabolic research. The experimental procedures currently employed to gather information on the degree of insulin resistance of a subject are the Oral Glucose Tolerance Test (OGTT), the Intra-Venous Glucose Tolerance Test (IVGTT), the Euglycemic Hyperinsulinemic Clamp (EHC), the Hyperglycemic Clamp, the insulin-induced hypoglycemia test (K_{ITT}), and less commonly used methods based on tracer administration ([11], [25], [26]). Of these, the EHC is considered the tool of choice in the diabetological community, in spite of its labor-intensive execution, due to the simple interpretation which is usually attributed to the obtained results ([8], [28]). The favor with which the EHC is viewed in this context stems in part from the belief that while mathematical models of the glucose insulin system make untenable assumptions, the EHC approach is relatively assumption-free, or model independent.

In the present work we study the dynamic behavior of glycemia and insulinemia recorded from human subjects during a EHC procedure. We firstly hypothesize a system of ODEs explaining this dynamics (inspired by a previously published deterministic model of the EHC [21]), and obtain the corresponding parameter estimates by numerically fitting the model to observed data. This simply deterministic model, however, does not accommodate random variations of metabolism. In fact a deterministic model assumes that: (i) the mathematical process generating the observed glycemias is smooth (continuous and continuously differentiable) in the considered time-frame; and (ii) the variability of the actual measurements is due to observation error, which does not influence the course of the underlying process. An alternative, stochastic, approach would result from the hypothesis that the underlying mathematical process itself is not smooth, at least when considered at the practicable time resolution. The glucose metabolizing organs and tissues are in fact subject to a variety of internal and external influences, which change over time (e.g. blood flow, energy requirements, hormone levels, the cellular metabolism of the tissues themselves) and which may affect the instantaneous glycemias. This second approach would maintain that some degree of randomness is already present in the glucose disposition process itself, and that observational error is superimposed to it. A natural extension of the deterministic model is given by a stochastic differential equations (SDEs) model (see e.g. [15] and [16]). We therefore define an SDE model by adding a suitable system variability to the simple deterministic model.

While SDE models are currently being employed in different applied fields (e.g. finance, engineering, physics), they are rarely used in biomedicine (except for specific fields like neuronal modelling, population growth models and recent contributions), even though it is generally recognized that biological data are fraught with many sources of intrinsic (system) error. This large amount of variability is often attributed to observation error exclusively, giving rise to inaccurate, but manageable, modelling representations of physiological phenomena.

The present work has two main goals: on one hand it purports to determine whether, in a particular physiological situation, system error is identifiable and necessary to explain observations, above and beyond commonly accepted levels of measurement error. The second goal is to show, by means of a practically occurring experimental situation, that SDE models are physiologically relevant and that their parameters can be numerically estimated using commonly available resources.

4.

2. Material and methods

2.1. Subjects

Sixteen subjects were enrolled in the study, 8 normal volunteers and 8 patients from the Obesity Outpatient Clinic of the Department of Internal Medicine at the Catholic University School of Medicine. For one normal subject the recorded glycemia values were accidentally lost and this subject was therefore discarded from the following considerations. The subjects had widely differing BMIs (from 18.5 to 63.6 kg/m^2). All subjects were clinically euthyroid, had no evidence of diabetes mellitus, hyperlipidemia, or renal, cardiac or hepatic dysfunction and were undergoing no drug treatments that could have affected carbohydrate or insulin metabolism. The subjects consumed a weight-maintaining diet consisting of at least 250 g. of carbohydrate per day for 1 week before the study. Table 1 reports anthropometric characteristics (BMI, BSA) of the subjects, measured plasma glucose and insulin concentrations (G_{fast}, I_{fast}) immediately before the EHC procedure, and the average levels of insulin after 80 min. of clamp insulinization (I_{max}). The study protocol followed the guidelines of the Medical Ethics Committee of the Catholic University of Rome Medical School; written informed consent was obtained from all subjects.

2.2. Experimental protocol

Each subject was studied in the postabsorptive state after a 12-14 h. overnight fast. Subjects were admitted to the Department of Metabolic Diseases at the Catholic University School of Medicine in Rome the evening before the study. At 07.00 hours on the following morning, the infusion catheter was inserted into an antecubital vein; the sampling catheter was introduced in the contralateral dorsal hand vein and this hand was kept in a heated box (60 $°C$) in order to obtain arterialized blood. A basal blood sample was obtained in which insulin and glucose levels were measured. At 08.00 hours, after a 12-14 h. overnight fast, the EHC was performed according to [8]. A priming dose of short-acting human insulin was given during the initial 10 min. in a logarithmically decreasing manner so that the plasma insulin was raised acutely to the desired level. During the five-hour clamp procedure, the glucose and insulin levels were monitored every 5 min. and every 20 min. respectively, and the rate of infusion of a 20% glucose solution was adjusted during the procedure following the published algorithm [8]. Because serum potassium levels tend to fall during this procedure, KCl was given at a rate of 15-20 mEq/h to maintain the serum potassium between 3.5 and 4.5 mEq/l. Serum glucose was measured by the glucose oxidase method using a Beckman Glucose Analyzer II (Beckman Instruments, Fullerton, Calif., USA). Plasma insulin was measured by microparticle enzyme immunoassay (Abbott Imx, Pasadena, Calif., USA).

2.3. Deterministic model

In order to explain the oscillations of glycemia occurring in response to hyperinsulinization and to continuous glucose infusion at varying speeds, we hypothesized the following system (inspired by a previously published deterministic model of the EHC [21]):

$$\frac{dG(t)}{dt} = \frac{(T_{gx}(t - \tau_g) + T_{gh}(t))}{V_g} - T_{xg}\frac{G(t)}{0.1 + G(t)} - K_{xgI}G(t)I(t), \qquad G(0) = G_b \quad (1)$$

$$\frac{dI(t)}{dt} = \frac{(T_{iG}G(t) + T_{ix}(t))}{V_i} - K_{xi}I(t), \qquad I(0) = I_b \tag{2}$$

$$T_{gh}(t) = T_{ghmax}\exp(-\lambda G(t)I(t)), \qquad T_{gh}(0) = T_{ghb} = T_{ghmax}\exp(-\lambda G_b I_b) \tag{3}$$

where

$$T_{gx}(s) = 0 \quad \forall s \in [-\tau_g, 0] \qquad and \qquad T_{ix}(0) = T_{ixb}$$

$T_{gx}(t)$ and $T_{ix}(t)$ are (input or forcing) state variables of which the values are known at each time; the state variables and the parameters are defined in Table 2 and Table 3. The model is diagrammatically represented in Figure 1.

Equations (1) and (2) express the variations of plasma glucose and plasma insulin concentrations. The variation of glucose concentration in its distribution space is attributed to the external glucose infusion rate, to liver glucose output and to insulin-dependent as well as insulin-independent glucose tissue uptake. Infused glucose raises glycemia after a delay τ_g due to the time required to equilibrate the intravenously infused quantity throughout the distribution space. The insulin-independent glucose tissue uptake process is modelled as a Hill function rapidly increasing to its (asymptotic) maximum value T_{xg}; thus for glycemia values appreciably larger than 0.1 mM the insulin-independent glucose tissue uptake is already close to its maximum. This formulation is intended to represent the aggregated apparent zero-order (fixed) glucose utilization mechanism at rest (mainly by the brain and heart [17] and [22] p. 320), with the mathematical and physiological requirement that glucose uptake tends to zero as glucose concentration in plasma approaches zero. The variation of insulin concentration in its distribution space (equation (2)) may be thought of as due to the external insulin infusion, to glucose-dependent pancreatic insulin secretion and to the apparently first-order insulin removal from plasma. Equation (3) represents the rate of net Hepatic Glucose Output, starting at maximal HGO at zero glucose and zero insulin and decaying monotonically with increases in both glucose and effective insulin concentrations in the plasma. The net HGO is assumed to be equal to T_{ghb} at the beginning of the experiment and to decrease toward zero as glycemia or insulinemia levels increase. Serum insulin affects glucose clearance through equation (1) and the glucose synthesis rate through equation (3).

Steady-state conditions are used to decrease the number of free parameters to be estimated: at steady state, before the start of the clamp ($G = G_b$, $I = I_b$, $T_{gx} = T_{ix} = 0$), we have

$$T_{ghb} = T_{ghmax}\exp(-\lambda G_b I_b)$$

$$0 = \frac{0 + T_{ghb}}{V_g} - \frac{T_{xg}G_b}{0.1 + G_b} - K_{xgI}I_b G_b \Rightarrow T_{xg} = \left(\frac{T_{ghb}}{V_g} - K_{xgI}I_b G_b\right)\frac{(0.1 + G_b)}{G_b}$$

$$0 = \frac{T_{iG}G_b + 0}{V_i} - K_{xi}I_b \Rightarrow T_{iG} = \frac{K_{xi}I_b V_i}{G_b}$$

Therefore the parameters T_{ghb}, T_{xg}, and T_{iG} are completely determined by the values of the other parameters.

2.3.1. Deterministic model estimation

The system (1)–(3) has been numerically integrated by means of a fourth–order Runge–Kutta scheme with constant stepsize equal to 0.5 min. In order to distinguish among the n observations (and corresponding predictions) between glucose and insulin, the indices j for glucose and k for insulin are used as follows: $j \in J$, $k \in K$, $J \cap K = \{\varnothing\}$, $J \cup K = \{1, ..., n\}$. We indicate with $\hat{G}(t, \theta) \equiv \hat{G}(t)$ and $\hat{I}(t, \theta) \equiv \hat{I}(t)$ the (numerically integrated) solutions of equations (1)–(3) for parameter θ at time t. The solutions have been fitted by Iteratively Re-Weighted Least Squares (IRWLS, see e.g. [6, chapter 2]) separately on each subject's glycemia and insulinemia time-points, estimating only the free parameters $\theta = (G_b, I_b, K_{xgI}, K_{xi}, T_{ghmax}, V_g, V_i, \tau_g, \lambda)$ by minimizing the following loss function

$$(y - \hat{y})'\Omega(y - \hat{y})$$

where y is the $n \times 1$ array containing both glycemias and insulinemias, observed at times $0 = t_1 \leq t_2 \leq \cdots \leq t_n$; $\hat{y}$ is the array of corresponding predictions obtained by numerical integration of the system (1)-(3), $\hat{y}(t_j) = \hat{y}_j = \hat{G}(t_j) \; \forall j \in J$, $\hat{y}(t_k) = \hat{y}_k = \hat{I}(t_k) \; \forall k \in K$; Ω is an $n \times n$ diagonal matrix of weights. Here Z' denotes the transpose of the matrix Z. The statistical weight associated with a generic glucose concentration point y_j has been defined as $1/(\hat{y}_j CV_G)^2$, where CV_G is the coefficient of variation for glucose. Similarly the statistical weight associated with a generic insulin concentration point has been defined as $1/(\hat{y}_k CV_I)^2$, where CV_I is the coefficient of variation for insulin. For each subject IRWLS parameter estimates of θ were obtained for several different values of the coefficients of variation CV_G and CV_I, namely:

$$(CV_G, CV_I) \in \quad \{(0.015, 0.07), (0.02, 0.10), (0.03, 0.10), (0.03, 0.15),$$
$$(0.04, 0.15), (0.05, 0.15), (0.15, 0.30)\}.$$

These sets of coefficients of variation were used in order to conduct a sensitivity analysis on the *diffusion coefficient* (defined in section 2.4 below) by considering different values for the variance of the measurement error, which is assumed proportional to the square of the coefficient of variation.

