2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Predictive power of indices derived from models of biological dynamic systems

Gianluigi Pillonetto and Claudio Cobelli

Abstract— **We are given a reliable parametric model of a system whose structure and parameter values can be obtained by an identification experiment. Often, one or more indices can be determined as a function of model parameters, i.e. an index is a function that maps the parameter space into the real line. The aim of these indices is to incorporate as much information as possible on a certain phenomenon of interest described by the model. The paper proposes an approach to compare competitive indices in terms of predictive power of system output. The new concept is applied to the minimal model of glucose kinetics, by comparing the performance of two different insulin sensitivity indices whose objective is to describe insulin ability to control glucose.**

Index Terms— **modeling methodology; biomedical systems; Markov chain Monte Carlo; glucose kinetics; diabetes**

I. Introduction

Parametric models of dynamic systems are ubiquitous in several fields of science, including engineering, physics, biology and medicine [1], [2], [3], [4]. Assume that we have a valid model structure. In particular, assume that its parameter values can be determined with an identification experiment and that validation tests suggest that the model is suitable for its intended use. Often, one or more indices are derived from the model. In this paper, with the term index, it is meant a function that maps the parameter space into the real line. The aim of an index is to provide a scalar able to incorporate as much information as possible on a certain phenomenon of interest described by the model. A classic example, also discussed in the sequel, is the minimal model of glucose kinetics (MM) [5], [6]. Since its inception in the late seventies MM has been employed in hundreds of papers to describe glucose and insulin dynamics after a glucose perturbation [7]. The model is very popular since it yields, in a relatively simple way, an index, named insulin sensitivity, which measures the ability of insulin to control glucose metabolism, thus turning out extremely important for physiological/clinical studies. For example, in human subjects it permits to assess or predict diseases such as diabetes as well as to monitor efficacy of therapies [8], [9], [10].

Consider now a scenario where two or more competitive indices are derived from the same mathematical model. To our knowledge, a problem that has not been fully appreciated so far consists of establishing which of the indices could be considered the most informative one to describe the phenomenon under study. Clearly, many different parameter values may lead to the same index value, i.e. in practice an index is never an injective map. The paper develops the idea that the index of choice should correspond to that possessing the best predictive power of the system output. For instance, in the case of MM the system output corresponds to glucose profile in plasma, often measured during an intravenous glucose tolerance test (IVGTT) [5]. Assume now that two sets of parameters, associated with two different subjects, gives the same value of insulin sensitivity. Then, the larger the predictive power of the index, the more similar the glucose time-courses and thus also the way glucose comes back to its basal value. Thus, according to our paradigm, a MM index with large predictive power will permit to accurately rank subjects on the basis of insulin efficacy in controlling glucose.

Gianluigi Pillonetto and Claudio Cobelli are with the Department of Information Engineering of the University of Padova, Italy.

The paper is so organized. In Section 2 first the notation used in the rest of the paper to describe a nonlinear dynamic model is illustrated and then index definition is provided. In Section 3 we show how the concept of predictive power of an index can be formalized mathematically and how index performance can be quantified by solving a suitable multidimensional integral. In Section 4 it is illustrated how such integral, that is in general analytically intractable, can be computed by Markov chain Monte Carlo (MCMC) techniques [11], [12]. In Section 5 the new concepts developed are applied to the MM of glucose kinetics to show that a new insulin sensitivity index recently proposed in [13] is more informative than the classical one that has been used in the literature for almost three decades. Conclusions can be found in Section 6 while further details related to some mathematical concepts introduced in the paper are reported in Appendices.