The starting point to fix reasonable values for CV_G and CV_I was suggested in [2] where it had been found $(CV_G, CV_I) = (0.015, 0.07)$; nevertheless, since these values refer to *in vitro* estimates of the variance of repeated laboratory measurements on the same preparation, it could be more realistic to re-estimate CV_G and CV_I from data. To this aim we adopted a General Least Squares approach (GLS [6, chapter 5], detailed in appendix).

2.4. Stochastic model

As an alternative to the above deterministic model, we may assume that the underlying tissue glucose uptake process is not smooth, subject as it is to a variety of metabolic and hormonal influences, which change over time. In fact, tissue glucose uptake is determined not only by the varying concentrations of certain hormones (e.g. cortisol or growth hormone) and by the rhythm of food intake, events which take place over periods of hours, but also by sudden changes in physical activity or emotional stresses induced by thought processes. We may thus imagine that the insulin-dependent glucose disposal rate may be subject to moment-by-moment variations and that the rate constant K_{xgI} is likely to exhibit substantial irregular oscillations over time. Thus we define two sources of noise: a dynamic noise term, which is a part of the process, such

that the value of the process at time t depends on this noise up to time t, and a measurement noise term, which does not affect the process itself, but only its observations.

We therefore allow the parameter K_{xgI} to vary randomly as $(K_{xgI} - \xi(t))$, where $\xi(\cdot)$ is a gaussian white-noise process. Then the system noise $\xi(t)dt$ can be written as $\sigma dW(t)$ (see e.g. [9], [15] and [16]), where $\sigma \geq 0$ represents the (unknown) diffusion coefficient and $W(\cdot)$ is the Wiener process (Brownian motion), which is a random process whose increments are independent and normally distributed with zero mean and with variance equal to the length of the time interval over which the increment take place. By incorporating the K_{xgI} variation into the deterministic model, we obtain the following (Itô) SDE:

$$dG(t) = \left[\frac{(T_{gx}(t - \tau_g) + T_{gh}(t))}{V_g} - T_{xg}\frac{G(t)}{0.1 + G(t)} - K_{xgI}G(t)I(t) \right]dt + \sigma G(t)I(t)dW(t) \quad (4)$$

$$dI(t) = \left[\frac{(T_{iG}G(t) + T_{ix}(t))}{V_i} - K_{xi}I(t) \right]dt, \quad (5)$$

$$T_{gh}(t) = T_{ghmax}\exp(-\lambda G(t)I(t)) \quad (6)$$

with $G(0) = G_b$, $I(0) = I_b$ and $T_{gh}(0) = T_{ghb} = T_{ghmax}\exp(-\lambda G_b I_b)$. Notice that this formulation has the theoretical advantage of never becoming negative in any of the coordinates.

2.5. SDE estimation

While the model estimation procedure for systems of ODEs is well established, estimating parameters in SDE models is not straightforward, except for simple cases. A variety of methods for statistical inference in discretely observed diffusion processes have been developed during the past decades (e.g. [1], [3], [4], [5], [10], [12], [13], [14], [19], [20], [23], [24], [27]). A natural approach would be maximum likelihood inference, but it is rarely possible to write the likelihood function explicitly. In our case it becomes further complicated since we are dealing with partially observed state variables. The estimation approach we follow is to first estimate by IR-WLS the parameters of the ODE system (1)–(3) (as explained in section 2.3), which represents the deterministic part (drift) of the SDE model and the mean of the corresponding stochastic process. We then use the Monte-Carlo approximation to the unknown likelihood function as suggested in [20], in order to estimate σ by keeping fixed the previously obtained drift parameter estimates. We make recourse to the method in [20] since plasma glucose concentrations G and serum insulin concentrations I are not observed at the same time-points (every 5 min. and every 20 min. respectively), and we therefore deal with partially observed state variables, and further because the concentrations are observed with measurement error. In the following the application of the method to the problem under investigation is detailed; for ease of notation we denote with Z_i the generic variable $Z(t_i)$ at time t_i. Consider the model (4)–(6) and observation times $0 = t_1 \leq t_2 \leq \cdots \leq t_n$: all the observations y are collected in a single array and distinguished using the following label-variable

$$\chi_i = \begin{cases} G, & \textit{if the observation at time } t_i \textit{ refers to glucose} \\ I, & \textit{if the observation at time } t_i \textit{ refers to insulin} \end{cases} \quad i = 1, ..., n$$

We consider the error-model

$$Y_i = H_{\chi_i} + \varepsilon_i, \quad (7)$$

where

$$H_{\chi_i} = \begin{cases} G_i, & \textit{if } \chi_i = G \\ I_i, & \textit{if } \chi_i = I \end{cases} \quad i = 1, ..., n$$

and the ε_i's are independent normal variables with mean 0 and variance $\sigma^2_{\chi_i}$ representing the measurement errors. We assume that (i)

$$\sigma^2_{\chi_i} = \begin{cases} (CV_G G_i)^2, & if \ \chi_i = G \\ (CV_I I_i)^2, & if \ \chi_i = I \end{cases} \qquad i = 1, ..., n$$

where CV_G and CV_I represent the coefficient of variations for the glucose and the insulin concentrations respectively, and that (ii) the measurement errors are independent of the process $W(\cdot)$. Equation (7) and the SDE system (4)-(6) provide a representation of the error-structure in our problem. Denote with y_i the observed value of Y_i at time t_i, then the likelihood function of σ can be written as

$$
\begin{aligned}
L(\sigma) &= \int_{\mathbb{R}^{n-2}} \left[\prod_{i=1}^{n} g_i(y_i | H_{\chi_i}; \sigma) \right] f(H_{\chi_3}, ..., H_{\chi_n}; H_{\chi_1}, H_{\chi_2}, \sigma) dH_{\chi_3} \cdots dH_{\chi_n} \\
&= \mathbb{E}_\sigma \prod_{i=1}^{n} g_i(y_i | H_{\chi_i}; \sigma)
\end{aligned}
$$

where $H_{\chi_1} = G_b$ and $H_{\chi_2} = I_b$ are the initial conditions of G and I respectively, f denotes the (unknown) joint density function of $H_{\chi_3}, ..., H_{\chi_n}$ given $(H_{\chi_1}, H_{\chi_2}, \sigma)$, $\mathbb{E}_\sigma$ denotes expectation w.r.t. the distribution of $H_{\chi_3}, ..., H_{\chi_n}$ for the indicated value of σ and

$$g_i(y_i | H_{\chi_i}; \sigma) = \left(2\pi\sigma^2_{\chi_i} \right)^{-1/2} \exp\left[-\frac{1}{2\sigma^2_{\chi_i}} \left(y_i - H_{\chi_i} \right)^2 \right]$$

is the normal density function with expectation H_{χ_i} and variance $\sigma^2_{\chi_i}$. If H^r $(r = 1, ..., R)$ are stochastically independent random vectors, each distributed as $(H_{\chi_3}, ..., H_{\chi_n})$, then it follows from the strong law of large numbers that the likelihood function can, for large values of R, be approximated by

$$L(\sigma) \simeq \frac{1}{R} \sum_{r=1}^{R} \prod_{i=1}^{n} g_i(y_i | H^r_{\chi_i}; \sigma) \tag{8}$$

In practice the approximation is obtained by simulating the $H^r_{\chi_i}$'s (see [15]) for a large finite number R.

We have initially simulated $R = 1000, 2000$ and 4000 trajectories of the process according to the Euler-Maruyama scheme [15] with an integration step-size of 0.1 min.: given a σ of the order of 10^{-5}, the step-size ensures a standard deviation for dG smaller than 0.02 in each integration step, which is very small compared to the order of magnitude of the glucose concentrations. The simulated likelihood functions did not appreciably change the location of their maximum when increasing the number of trajectories beyond 2000. Therefore the reported estimates of σ were obtained by maximizing the approximated likelihood (8), based on $R = 2000$ trajectories, when keeping fixed the parameters entering the drift part of the model and using different combinations of levels of CV_G and CV_I (see section 2.3), in order to explore the sensitivity of the obtained estimates $\hat{\sigma}$ to mis-specification of the observation error.

3. Results

3.1. Deterministic differential model

We first estimated by IRWLS, separately on each subject, the structural (free) parameters G_b, I_b, K_{xgI}, K_{xi}, T_{ghmax}, V_g, V_i, τ_g, λ, and computed the corresponding structural (determined)

parameters T_{ghb}, T_{xg} and T_{iG} entering the deterministic model (1)–(3): the estimates corresponding to three different choices for the coefficients of variation (see section 2.3) are reported in tables 4, 5 and 6, whereas graphical results of the IRWLS fitting are shown in figures 2 and 3 only for the case $(CV_G, CV_I) = (0.05, 0.15)$ (in fact these values are very similar to the GLS population-estimates of CV_G and CV_I, as explained below), since in our problem the predicted curves do not vary substantially with different measurement error values.

Secondarily, we estimated the individual structural parameters, as well as the population parameters CV_G and CV_I, simultaneously by GLS (Table 7): in this way we found that $\widehat{CV}_G = 0.071$ and $\widehat{CV}_I = 0.1702$.

3.2. Stochastic differential model

The stochastic model (4)–(6) was adapted to our data and σ was estimated as described in section 2.5. The estimates of σ corresponding to the different sets of coefficients of variation are reported in Table 8 for each subject. In this table we notice that the σ estimates are stable when considered in a reasonable region of the coefficient of variations values, that is when considered in $(CV_G, CV_I) \in [0.02, 0.05] \times [0.10, 0.15]$. At the smallest level $(CV_G, CV_I) = (0.015, 0.07)$ the σ estimates result numerically unidentifiable for three subjects, and are thus marked with an 'NA'.

While theoretically we could have tried to estimate simultaneously all parameters appearing in the SDE model (drift parameters, diffusion coefficient determining the system noise variance, observation error variance), from a purely computational point of view this proved exceedingly expensive and we had to be content with a sequential estimation approach. In this way, several combinations of observation error levels were hypothesized, the drift parameters were estimated under each hypothesis, and the corresponding diffusion was then estimated in each case.

For illustration purposes, graphical results of the fitting only for the cases $(CV_G, CV_I) = (0.05, 0.15)$ and $(CV_G, CV_I) = (0.03, 0.15)$ are shown in figures 4 and 5 respectively, only for the glycemia values since the insulin curves are almost identical to those produced by the deterministic model. For each subject figures 4 and 5 report the observed glycemias and the empirical mean of $R = 2000$ simulated trajectories of the $G(t)$ process, their empirical 95% confidence limits (from the 2.5th percentile to the 97.5th percentile) and one simulated trajectory. According to [18]-[20] we are able to check the plausibility of our stochastic model by simulating uniform residuals; the q-q plots of the simulated uniform residuals are reported in figures 6 and 7, where the residuals are plotted against percentiles from the $U(0, 1)$ distribution. The caption of each subfigure also reports the p-value from the two-tailed Kolmogorov-Smirnov goodness-of-fit test; if $p < 0.05$ the simulated residuals do not conform to the hypothesis of $U(0, 1)$ distribution at a 5% confidence level. The tests have not been subjected to correction for simultaneous inference (Bonferroni or similar) in order to be more conservative. All tests had $p > 0.05$, except for the glycemia residuals for subjects 4 and 6 and for the insulinemia residuals for subjects 8 and 15.

4. Discussion

The Euglycemic Hyperinsulinemic Clamp is the procedure most commonly employed by research diabetologists in their quest for the determination of the degree of insulin sensitivity (or resistance) exhibited by a given experimental subject. In its common usage, after having "clamped", i.e. stabilized, the subject's glycemia to pre-insulinization levels, the average rate of

glucose infusion necessary to maintain euglycemia is measured, and is directly employed (once normalized by the subject's body mass) as an index of insulin sensitivity.

As an alternative to the above, a model may be drawn to describe the mass flow of glucose into and out the central (sampling) plasma compartment of the subject, explicitly representing the physiological mechanisms known to intervene in the process.

In the present work, a (simple) deterministic model of the clamp procedure is studied first. The main result of this study is that the level of error around the predicted curve is very large, in particular it is much larger than the (0.015,0.07) commonly accepted levels of measurement error in *in vitro* repeated testing of the same laboratory preparation. This result would theoretically be compatible with either one of the following alternatives: the model is mis-specified; the *in vivo* measurement error is in reality much larger than (0.015,0.07); or, there is some additional source of noise, besides measurement error, which substantially impacts observations.

From an examination of Figure 2, it would not seem that the average model prediction is systematically wrong. Similarly, the coefficient of variations estimated by GLS around the deterministic prediction are much too large to be compatible with measurement error. The idea that glucose absorption by tissues varies in time is, on the other hand, rather natural: it seems evident that, subject to variable hormonal concentrations, variable stress levels, even to minor posturali changes, muscle uptake and liver output of glucose may vary from moment to moment. What remains to be seen is if a mathematical model incorporating this idea would be supported by the actual observations.

A random fluctuation in the net tissue glucose uptake rate is a reasonable approximation to the effect of a host of the poorly controlled, additive influences mentioned. When considering this random fluctuations as well, the original deterministic model (1)–(3) is thus transformed into the SDE model (4)–(6).

The approach followed for the estimation of the relevant quantities of the model (structural parameters and diffusion) is motivated by the computing–intensive algorithms necessary for the estimation of the diffusion, which require the simulation of thousands of possible trajectories of the process for every evaluation of the merit function.

Essentially, it has been shown that: (i) for any reasonable level of observation error, the estimated diffusion has more or less the same value. For "reasonable" it is here meant larger than pure measurement error and smaller than the total error around the expected trajectory as estimated by GLS. Adopting the lowermost observation error level (0.015,0.07) would be equivalent to stating that the same variability exists on repeated laboratory measurements on the same sample as on repeated sampling/measurement procedures at the same actual glycemia and insulinemia, disregarding further potential sources of variation accruing to the sampling procedure itself (volume of blood vs. volume of anticoagulant, degree of coagulation, variation in spinning time etc.): this seems extreme in one direction. In the other direction, the largest error level considered (0.15,0.30) is much higher than the total observation variability actually estimated with the GLS procedure around the deterministic prediction (0.071,0.1702), and for this reason should also be discarded. Having excluded these extreme cases, it can be seen that, in the present situation, the estimation of the diffusion is very robust to changes in the likely value of the observation error, as similarly robust are the estimates of the structural (drift) parameters. Further (ii) we showed that the diffusion coefficient estimates are generally strictly positive: this means that the dynamical process which most likely represents the glycemia time-course (given the estimated deterministic differential model) is a stochastic process with a non-negligible sys-

tem noise, whose intensity factor is represented by the diffusion coefficient. Pictorial evidence of the magnitude of diffusion is given in figures 4 and 5. This system noise represents the additive action of many factors, each with a small individual effect, which are not explicitly represented in the deterministic model (that is in the drift term of the SDE), and which instantaneously affect glucose uptake rate. Therefore, in the stochastic differential model the collective influence of many individually neglected effects is added to the average drift term, which, on the other side, represents the most relevant and generally well-recognized factors affecting glycemia.

It is interesting to note from figures 4 and 5 that when the process average (which also represents the ordinary differential model solution) fits well the observed glycemias, its 95% confidence band is narrow (e.g. subjects 1, 2, 3, 10 and 12). On the other hand, when the process average itself is not able to meaningfully capture the general trend of the observations, the corresponding confidence band is much larger (e.g. subjects 4, 5 and 9). In these last cases, the system is driven more by the erratic oscillations in tissue glucose transport rather than by the smooth dynamics of the overall actual system, of course under the hypothesis that the proposed deterministic term is correctly specified. This finding would prompt us to re-consider the quantitative relevance of the unexpressed effects, in other situations, represented so far only with deterministic models, especially when less than perfectly satisfactory fits to data have been obtained.

The particular behavior of the estimated diffusion (Table 8), for the different choices of coefficient of variation values $\xi \equiv (CV_G, CV_I)$ could seem counterintuitive: one would expect that as the observation error is assumed to increase, the estimated system error should decrease (this actually happens for macroscopically exaggerated values of the observation error). This unexpected result may be due to the estimation method we adopted: since the array of free structural parameters θ for the drift part of the stochastic model is estimated anew for different values of ξ, the $\hat{\theta}$'s are depending on these levels of error, so we can write $\hat{\theta} \equiv \hat{\theta}_\xi$. As a consequence of that, the (numerical) solution of the SDE system (4)–(6) and the measurement error ε are not independent and, by means of equation (7), we can write

$$
\begin{aligned}
Var(Y_i) &= Var(H_{\chi_i}(\hat{\theta}_\xi)) + Var(\varepsilon_i) + 2Cov(H_{\chi_i}(\hat{\theta}_\xi), \varepsilon_i) \\
&= Var(H_{\chi_i}(\hat{\theta}_\xi)) + \sigma^2_{\chi_i} + 2\mathbb{E}(H_{\chi_i}(\hat{\theta}_\xi)\varepsilon_i).
\end{aligned}
$$

Since $H(\cdot)$ is unknown, so is $\mathbb{E}(H_{\chi_i}(\hat{\theta}_\xi)\varepsilon_i)$ and we cannot compute the covariance analytically: however it is generally not zero, and the variance of the observations is not the simple sum of the variance of the trajectories and of the variance of the measurement error.

It is of interest to note that even when glycemia is allowed to vary stochastically, responding to the Wiener process introduced in equation (4), the corresponding oscillations in insulinemia are very small, and the insulin process does not substantially differ from its own expected value. This may very well be explained when considering the relative inertia of the pancreatic insulin secretion mechanism, coupled with the large volume of distribution of the hormone, which combine to minimize oscillations in insulin concentrations in response to rapidly varying glucose concentrations.

We conclude therefore that the stochastic differential model (4)–(6) is statistically robust, physiologically meaningful and represents well the glucose metabolism occurring during a clamp study. More generally, it can be concluded that stochastic differential equations are theoretically

12.

useful and practically applicable, and deserve to be considered more often as a valuable addition to the biomedical modeller's toolbox.

Appendix

To obtain subject-specific regression parameters and population estimates of CV_G and CV_I the GLS method was performed [6, chapter 5]. The GLS is a *two-stage* method: (stage 1) at first individual estimates for each subject i ($i = 1, ..., 15$) were obtained; then (stage 2) these estimates were used as building blocks to construct the population estimates of CV_G and CV_I.

Suppose that y_i and θ_i represent the n_i-dimensional array of recorded data and the array of (structural) individual parameters for subject i respectively ($i = 1, ..., 15$), i.e. θ_i contains the values of the (free) parameters $\theta = (G_b, I_b, K_{xgI}, K_{xi}, T_{ghmax}, V_g, V_i, \tau_g, \lambda)$ entering the model (1)–(3) for subject i. Consider now the model

$$y_i = f_i(\theta_i) + \varepsilon_i$$

such that

$$\mathbb{E}(\varepsilon_i|\theta_i) = 0, \quad Cov(\varepsilon_i|\theta_i) = \Omega_i(\theta_i, \xi)$$

with $f_i(\cdot)$ representing the numerical solution of the system (1)–(3) for subject i, and assuming that the functional form of $\Omega_i(\cdot, \cdot)$ and the intra-individual covariance parameter $\xi = (CV_G, CV_I)$ are the same across individuals. If we denote with G and I the state variable Glucose and Insulin respectively, the covariance matrix $\Omega_i(\theta_i, \xi)$ in the present application has the structure of an $n_i \times n_i$ block-diagonal matrix

$$\Omega_i(\theta_i, \xi) = \begin{pmatrix} \Omega_{i,G} & 0 \\ 0 & \Omega_{i,I} \end{pmatrix} \qquad i = 1, ..., 15$$

where

$$\Omega_{i,G} = \begin{pmatrix} CV_G^2 f_{iG}^2(\theta_i, t_{i,1}) & 0 & \cdots & 0 \\ & \cdots & \cdots & \cdots & \cdots \\ 0 & & \cdots & 0 & CV_G^2 f_{iG}^2(\theta_i, t_{i,n_{iG}}) \end{pmatrix}$$

$$\Omega_{i,I} = \begin{pmatrix} CV_I^2 f_{iI}^2(\theta_i, t_{i,1}) & 0 & \cdots & 0 \\ & \cdots & \cdots & \cdots & \cdots \\ 0 & & \cdots & 0 & CV_I^2 f_{iI}^2(\theta_i, t_{i,n_{iI}}) \end{pmatrix}$$

with $f_{iG}(\theta_i, t_{i,j_1})$ and $f_{iI}(\theta_i, t_{i,j_2})$ representing the predicted glycemia and insulinemia values at times $t_{i,j1}$ and $t_{i,j2}$ respectively ($j_1 = 1, ..., n_{iG}$; $j_2 = 1, ..., n_{iI}$; $n_{iG} + n_{iI} = n_i$). Then the GLS algorithm is given by the following scheme:

1. in $m = 15$ separate regressions, obtain preliminary estimates $\hat{\theta}_i^{(p)}$ for each individual, $i = 1, ..., m$;

2. use residuals from these preliminary fits to estimate ξ by minimizing the following functional

$$\sum_{i=1}^m PL_i(\hat{\theta}_i^{(p)}, \xi) = \sum_{i=1}^m \log |\Omega_i(\hat{\theta}_i^{(p)}, \xi)| + (y_i - f_i(\hat{\theta}_i^{(p)}))' \Omega_i^{-1}(\hat{\theta}_i^{(p)}, \xi)(y_i - f_i(\hat{\theta}_i^{(p)}))$$

where PL_i is the pseudolikelihood of ξ for the ith individual. Form estimated weight matrices based on the estimate $\hat{\xi}$ obtained from this procedure, along with the preliminary $\hat{\theta}_i^{(p)}$, to form

$$\hat{\Omega}_i(\hat{\theta}_i^{(p)}, \hat{\xi})$$

3. using the estimated weight matrices from step 2, re-estimate the θ_i's by m separate minimizations: for individual i, minimize in θ_i

$$(y_i - f_i(\theta_i))'\hat{\Omega}_i^{-1}(y_i - f_i(\theta_i))$$

Treating the resulting estimators as new preliminary estimators, return to step 2.

The algorithm should be iterated at least once to eliminate the effect of potentially inefficient preliminary estimates in step 1.

Subject	BMI $[kg/m^2]$	BSA $[m^2]$	G_{fast} $[mM]$	I_{fast} $[pM]$	I_{max} $[pM]$
1	20.20	1.60	3.61	36.84	472.00
2*	35.93	2.38	4.83	108.42	607.30
3*	27.77	2.04	5.39	79.23	683.18
4*	38.10	2.11	5.11	139.00	625.76
5	20.03	1.49	4.83	15.29	506.89
6	19.33	1.55	3.67	21.55	464.54
7*	48.07	2.16	5.39	139.00	592.02
8	18.51	1.55	4.06	13.90	527.00
9*	63.57	2.08	5.94	152.90	522.25
10	18.59	1.46	5.44	49.34	482.42
11*	42.19	2.02	5.28	139.00	497.31
12	22.59	1.71	3.39	32.66	469.22
13*	31.35	2.11	4.94	79.23	605.22
14*	27.91	1.83	4.61	83.40	679.62
15	22.68	1.73	3.50	27.80	482.14

Table 1: Anthropometric and metabolic characteristics for the normal and obese (*) subjects; BSA is the Body Surface Area $[m^2]$ calculated via the DuBois formula ($BSA = 0.20247 \cdot height^{0.725}[m] \cdot weight^{0.425}[kg]$).

Variables	
t $[min]$	time from insulin infusion start
$G(t)$ $[mM]$	plasma glucose concentration at time t
$I(t)$ $[pM]$	serum insulin concentration at time t
$T_{gx}(t)$ $[mmol/min/kgBW]$	glucose infusion rate at time t
$T_{ix}(t)$ $[pmol/min/kgBW]$	insulin infusion rate at time t
$T_{gh}(t)$ $[mmol/min/kgBW]$	net Hepatic Glucose Output (HGO) at time t

Table 2: Definitions of the state variables.

Parameters	
$G_b \ [mM]$	basal glycemia
$I_b \ [pM]$	basal insulinemia
$T_{xg} \ [mM/min]$	maximal insulin-independent rate constant for glucose tissue uptake
$K_{xgI} \ [min^{-1}/pM]$	insulin-dependent apparent first-order rate constant for glucose tissue uptake at insulinemia I
$K_{xi} \ [min^{-1}]$	apparent first-order rate constant for insulin removal from plasma
$T_{iG} \ [pM/min/mM]$	apparent zero-order net insulin synthesis rate at unit glycemia (after liver first-pass effect)
$T_{ixb} \ [pmol/min/kgBW]$	basal insulin infusion rate, which is given by the measured value of T_{ix} at time zero according to [7]
$T_{ghmax} \ [mmol/min/kgBW]$	maximal Hepatic Glucose Output at zero glycemia, zero insulinemia
$T_{ghb} \ [mmol/min/kgBW]$	basal value of T_{gh}
$V_g \ [L/kgBW]$	volume of distribution for glucose
$V_i \ [L/kgBW]$	volume of distribution for insulin
$\tau_g \ [min]$	discrete (distributional) delay of the change in glycemia following glucose infusion
$\lambda \ [mM^{-1}pM^{-1}]$	rate constant for Hepatic Glucose Output decrease with increase of glycemia and insulinemia

Table 3: Definitions of the parameters.

Subject	$\hat{G}_b$	$\hat{I}_b$	$\hat{K}_{xgI}$	$\hat{K}_{xi}$	$\hat{T}_{ghmax}$	$\hat{V}_g$	$\hat{V}_i$	$\hat{\tau}_g$	$\hat{\lambda}$	$\hat{T}_{ghb}$	$\hat{T}_{xg}$	$\hat{T}_{iG}$
1	3.662	42.07	6.05E-5	0.018	0.016	0.464	0.829	0.54	7.91E-3	0.005	1.08E-3	0.17
2	4.535	217.16	4.07E-5	0.010	0.012	0.266	0.990	2.50	9.33E-8	0.012	5.99E-3	0.45
3	6.153	588.73	9.94E-5	0.028	1.000	0.990	0.990	0.13	19.18E-8	0.999	659.84E-3	2.63
4	5.764	141.49	0.20E-8	0.040	0.320	0.990	0.274	0.00	3.28E-3	0.022	22.69E-3	0.27
5	5.043	15.52	2.16E-5	0.026	0.002	0.988	0.578	17.00	0.09E-3	0.002	2.01E-7	0.05
6	3.755	27.31	2.48E-5	0.033	0.044	0.719	0.535	0.50	31.06E-3	0.002	0.02E-7	0.13
7	5.547	141.74	0.75E-5	0.030	0.033	0.988	0.341	0.00	2.20E-3	0.006	0.09E-7	0.26
8	4.227	15.64	1.85E-5	0.019	0.005	0.990	0.714	0.00	20.00E-3	0.001	0.01E-7	0.05
9	6.645	219.74	0.95E-5	0.010	0.012	0.854	0.990	8.00	0.01E-3	0.012	11.80E-7	0.33
10	5.500	257.45	26.56E-5	0.163	0.055	0.141	0.229	5.49	0.03E-3	0.053	2.62E-7	1.75
11	5.589	140.83	0.64E-5	0.029	0.007	0.874	0.480	0.00	0.06E-3	0.004	4.79E-7	0.35
12	3.146	34.14	3.57E-5	0.019	0.004	0.990	0.731	0.00	0.06E-3	0.004	0.40E-7	0.15
13	5.332	84.34	0.63E-5	0.025	0.025	0.990	0.422	0.38	4.91E-3	0.003	1.92E-7	0.17
14	5.136	246.11	1.97E-5	0.010	0.646	0.948	0.990	12.49	1.25E-8	0.646	0.66	0.48
15	3.629	48.01	2.06E-5	0.017	0.010	0.990	0.758	0.50	5.81E-3	0.004	0.11E-7	0.17

Table 4: IRWLS parameter estimates for the ODE model (1)–(3) when $CV_G = 0.015$ and $CV_I = 0.07$; the notation E±p is used for $10^{\pm p}$.

Subject	$\hat{G}_b$	$\hat{I}_b$	$\hat{K}_{xgI}$	$\hat{K}_{xi}$	$\hat{T}_{ghmax}$	$\hat{V}_g$	$\hat{V}_i$	$\hat{\tau}_g$	$\hat{\lambda}$	$\hat{T}_{ghb}$	$\hat{T}_{xg}$	$\hat{T}_{iG}$
1	3.695	39.42	6.07E-5	0.020	0.016	0.472	0.779	0.50	9.06E-3	0.004	1.60E-12	0.16
2	4.433	215.92	4.36E-5	0.010	0.011	0.254	0.990	0.50	2.30E-7	0.011	0.01E-3	0.49
3	7.101	490.98	52.05E-5	0.019	0.976	0.130	0.990	3.08	0.16E-3	0.554	247.88E-3	1.30
4	5.758	141.16	4.66E-9	0.040	0.319	0.990	0.270	0.01	3.29E-3	0.022	22.60E-3	0.27
5	5.112	15.65	2.23E-5	0.029	0.073	0.990	0.532	16.50	4.66E-7	0.073	73.70E-3	0.05
6	3.794	21.52	2.48E-5	0.037	0.024	0.726	0.483	0.71	34.45E-3	0.001	3.15E-12	0.10
7	5.574	140.98	0.75E-5	0.034	0.031	0.988	0.308	0.50	2.11E-3	0.006	3.17E-8	0.26
8	4.318	14.32	1.94E-5	0.022	0.004	0.976	0.646	0.49	20.48E-3	0.001	4.07E-6	0.05
9	6.661	220.06	0.91E-5	0.012	0.013	0.967	0.912	2.01	0.03E-3	0.013	1.10E-7	0.36
10	5.516	254.06	26.49E-5	0.165	0.054	0.138	0.222	5.04	0.04E-3	0.051	1.30E-7	1.69
11	5.600	139.75	0.77E-5	0.032	0.006	0.905	0.432	0.24	0.06E-3	0.006	0.15E-3	0.35
12	3.209	34.12	3.66E-5	0.023	0.004	0.990	0.663	0.02	0.02E-4	0.004	0.04E-3	0.16
13	5.368	79.21	0.63E-5	0.03	0.026	0.990	0.359	0.36	5.42E-3	0.003	1.80E-10	0.16
14	5.648	467.41	6.21E-5	0.021	0.969	0.943	0.985	22.00	0.03E-3	0.902	806.10E-3	1.68
15	3.687	28.97	2.18E-5	0.020	0.010	0.980	0.681	0.49	13.5E-3	0.002	0.20E-3	0.11

Table 5: IRWLS parameter estimates for the ODE model (1)–(3) when $CV_G = 0.05$ and $CV_I = 0.15$; the notation E$\pm p$ is used for $10^{\pm p}$.

18.