II. Model and index definition

Consider a nonlinear dynamical system

$$\begin{cases} \dot{x}(t) = A[x(t), u(t); \theta] \\ x(0) = x_0(\theta) \\ y(t) = B[x(t), u(t); \theta] \quad t \in [0, T] \end{cases} \tag{1}$$

where x is the n-dimensional state variable, $u(t)$ is a fixed and deterministic m-dimensional input while $y(t)$ is the scalar system output. As it will be also clear in the sequel, the theory developed could also easily deal with a scenario where the input is a stochastic process and $y(t)$ is multi-dimensional but we prefer to maintain our notation as simple as possible. In addition, θ is a p-dimensional parameter vector, modeled as a random vector of density $p_\theta : \Re^p \mapsto \Re$. Let also Y be

the space containing the output functions mapping $[0,T]$ into $\Re$. Then, we assume that, given $u(t)$, map induced by the model from parameter space to output space is well-defined and denote it with $h:\theta \mapsto Y$.
Our definition of index is provided below.

Definition 1: We call an index any function g that maps the parameter space into the real line, i.e. $g:\theta \mapsto \Re$
The aim of the next Section is to associate with g a measure of predictive capability of the system output y.

III. Predictive power of an index

A. Definition

Assume that an index g is given and let $\Omega \subseteq \Re^{2p}$. We denote with $\pi_g : \Re^{2p} \mapsto \Re$ the probability density function defined for every $\theta_1 \in \Re^p$ and $\theta_2 \in \Re^p$ by the following equation

$$\pi_g(\theta_1,\theta_2) = \frac{p_\theta(\theta_1)p_\theta(\theta_2)\chi_g(\theta_1,\theta_2)\chi_\Omega(\theta_1,\theta_2)}{c} \quad (2)$$

where here, and in the sequel, c denotes a suitable normalization factor while χ_g is given by

$$\begin{cases} \chi_g(\theta_1,\theta_2)=1 & if \quad g(\theta_1)=g(\theta_2) \\ \chi_g(\theta_1,\theta_2)=0 & otherwise \end{cases} \quad (3)$$

In addition, given a set S, χ_S is used to denote the indicator function of S. Thus, χ_Ω is defined as follows

$$\begin{cases} \chi_\Omega(\theta_1,\theta_2)=1 & if \quad (\theta_1,\theta_2)\in\Omega \\ \chi_\Omega(\theta_1,\theta_2)=0 & otherwise \end{cases} \quad (4)$$

One can think of realizations of θ_1 and θ_2 drawn from π_g in eq.(2) as parameters associated e.g. with two different subjects. Then, notice that χ_g forces to select subjects with the same index value. In addition, χ_Ω allows one to include extra constraints on the parameter space, e.g. to sample subjects having the same value for those components of θ the index is independent of. Usefulness of including χ_Ω in π_g will be in particular further elucidated in the next Section, when two different indices coming from MM will be compared.

We are now in the position to provide our definition of predictive power of an index.

Definition 2: Let index g and a set $\Omega \subseteq \Re^{2p}$ be given. Let also L^2 the classical Lebesque space of square integrable functions equipped with the usual norm $\|.\|_{L^2}$. We define the predictive power of g, and denote it with $pr[g]$, as

$$pr[g] = V^{-1}$$

where

$$V = \int_{\Re^{2p}} \frac{\|h(\theta_1)-h(\theta_2)\|_{L^2}}{\sqrt{T}} \pi_g(\theta_1,\theta_2)d\theta_1 d\theta_2$$

B. Reformulation of V

In this subsection we provide a reformulation of V that will be especially useful to elucidate the meaning of $pr[g]$. To this aim, being $p_\theta : \Re^p \mapsto \Re$ the probability density function of θ, we use the notation $p_{\theta,i} : \Re^{p+1} \mapsto \Re$ to denote the joint probability density of θ and i, where $i = g(\theta)$. We also use $p_{\theta|i} : \Re^{p+1} \mapsto \Re$ to denote the probability density of θ conditioned on i.
Now, let $\nu_g : \Re^{2p+2} \mapsto \Re$ the probability density function that is defined for every $\theta_1 \in \Re^p, i_1 \in \Re, \theta_2 \in \Re^p$ and $i_2 \in \Re$ by the following equation