Subject	$\hat{G}_b$	$\hat{I}_b$	$\hat{K}_{xgl}$	$\hat{K}_{xi}$	$\hat{T}_{ghmax}$	$\hat{V}_g$	$\hat{V}_i$	$\hat{\tau}_g$	$\hat{\lambda}$	$\hat{T}_{ghb}$	$\hat{T}_{xg}$	$\hat{T}_{iG}$
1	3.807	38.69	6.16E-5	0.022	0.011	0.471	0.730	1.45	6.41E-3	0.004	0.02E-3	0.16
2	4.616	239.06	4.25E-5	0.012	0.014	0.290	0.990	0.35	0.02E-4	0.014	0.70E-3	0.61
3	6.943	507.86	24.88E-5	0.021	0.814	0.908	0.990	0.00	0.01E-3	0.796	0.99E-7	1.52
4	5.752	140.18	9.0E-10	0.041	0.308	0.990	0.266	0.49	3.28E-3	0.022	0.02	0.26
5	5.193	15.33	2.27E-5	0.035	0.002	0.990	0.449	14.52	0.02E-3	0.002	0.14E-3	0.05
6	3.841	21.55	2.44E-5	0.041	0.034	0.738	0.443	0.91	37.88E-3	0.001	0.1E-7	0.10
7	5.600	139.77	0.75E-5	0.038	0.026	0.990	0.276	0.50	1.91E-3	0.006	0.05E-7	0.26
8	4.448	13.96	1.96E-5	0.027	0.005	0.990	0.558	0.37	21.27E-3	0.001	0.1E-9	0.05
9	6.825	154.79	0.70E-5	0.024	0.008	0.990	0.433	0.50	0.06E-3	0.007	0.24E-5	0.24
10	5.550	250.52	26.57E-5	0.166	0.053	0.137	0.216	5.50	0.03E-3	0.051	1.36E-3	1.62
11	5.612	139.41	0.76E-5	0.035	0.006	0.923	0.404	0.02	0.04E-3	0.006	0.1E-3	0.35
12	3.335	32.91	3.67E-5	0.029	0.676	0.989	0.539	1.05	0.01E-4	0.676	0.700	0.16
13	5.414	80.64	0.63E-5	0.035	0.024	0.990	0.309	0.24	5.00E-3	0.003	0.029E-9	0.16
14	5.546	447.67	5.68E-5	0.020	0.996	0.990	0.990	12.00	0.01E-3	0.962	0.85	1.63
15	3.890	27.66	2.27E-5	0.028	0.007	0.990	0.532	0.50	9.67E-3	0.002	0.16E-6	0.11

Table 6: IRWLS parameter estimates for the ODE model (1)–(3) when $CV_G = 0.15$ and $CV_I = 0.30$; the notation E$\pm p$ is used for $10^{\pm p}$.

	Individual estimates											
Subject	$\hat{G}_b$	$\hat{I}_b$	$\hat{K}_{xgI}$	$\hat{K}_{xi}$	$\hat{T}_{ghmax}$	$\hat{V}_g$	$\hat{V}_i$	$\hat{\tau}_g$	$\hat{\lambda}$	$\hat{T}_{ghb}$	$\hat{T}_{xg}$	$\hat{T}_{iG}$
1	3.662	42.07	6.40E-5	0.019	0.016	0.464	0.829	0.54	8.19E-3	0.005	0.57E-7	0.18
2	4.737	238.23	4.16E-5	0.011	0.340	0.297	0.990	2.50	3.68E-8	0.340	1.12	0.57
3	6.183	588.73	27.65E-5	0.039	1.000	0.990	0.849	0.13	1.02E-6	0.996	5.53E-7	3.16
4	5.764	141.49	0.52E-10	0.040	0.321	0.990	0.274	0.00	3.28E-3	0.022	0.02	0.27
5	5.118	15.49	2.28E-5	0.033	1.000	0.990	0.471	17.00	7.61E-8	1.000	1.03	0.05
6	3.754	27.31	2.51E-5	0.035	0.045	0.720	0.518	0.50	31.03E-3	0.002	0.11E-7	0.13
7	5.547	141.74	0.76E-5	0.031	0.033	0.988	0.338	0.00	2.19E-3	0.006	0.95E-7	0.27
8	4.227	15.64	1.95E-5	0.020	0.005	0.990	0.713	0.00	19.40E-3	0.001	0.12E-7	0.05
9	6.759	219.60	0.74E-5	0.016	0.013	0.990	0.706	8.00	0.13E-3	0.011	12.83E-7	0.38
10	5.498	257.37	21.07E-5	0.171	0.059	0.140	0.215	5.49	0.25E-3	0.042	1.58E-7	1.73
11	5.600	140.83	0.73E-5	0.033	0.006	0.899	0.421	0.00	0.14E-3	0.005	0.11E-7	0.35
12	3.247	33.93	3.68E-5	0.028	0.860	0.990	0.569	0.00	2.61E-8	0.860	0.89	0.17
13	5.332	84.34	0.63E-5	0.025	0.025	0.990	0.422	0.38	4.90E-3	0.003	0.79E-7	0.17
14	5.168	247.46	2.13E-5	0.013	1.000	0.990	0.921	12.49	5.45E-9	1.000	1.00	0.57
15	3.629	48.01	2.18E-5	0.018	0.010	0.982	0.758	0.50	5.69E-3	0.004	0.08E-7	0.18
Population estimates												
$\widehat{CV}_G = 0.071,\ \widehat{CV}_I = 0.1702$												

Table 7: GLS individual and population parameter estimates for the ODE model (1)–(3); the notation E$\pm p$ is used for $10^{\pm p}$.

Subjects	$\hat{\sigma}^{(1)}$	$\hat{\sigma}^{(2)}$	$\hat{\sigma}^{(3)}$	$\hat{\sigma}^{(4)}$	$\hat{\sigma}^{(5)}$	$\hat{\sigma}^{(6)}$	$\hat{\sigma}^{(7)}$
1	1.60E-5	1.78E-5	2.25E-5	1.59E-5	2.10E-5	2.25E-5	0
2	NA	1.38E-5	1.47E-5	1.38E-5	1.38E-5	1.15E-5	2.88E-7
3	2.39E-5	4.55E-5	5.71E-5	2.54E-5	3.95E-5	2.58E-5	0
4	NA	1.00E-5	1.00E-5	1.00E-5	1.00E-5	0.95E-5	3.68E-8
5	1.83E-5	1.97E-5	2.00E-5	1.93E-5	1.93E-5	1.93E-5	0.91E-5
6	2.72E-5	2.65E-5	2.68E-5	2.71E-5	2.72E-5	2.73E-5	5.29E-8
7	0.80E-5	0.80E-5	0.80E-5	0.80E-5	2.35E-8	2.35E-8	1.47E-7
8	0.72E-5	0.76E-5	0.73E-5	0.72E-5	0.60E-5	0.42E-5	2.12E-7
9	NA	2.42E-5	2.65E-5	2.50E-5	2.60E-5	2.69E-5	4.77E-7
10	3.08E-5	3.04E-5	3.04E-5	3.04E-5	3.00E-5	2.88E-5	4.25E-7
11	0.62E-5	0.59E-5	3.68E-8	3.68E-8	2.35E-8	3.68E-8	4.77E-7
12	1.44E-5	0.98E-5	1.36E-5	1.41E-5	1.47E-5	1.53E-5	8.47E-7
13	1.23E-5	0.82E-5	0.86E-5	0.74E-5	0.84E-5	0.87E-5	2.12E-7
14	1.73E-5	1.65E-5	1.64E-5	1.62E-5	1.56E-5	0.12E-5	7.21E-8
15	1.87E-5	1.23E-5	1.44E-5	1.86E-5	1.47E-5	1.50E-5	0.84E-5

Table 8: Estimates of σ in the cases $(CV_G, CV_I) = (0.015, 0.07)$, $(CV_G, CV_I) = (0.02, 0.10)$, $(CV_G, CV_I) = (0.03, 0.10)$, $(CV_G, CV_I) = (0.03, 0.15)$, $(CV_G, CV_I) = (0.04, 0.15)$, $(CV_G, CV_I) = (0.05, 0.15)$ and $(CV_G, CV_I) = (0.15, 0.30)$ given by $\hat{\sigma}^{(1)}$, $\hat{\sigma}^{(2)}$, $\hat{\sigma}^{(3)}$, $\hat{\sigma}^{(4)}$, $\hat{\sigma}^{(5)}$, $\hat{\sigma}^{(6)}$ and $\hat{\sigma}^{(7)}$ respectively. The notation E$\pm p$ is used for $10^{\pm p}$.
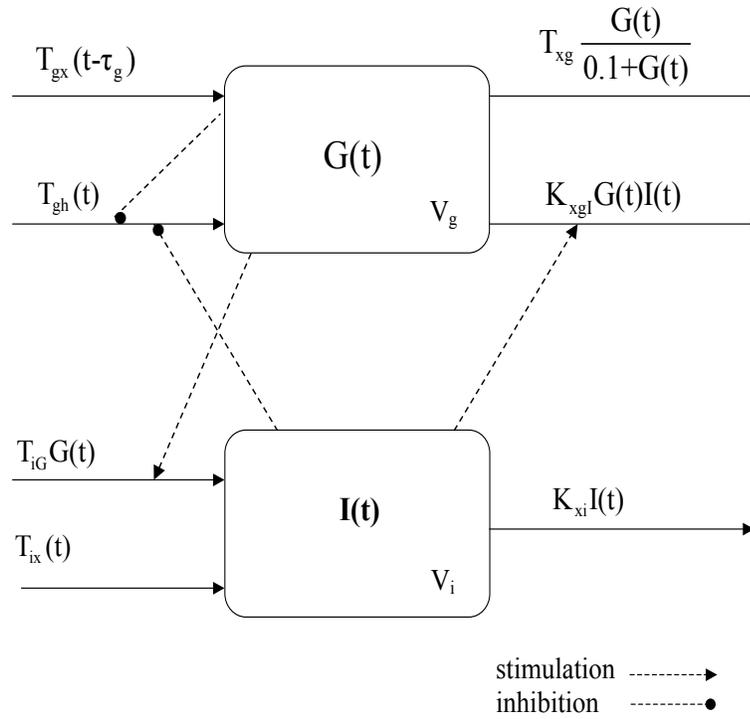
Figure 1: Schematic representation of the model (1)–(3).

(a) Subject 1

(b) Subject 2

(c) Subject 3

(d) Subject 4

(e) Subject 5

(f) Subject 6

(g) Subject 7

(h) Subject 8

(i) Subject 9

(j) Subject 10

(k) Subject 11

(l) Subject 12

(m) Subject 13

(n) Subject 14
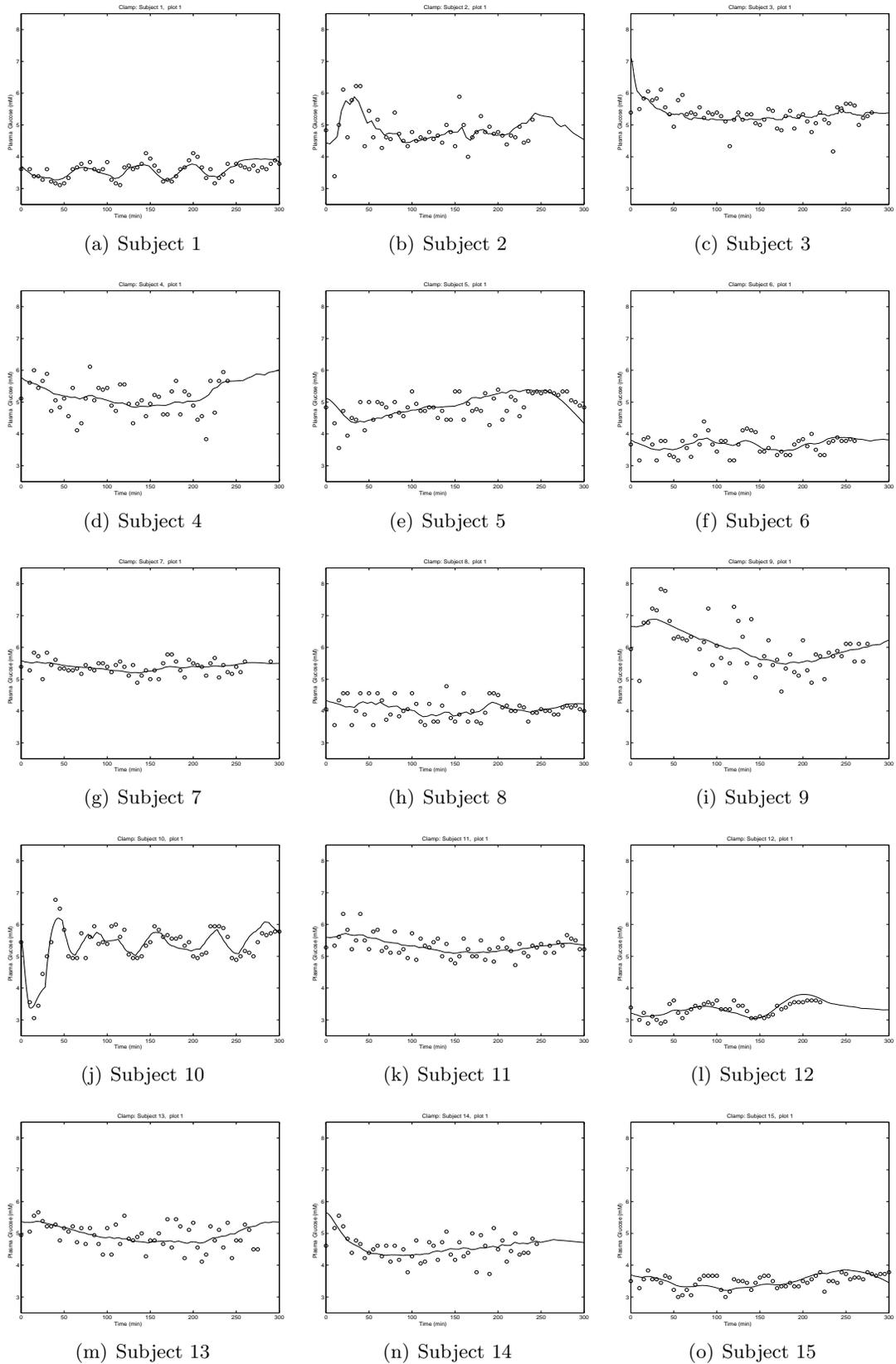
(o) Subject 15

Figure 2: ODE model: observed (○) and predicted (solid line) glycemia corresponding to the IRWLS estimates for the case $(CV_G, CV_I) = (0.05, 0.15)$ (see Table 5).
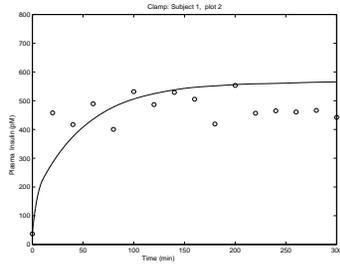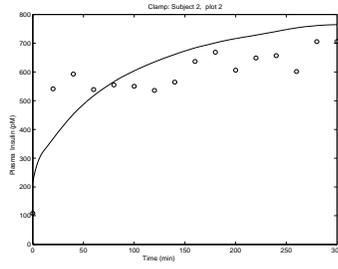
(a) Subject 1
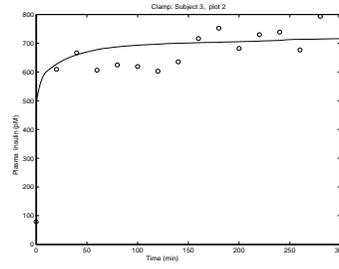
(b) Subject 2
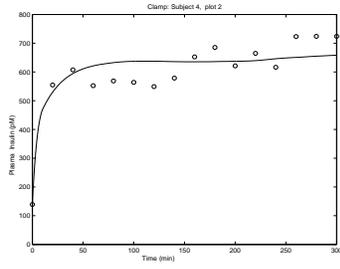
(c) Subject 3
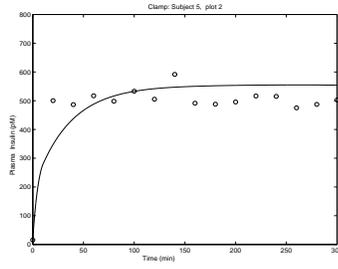
(d) Subject 4
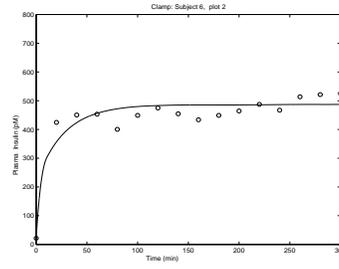
(e) Subject 5
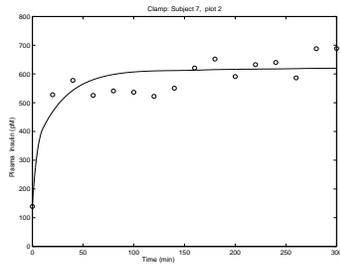
(f) Subject 6
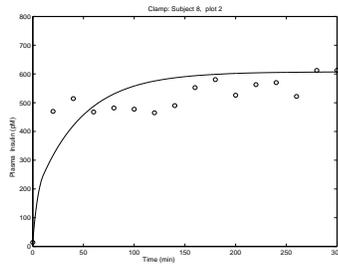
(g) Subject 7
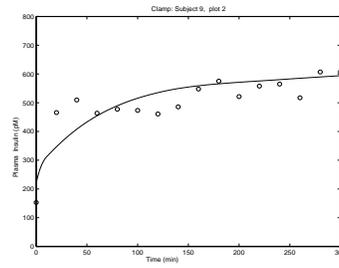
(h) Subject 8
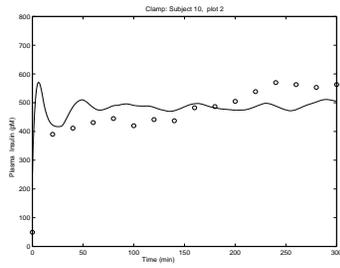
(i) Subject 9
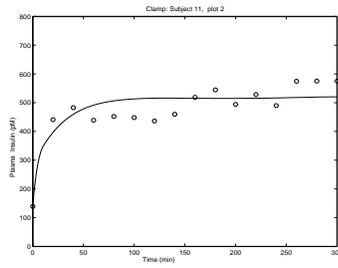
(j) Subject 10
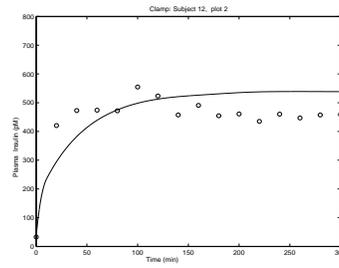
(k) Subject 11
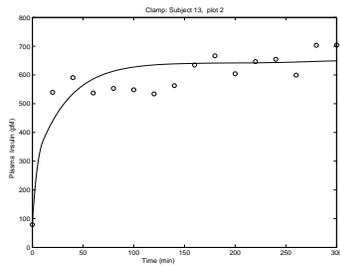
(l) Subject 12
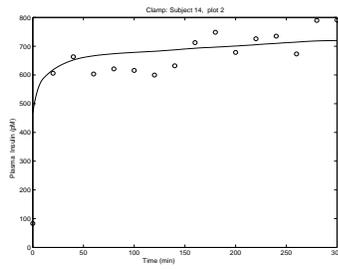
(m) Subject 13
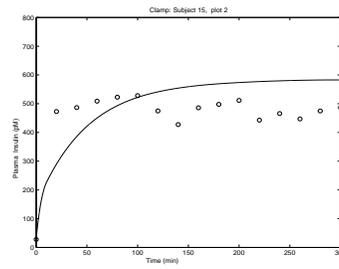
(n) Subject 14

(o) Subject 15

Figure 3: ODE model: observed (∘) and predicted (solid line) insulinemia corresponding to the IRWLS estimates for the case $(CV_G, CV_I) = (0.05, 0.15)$ (see Table 5).