$$\nu_g(\theta_1,i_1,\theta_2,i_2) = \frac{p_{\theta,i}(\theta_1,i_1)p_{\theta,i}(\theta_2,i_2)\chi_g(\theta_1,\theta_2)\chi_\Omega(\theta_1,\theta_2)}{c} \quad (5)$$

where

$$i_1 = g(\theta_1) \qquad i_2 = g(\theta_2) \quad (6)$$

Notice from eq.(3) and eq.(6) that the probability density function ν_g may assume values different from zero only if $i_1 = i_2$. It can also be easily seen that π_g corresponds to the marginal density of ν_g once i_1 and i_2 are integrated out, i.e.

$$\int_{\Re^2} \nu_g(\theta_1,i_1,\theta_2,i_2)di_1 di_2 = \pi_g(\theta_1,\theta_2) \quad (7)$$

In the light of eq. (7), V can be rewritten as follows

$$V = \int_{\Re^{2p+2}} \frac{\|h(\theta_1)-h(\theta_2)\|_{L^2}}{\sqrt{T}} \nu_g(\theta_1,i_1,\theta_2,i_2)d\theta_1 di_1 d\theta_2 di_2 \quad (8)$$

Denote with f the marginal density of i_1 and i_2, i.e.

$$f(i_1,i_2) = \int_{\Re^{2p}} \nu_g(\theta_1,i_1,\theta_2,i_2)d\theta_1 d\theta_2$$

Notice how f may assume values different from zero only in I where $I = \{i_1,i_2 | i_1 = i_2\}$. Just for sake of simplicity, assume that f takes on strictly positive values in I (otherwise a suitable subset of I can be considered). If θ_1,θ_2,i_1 and i_2 have density ν_g, we use $\nu_g(\theta_1,\theta_2|i_1,i_2)$ to denote the conditional density that turns out well defined in I and given by

$$\nu_g(\theta_1,\theta_2|i_1,i_2) = \frac{\nu_g(\theta_1,i_1,\theta_2,i_2)}{f(i_1,i_2)} \qquad for \qquad (i_1,i_2)\in I$$

Proposition 3: We have

$$V = \int_I \alpha(i_1,i_2)f(i_1,i_2)di_1 di_2$$

where, if $i_1 = i_2 = i$, $\alpha(i_1,i_2)$ is given by

$$\int_{\Re^{2p}} \frac{\|h(\theta_1)-h(\theta_2)\|_{L^2}}{\sqrt{T}} \frac{p_{\theta|i}(\theta_1|i)p_{\theta|i}(\theta_2|i)\chi_\Omega(\theta_1,\theta_2)}{c} d\theta_1 d\theta_2$$

Proof: Let $E[v]$ denote expectation of a random variable v. From eq. (8) it holds that

$$V = \frac{E[\|h(\theta_1) - h(\theta_2)\|_{L^2}]}{\sqrt{T}} \tag{9}$$

where expectation is taken with respect to density ν_g. In terms of conditional expectations (see e.g. [14]) we also have

$$V = \frac{E[E[\|h(\theta_1) - h(\theta_2)\|_{L^2}|i_1 = i, i_2 = i]]}{\sqrt{T}} \tag{10}$$

where expectations are computed with respect first to $\nu_g(\theta_1, \theta_2|i_1, i_2)$ and then to $f(i_1, i_2)$. Using eq.(5), for $i_1 = i_2 = i$, we obtain immediately

$$\nu_g(\theta_1, \theta_2|i_1 = i, i_2 = i) = \frac{p_{\theta|i}(\theta_1|i)p_{\theta|i}(\theta_2|i)\chi_\Omega(\theta_1, \theta_2)}{c} \tag{11}$$

This completes the proof.

The new expression of V so obtained elucidates the rationale underlying our definition of index predictive power. It shows how a particular realization i of an index is regarded as informative if, on average, samples independently drawn from $p_{\theta|i}(\theta_1|i)$ and $p_{\theta|i}(\theta_2|i)$, and falling in Ω, are mapped by h into similar output profiles, where similarity is regulated by the distance in L^2. Thus, α is a predictive measure since it quantifies information that i incorporates on the system output. Finally, since i is random as well, expression of V is obtained by averaging α with respect to density f.

IV. Computational issues

Density π_g (or ν_g) is possibly known only up to a normalizing constant. In addition, its shape may be much complicated since model described by eq.(1) as well as index g may be highly nonlinear. Thus, integral in eq.(2) may turn out analytically intractable and numerical integration may be unfeasible if dimension of θ is large. In this case Markov chain Monte Carlo (MCMC) techniques may be used to obtain V, and hence $pr[g]$, by the ergodic average

$$V \approx \frac{\sum_{k=1}^{N} \|h(\theta_1^k) - h(\theta_2^k)\|_{L^2}}{N\sqrt{T}} \tag{12}$$

where $\{\theta_1^k\}$ and $\{\theta_1^k\}$ are correlated samples collected along the path of a Markov chain converging in distribution to π_g [15].

All the different MCMC strategies proposed in the literature are special cases of the Metropolis-Hastings algorithm [12] that consists of two steps. First, a candidate sample is drawn from a proposal distribution whose choice is important since it establishes the rate of convergence of the algorithm. Then, the candidate point is accepted with a suitable probability, see e.g. [11] for details.

In our case, the choice of the proposal depends on the particular structure of model (1) as well as on the nature of index g and of set Ω. In the next Section we introduce two MCMC algorithms, which rely upon random walks proposals, able to return predictive capabilities of two MM insulin sensitivity indices.

V. Minimal model insulin sensitivity indices

A. Classical and new insulin sensitivity index

During an intravenous glucose tolerance test (IVGTT) MM equations are:

$$\begin{cases} \dot{G}(t) = -\left[S_G + X(t)\right] G(t) + G_b S_G \\ \dot{X}(t) = -p_2 \left[X(t) - S_I(I(t) - I_b)\right] \\ G(0) = G_0, \quad X(0) = 0 \\ y(t) = G(t) \end{cases} \tag{13}$$

In eq.(13), $G(t)$ $(mgdl^{-1})$ is glucose concentration in plasma and represents the output of the system. $I(t)$ (μUml^{-1}) is insulin concentration in plasma and is regarded as the forcing input. G_b and I_b are glucose and insulin baseline values, respectively, while G_0 accounts for the intravenous glucose dose injected at time 0. Glucose effectiveness S_G instead describes the glucose *per se* control, i.e. its ability to enhance its own rate of disappearance and to inhibit its endogenous production. Furthermore $X(t)$ denotes remote insulin, i.e. insulin action, whose dynamics are regulated by parameters p_2 (min^{-1}) and S_I $(min^{-1}\mu U^{-1}ml)$ that thus both provide information regarding efficacy of insulin control on glucose.

In the sequel, we use $h_M : \theta \mapsto \Re$, with $\theta = [S_I, S_G, G_0, p_2]$ to denote the map between parameter space and glucose profile. Furthermore, we assume that t varies on the interval $[0, 240]$ min, which is the typical duration of an IVGTT experiment.

Starting from MM, two different indices of insulin sensitivity have been derived. The classical one is S_I and it is well known that this index provides quantitative information on insulin sensitivity when insulin action is at steady state, i.e. $\dot{X}(t) = 0$ in eq.(13), see [6].

Recently, in [13] a new insulin sensitivity index has been derived that, at variance with the classic one, also accounts for how insulin action reaches its plateau value. This new index has been named Dynamic Insulin Sensitivity (S_I^D) and is defined in mathematical terms as

$$S_I^D = \eta(p_2) S_I \tag{14}$$

where

$$\eta(p_2) = \left[1 - \frac{1 - e^{-60p_2}}{60p_2}\right] \tag{15}$$

Notice how S_I^D is given by multiplication of S_I by η, a correction factor that was termed efficiency. Thus, while S_I measures the maximal metabolic response capacity of a given individual, S_I^D represents intuitively that fraction of the maximal capacity that is promptly available by virtue of the dynamic properties of insulin action, see [16], [13] for details.