(a) Subject 1   (b) Subject 2   (c) Subject 3

(d) Subject 4   (e) Subject 5   (f) Subject 6

(g) Subject 7   (h) Subject 8   (i) Subject 9

(j) Subject 10   (k) Subject 11   (l) Subject 12

(m) Subject 13   (n) Subject 14   (o) Subject 15
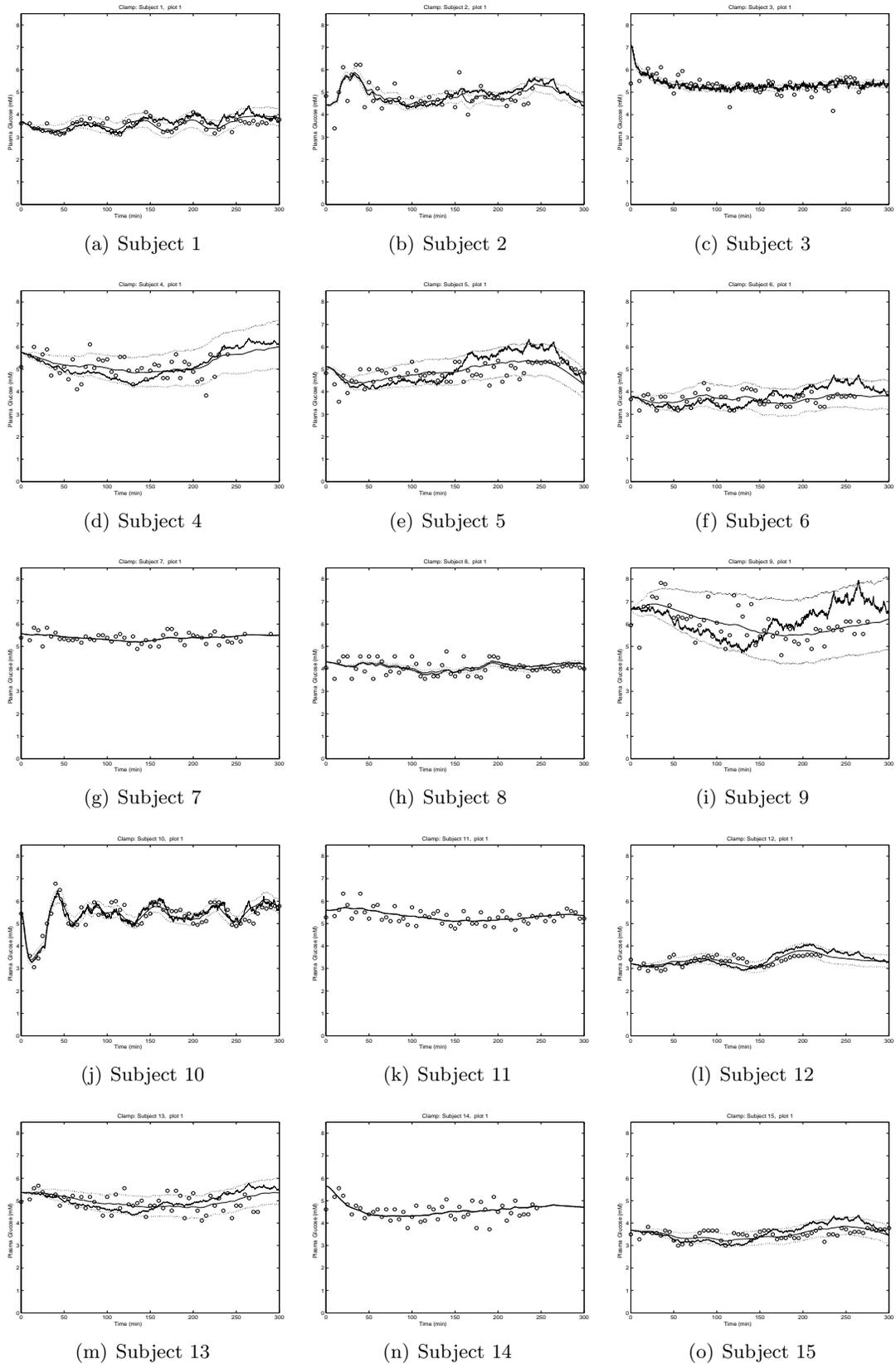
Figure 4: SDE model: a simulated trajectory of $G(t)$, empirical mean curve of the $G(t)$ process (smooth solid lines), empirical 95% confidence limits of the mean process (dashed lines) for the case $(CV_G, CV_I) = (0.05, 0.15)$ and glycemia observations.

(a) Subject 1 (b) Subject 2 (c) Subject 3

(d) Subject 4 (e) Subject 5 (f) Subject 6

(g) Subject 7 (h) Subject 8 (i) Subject 9

(j) Subject 10 (k) Subject 11 (l) Subject 12

(m) Subject 13 (n) Subject 14 (o) Subject 15
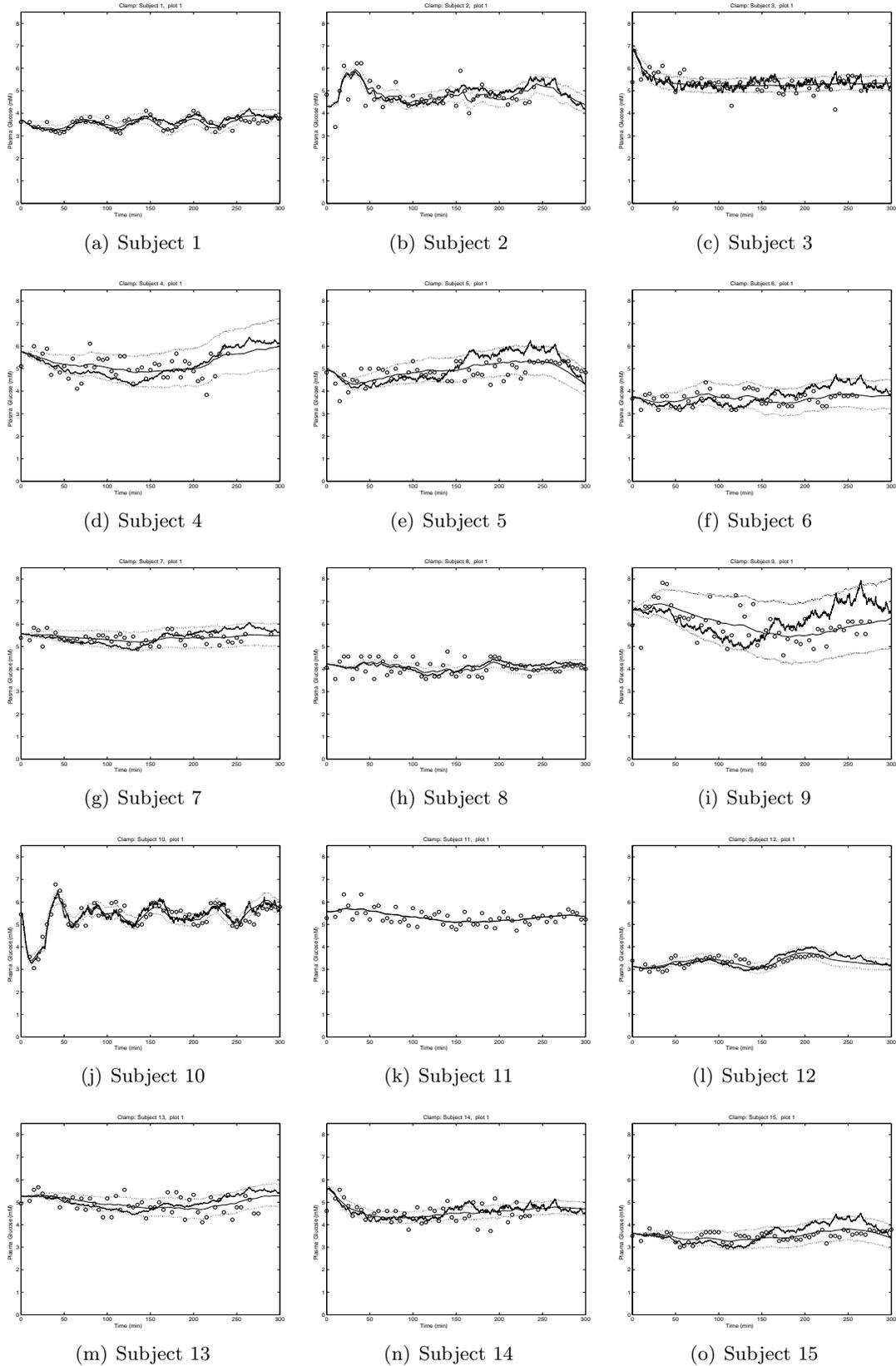
Figure 5: SDE model: a simulated trajectory of $G(t)$, empirical mean curve of the $G(t)$ process (smooth solid lines), empirical 95% confidence limits of the mean process (dashed lines) for the case $(CV_G, CV_I) = (0.03, 0.15)$ and glycemia observations.
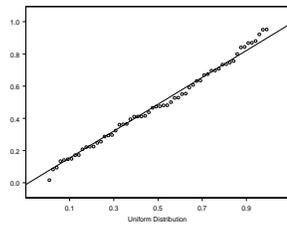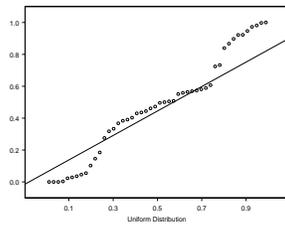
(a) Subject 1, $p = 0.658$ (b) Subject 2, $p = 0.2545$ (c) Subject 3, $p = 0.4677$
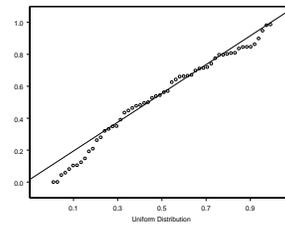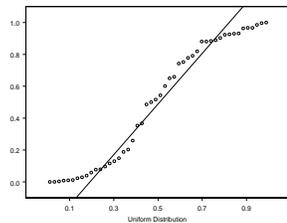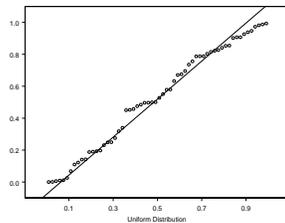
(d) Subject 4, $p = 0.0485$ (e) Subject 5, $p = 0.3502$ (f) Subject 6, $p = 0.0452$
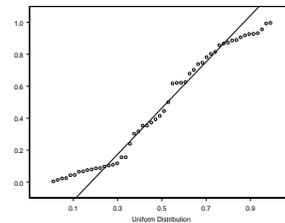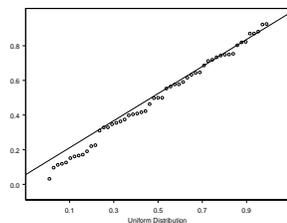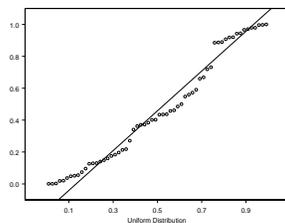
(g) Subject 7, $p = 0.7241$ (h) Subject 8, $p = 0.1418$ (i) Subject 9, $p = 0.2443$
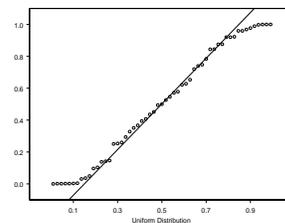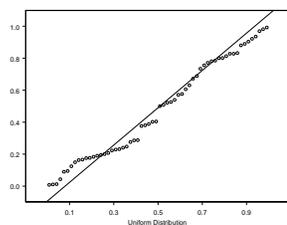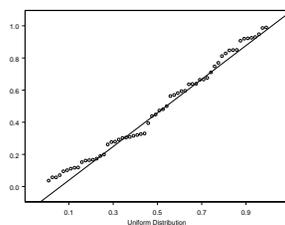
(j) Subject 10, $p = 0.2676$ (k) Subject 11, $p = 0.3553$ (l) Subject 12, $p = 0.1732$
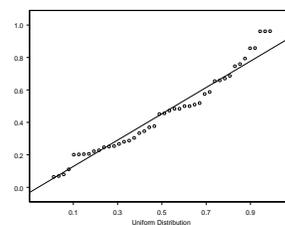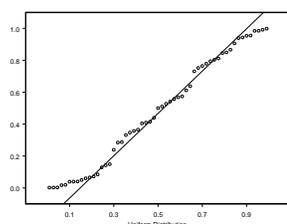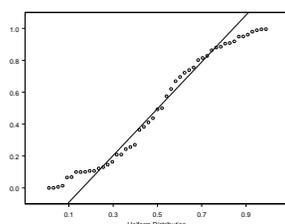
(m) Subject 13, $p = 0.1585$ (n) Subject 14, $p = 0.2475$ (o) Subject 15, $p = 0.2545$
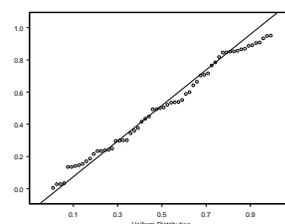
Figure 6: SDE model: q-q plots of the simulated glycemia residuals in the case $(CV_G, CV_I) = (0.05, 0.15)$ and p-values from the Kolmogorov-Smirnov goodness-of-fit test. If $p < 0.05$ the residuals do not conform the hypothesis of $U(0,1)$ distribution at a 5% confidence level.
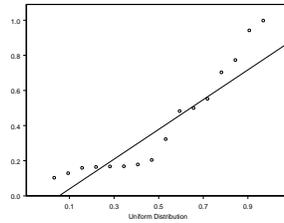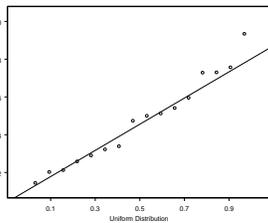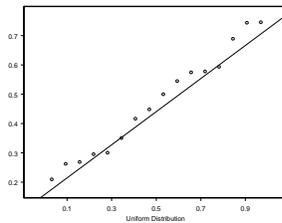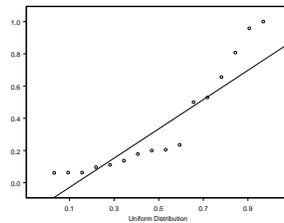
26.
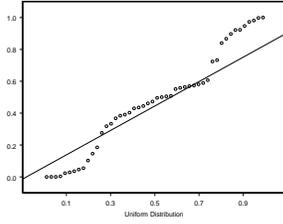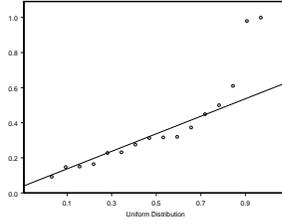
(a) Subject 1, $p = 0.0922$ (b) Subject 2, $p = 0.0657$ (c) Subject 3, $p = 0.3172$
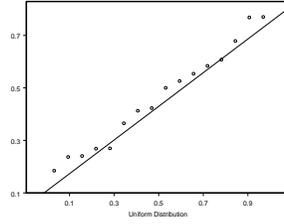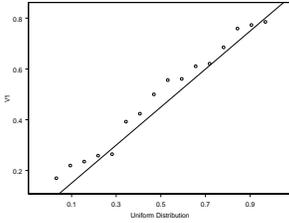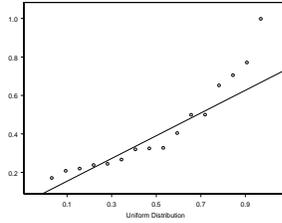
(d) Subject 4, $p = 0.3954$ (e) Subject 5, $p = 0.27$ (f) Subject 6, $p = 0.2847$
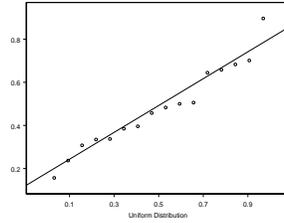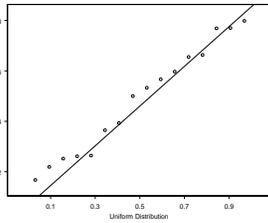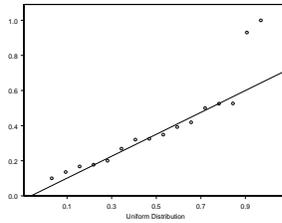
(g) Subject 7, $p = 0.4722$ (h) Subject 8, $p = 0.0305$ (i) Subject 9, $p = 0.1096$
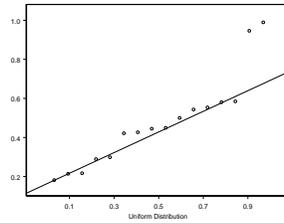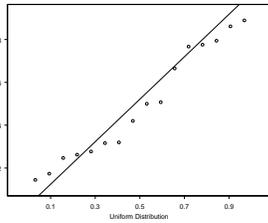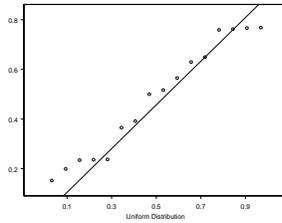
(j) Subject 10, $p = 0.8462$ (k) Subject 11, $p = 0.3077$ (l) Subject 12, $p = 0.0981$