B. Computation of S_I and S_I^D predictive power

By exploiting the many MM results coming from IVGTT studies performed in normal and diabetic subjects, see e.g. [17], [18], parameter vector θ can be modeled as a normal

random vector of mean μ and covariance Σ (values reported in Appendix A) whose realizations are constrained to be non-negative. Denoting with p_θ the probability density function of θ, it holds that

$$p_\theta(\theta) \propto N(\theta;\mu,\Sigma)\chi_{\{\theta>0\}}(\theta) \tag{16}$$

where $\propto$ stays for "proportional to" and $N(\theta;\mu,\Sigma)$ denotes a Gaussian density of mean μ and covariance Σ evaluated at θ. In addition, $\chi_{\{\theta>0\}}$ denotes the indicator function which equals 1 if all values of components of θ are positive and 0 otherwise.

In order to compare S_I and S_I^D, we consider couple of subjects where S_G and G_0 take on the same values. Our aim is to investigate if glucose profiles are more similar to each other when also S_I^D or S_I is the same in each couple. In this way it is possible to assess which index is able to better predict glucose time-course and thus to better measure insulin hypoglycaemic effect. This can be formalized in mathematical terms by exploiting the general framework developed in the previous Section. To this aim, define

$$\begin{aligned}\theta_1 &= [S_{I,1}, S_{G,1}, G_{0,1}, p_{2,1}]\\ \theta_2 &= [S_{I,2}, S_{G,2}, G_{0,2}, p_{2,2}]\end{aligned}$$

Since comparison has to be made among subjects with the same glucose *per se* control S_G and with the same glucose value at time zero G_0 coming from the injected dose, define

$$\Omega = \{\theta_1,\theta_2 | S_{G,1} = S_{G,2}, G_{0,1} = G_{0,2}\}$$

Now, let's define

$$\begin{cases}\chi_{S_I}(\theta_1,\theta_2) = 1 & if \quad S_{I,1} = S_{I,2}\\ \chi_{S_I}(\theta_1,\theta_2) = 0 & otherwise\end{cases}$$

and let also

$$\begin{cases}\chi_{S_I^D}(\theta_1,\theta_2) = 1 & if \quad \eta(p_{2,1})S_{I,1} = \eta(p_{2,2})S_{I,2}\\ \chi_{S_I^D}(\theta_1,\theta_2) = 0 & otherwise\end{cases}$$

Finally, according to eq.(2) the two probability density functions induced by indexes S_I and S_I^D are given, respectively, by

$$\pi_{S_I}(\theta_1,\theta_2) \propto p_\theta(\theta_1)p_\theta(\theta_2)\chi_{S_I}(\theta_1,\theta_2)\chi_\Omega(\theta_1,\theta_2) \tag{17}$$

$$\pi_{S_I^D}(\theta_1,\theta_2) \propto p_\theta(\theta_1)p_\theta(\theta_2)\chi_{S_I^D}(\theta_1,\theta_2)\chi_\Omega(\theta_1,\theta_2) \tag{18}$$

MCMC schemes can now be worked out in order to generate Markov chains converging in distribution to $\pi_{S_I}(\theta_1,\theta_2)$ and $\pi_{S_I^D}(\theta_1,\theta_2)$, thus permitting computation of $pr[S_I]$ and $pr[S_I^D]$. Details on our computational schemes as well as the pseudo-codes describing the two MCMC algorithms are reported in Appendix B.

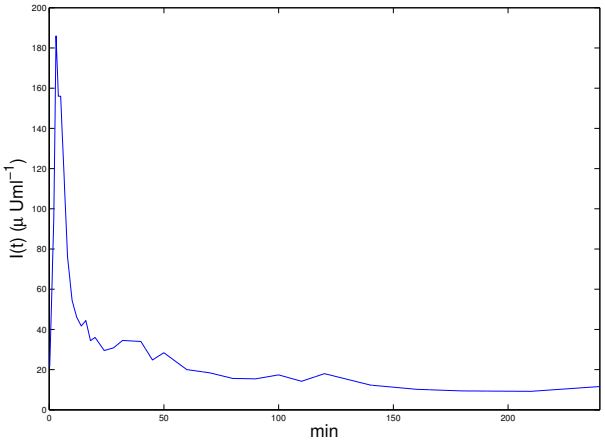

Fig. 1. Insulin time-course used during MCMC simulations

C. Results

During the simulation, the forcing input $I(t)$ in eq.(13) was set to the mean of real insulin time-courses measured in plasma during IVGTT experiments (see Figure 1)[1], while G_b was set to 80 $(mgdl^{-1})$.

Let $\{\theta_{1,S_I}^{(k)}, \theta_{2,S_I}^{(k)}\}$ and $\{\theta_{1,S_I^D}^{(k)}, \theta_{2,S_I^D}^{(k)}\}$ denote samples collected along the path of the Markov chain converging in distribution to $\pi_{S_I}(\theta_1,\theta_2)$ and $\pi_{S_I^D}(\theta_1,\theta_2)$, respectively. Then, we define

$$V_{S_I}^{(K)} = \sum_{k=1}^{K} \frac{\|h_M(\theta_{1,S_I}^{(k)}) - h_M(\theta_{2,S_I}^{(k)})\|_{L^2}}{K\sqrt{T}} \tag{19}$$

and

$$V_{S_I^D}^{(K)} = \sum_{k=1}^{K} \frac{\|h_M(\theta_{1,S_I^D}^{(k)}) - h_M(\theta_{2,S_I^D}^{(k)})\|_{L^2}}{K\sqrt{T}} \tag{20}$$

Figure 2 plots $V_{S_I}^{(K)}$ (solid line) and $V_{S_I^D}^{(K)}$ (dashed line), respectively, as a function of K. Estimated values of $pr[S_I]$ and $pr[S_I^D]$ turn out 0.093 and 0.19, respectively. As a matter of fact, S_I^D turns out to be twice as predictive as S_I, i.e. the sole knowledge of S_I^D carries much more information on the time-course of glucose profile in plasma than S_I does. This outcome appears relevant if one considers that MM index S_I has been used extensively in the literature for almost 30 years. These results show instead how MM mathematical structure allows definition of an index S_I^D which is a better summary of insulin ability to control glucose.

VI. Conclusions

Obtaining the most informative indices from a mathematical description of a certain phenomenon is

[1]We also considered other insulin profiles, e.g. measured in diabetic subjects during insulin modified IVGTT experiments. Obtained results (not shown) are in line with those presented in the paper.