(m) Subject 13, $p = 0.6089$ (n) Subject 14, $p = 0.2132$ (o) Subject 15, $p = 0.0105$

Figure 7: SDE model: q-q plots of the simulated insulinemia residuals in the case $(CV_G, CV_I) = (0.05, 0.15)$ and p-values from the Kolmogorov-Smirnov goodness-of-fit test. If $p < 0.05$ the residuals do not conform the hypothesis of $U(0, 1)$ distribution at a 5% confidence level.

References

[1] Y. Aït-Sahalia, "Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach," *Econometrica*, vol. 70, pp. 223–262, January 2002.

[2] R. Bergman, Y. Ider, C. Bowden, and C. Cobelli, "Quantitative estimation of insulin sensitivity," *Am. J. Physiol.*, vol. 236, pp. E667–E677, 1979.

[3] B. Bibby, M. Jacobsen, and M. Sørensen, "Estimating functions for discretely sampled diffusion-type models," in *Handbook of Financial Econometrics* (Y. Aït-Sahalia and L. Hansen, eds.), Amsterdam: North-Holland, 2005.

[4] B. Bibby and M. Sørensen, "Martingale estimation functions for discretely observed diffusion processes," *Bernoulli*, vol. 1, no. 1/2, pp. 17–39, 1995.

[5] D. Dacunha-Castelle and D. Florens-Zmirnou, "Estimation of the coefficients of a diffusion from discrete observations," *Stochastics*, vol. 19, pp. 263–284, 1986.

[6] M. Davidian and D. Giltinan, *Nonlinear models for repeated measurement data.* Chapman & Hall, 1995.

[7] R. DeFronzo and E. Ferrannini, "Insulin resistance. a multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease," *Diabetes Care*, vol. 14, pp. 173–194, 1991.

[8] R. DeFronzo, J. Tobin, and R. Andres, "Glucose clamp technique: a method for quantifying insulin secretion and resistance," *Am. J. Physiol.*, vol. 237, pp. E214–E223, 1979.

[9] S. Ditlevsen and A. De Gaetano, "Stochastic vs. deterministic uptake of dodecanedioic acid by isolated rat livers," *Bull. Math. Biol.*, vol. 67, pp. 547–561, 2005.

[10] O. Elerian, N. Shephard, and S. Chib, "Likelihood inference for discretely observed nonlinear diffusions," *Econometrica*, vol. 69, pp. 959–993, 2001.

[11] E. Ferrannini and A. Mari, "How to measure insulin sensitivity," *J. Hypertens.*, vol. 16, pp. 895–906, 1998.

[12] A. Gallant and J. Long, "Estimating stochastic differential equations efficiently by minimum chi-square," *Biometrika*, vol. 84, pp. 124–141, 1997.

[13] C. Gouriroux, A. Monfort, and E. Renault, "Indirect inference," *Journal of Applied Econometrics*, vol. 8, pp. 85–118, 1993.

[14] A. Hurn, K. Lindsay, and V. Martin, "On the efficacy of simulated maximum likelihood for estimating the parameters of stochastic differential equations," *Journal of Time Series Analysis*, vol. 24, no. 1, pp. 45–63, 2003.

[15] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations.* Springer, 1992.

[16] B. Øksendal, *Stochastic differential equations.* Springer, second ed., 2000.

28.

[17] A. Olson and J. Pessin, "Structure, function, and regulation of the mammalian facilitative glucose transporter gene family," *Annu. Rev. Nutr.*, vol. 16, pp. 235–256, 1996.

[18] A. R. Pedersen, "Uniform residuals for discretely observed diffusion processes," Tech. Rep. 292, Department of Theoretical Statistics, University of Aarhus, Denmark, 1994.

[19] A. R. Pedersen, "A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations," *Scand. J. Statist.*, vol. 22, no. 1, pp. 55–71, 1995.

[20] A. R. Pedersen, "Likelihood inference by monte carlo methods for incompletely discretely observed diffusion processes," Tech. Rep. 1, Department of Biostatistics, University of Aarhus, Denmark, 2001.

[21] U. Picchini, A. De Gaetano, S. Panunzi, S. Ditlevsen, and G. Mingrone, "A mathematical model of the euglycemic hyperinsulinemic clamp," *Theor. Biol. Med. Model.*, vol. 2, no. 44, 2005.

[22] W. Sacks in *Handbook of Neurochemistry* (A. Lajtha, ed.), vol. 1, New York: Plenum, 1969.

[23] I. Shoji and T. Ozaki, "Estimation for nonlinear stochastic differential equations by a local linearization method," *Stochastic Analysis and Applications*, vol. 16, pp. 733–752, 1998.

[24] M. Sørensen, "Prediction-based estimating functions," *Econometrics Journal*, vol. 3, pp. 123–147, 2000.

[25] A. Starke, "Determination of insulin sensitivity: methodological considerations," *J. Cardiovasc. Pharmacol.*, vol. 20, pp. S17–S21, 1992.

[26] T. Wallace and D. Matthews, "The assessment of insulin resistance in man," *Diabet. Med.*, vol. 19, pp. 527–534, 2002.

[27] N. Yoshida, "Estimation of diffusion processes from discrete observations," *Journal of Multivariate Analysis*, vol. 41, pp. 220–242, 1992.

[28] K. Zierler, "Whole body glucose metabolism," *Am. J. Physiol.*, vol. 276, pp. E409–E426, 1999.