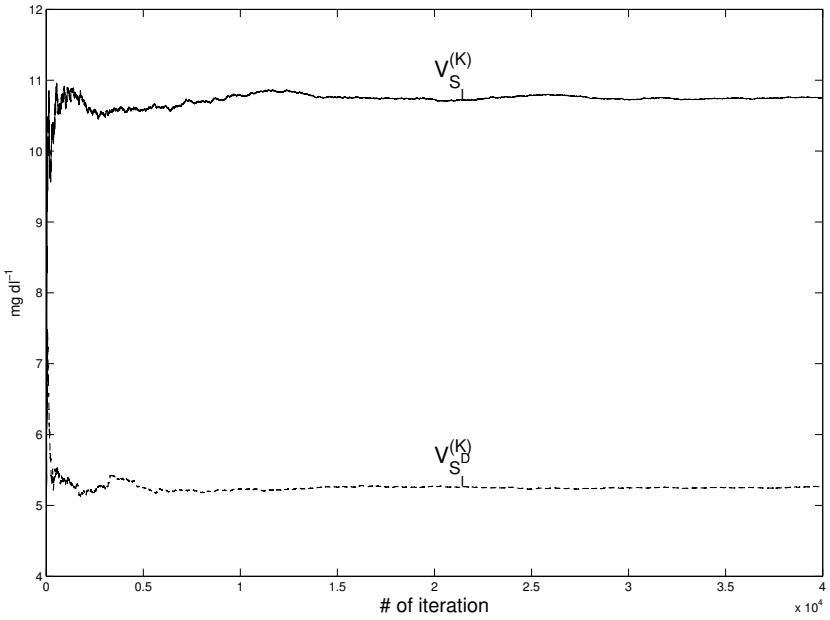

Fig. 2. $V_{S_I}^{(K)}$ (solid line) and $V_{S_I^D}^{(K)}$ (dashed line) as a function of K obtained by MCMC computation

an important problem that has not received the sufficient attention so far. In this paper we have developed a novel framework that allows comparison of competitive indices in terms of predictive power of system output. In addition, we have illustrated how these new concepts can be implemented by MCMC techniques.

The developed framework has been then exploited to illustrate how a recently proposed new insulin sensitivity index derived from MM, S_I^D, outperforms the classical one, S_I. Results obtained thus suggest that S_I^D should be used in place of S_I since it provides a more comprehensive picture of insulin control on glucose.

ACKNOWLEDGMENTS

This research was partially supported by the Italian Ministry of University and Research through the PRIN Project entitled "Tecniche ed applicazioni innovative di identificazione e controllo adattativo" and by the NIH grant EB-01975.

Appendix A

Numerical values of vector μ and matrix Σ, that allow one to fully define density p_θ in eq.(16), are reported below

$$\mu = \begin{pmatrix} 5e-4 & 2e-2 & 330 & 3e-2 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 1.6e-7 & 1.36e-6 & -0.01 & -1.55e-7 \\ 1.36e-6 & 1e-4 & -0.22 & -3.4e-5 \\ -0.01 & -0.22 & 2500 & -0.032 \\ -1.55e-7 & -3.4e-5 & -0.032 & 6.25e-4 \end{pmatrix}$$

Appendix B

We consider the problem of sampling the probability density function $\pi_{S_I}(\theta_1, \theta_2)$ reported in eq.(17). Recall that $\theta_1 = [S_{I,1}, S_{G,1}, G_{0,1}, p_{2,1}]$ and $\theta_2 = [S_{I,2}, S_{G,2}, G_{0,2}, p_{2,2}]$. Recall also that $\pi_{S_I}(\theta_1, \theta_2)$ assumes values equal to zero if the first three components of θ_1 do not exactly match the values of the first three components of θ_2. Our strategy then consists of designing a random walk converging in distribution to $\pi_{S_I}(\theta_1, \theta_2)$ by generating at every step identical proposal values for all the components of θ_1 and θ_2 but for $p_{2,1}$ and $p_{2,2}$. The resulting pseudo-code, able to return $pr[S_I]$, is reported below.

MCMC ALGORITHM FOR COMPUTING $pr[S_I]$

1) Define $\theta_1^0 = [S_I^0, S_g^0, G_0^0, p_{2,1}^0]$ and $\theta_2^0 = [S_I^0, S_g^0, G_0^0, p_{2,2}^0]$
2) Initialization: set $\Sigma_1 \in \Re^{3\times 3}$ and $\Sigma_2 \in \Re$ to suitable covariance matrices, θ_1^0 and θ_2^0 to the mean of p_θ, k to 1.
3) Iteration k
   - Sample $\psi_1 \in \Re^3$ from $N([S_I^{(k-1)}, S_g^{(k-1)}, G_0^{(k-1)}], \Sigma_1)$
   - Sample $\psi_2 \in \Re$ from $N(p_{2,1}^{(k-1)}, \Sigma_2)$
   - Sample $\psi_3 \in \Re$ from $N(p_{2,2}^{(k-1)}, \Sigma_2)$
   - Set $\xi_1 = [\psi_1 \quad \psi_2]$ and $\xi_2 = [\psi_1 \quad \psi_3]$
   - Sample w from a uniform distribution on $[0, 1]$
   - Set $\theta_1^{(k)}$ and $\theta_2^{(k)}$ as follows

$$\theta_1^{(k)} = \begin{cases} \xi_1 & \text{if } w \le \frac{\pi_{S_I}(\xi_1, \xi_2)}{\pi_{S_I}(\theta_1^{(k-1)}, \theta_2^{(k-1)})} \\ \theta_1^{(k-1)} & \text{otherwise} \end{cases}$$

$$\theta_2^{(k)} = \begin{cases} \xi_2 & \text{if } w \le \frac{\pi_{S_I}(\xi_1, \xi_2)}{\pi_{S_I}(\theta_1^{(k-1)}, \theta_2^{(k-1)})} \\ \theta_2^{(k-1)} & \text{otherwise} \end{cases}$$

4) Compute V and $pr[S_I]$ as

$$\begin{aligned} V &\approx \sum_{k=1}^{N} \frac{\|h_M(\theta_1^{(k)}) - h_M(\theta_2^{(k)})\|_{L^2}}{N\sqrt{T}} \\ pr[S_I] &= V^{-1} \end{aligned}$$

with N sufficiently large.

Now, let's consider the problem of drawing samples from the probability density function $\pi_{S_I^D}(\theta_1, \theta_2)$ reported in eq.(18). Recall that $S_I^D = \eta(p_2)S_I$ and denote ρ as the vector $[S_I^D, S_G, G_0, p_2]$. Define also

$$\begin{aligned} \rho_1 &= [S_{I,1}^D, S_{G,1}, G_{0,1}, p_{2,1}] \qquad S_{I,1}^D = \eta(p_{2,1})S_{I,1} \\ \rho_2 &= [S_{I,2}^D, S_{G,2}, G_{0,2}, p_{2,2}] \qquad S_{I,2}^D = \eta(p_{2,2})S_{I,2} \end{aligned}$$

Notice how vector ρ is given by an injective function that acts on θ by just mapping its first component S_I into S_I^D and maintaining the other components unchanged. By exploiting a well known result regarding transformation of random vectors (see e.g. [14]), the probability density function of ρ, given the density of θ reported in eq.(16), is

$$p_\rho(\rho) \propto \frac{\mathcal{N}([S_I^D/\eta(p_2), S_G, G_0, p_2]; \mu, \Sigma)}{\eta(p_2)} \chi_{\{\rho>0\}}(\rho) \quad (21)$$

where $\chi_{\{\rho>0\}}$ denotes the indicator function which equals 1 if all values of components of ρ are positive and 0 otherwise.

Recall from eq.(7) that $\pi_{S_I^D}(\theta_1, \theta_2)$ is a marginal density obtained by integrating out $S_{I,1}^D$ and $S_{I,2}^D$ from the joint density $\nu_g(\theta_1, i_1, \theta_2, i_2)$ defined in eq.(5) where $i_1 = S_{I,1}^D$ and $i_2 = S_{I,2}^D$. Now, it is useful to consider the problem of drawing samples from $\nu_g(\theta_1, i_1, \theta_2, i_2)$ once $S_{I,1}$ and $S_{I,2}$ are integrated out. As also stressed in the sequel, in this way the problem assumes the same structure of that previously solved to compute $pr[S_I]$ where the role of S_I is taken by S_I^D. Recalling the definition of $\nu_g(\theta_1, i_1, \theta_2, i_2)$ given in eq.(5), the desired probability density function denoted $\nu_{S_I^D}^\rho$ is

$$\begin{aligned} \nu_{S_I^D}^\rho(\rho_1, \rho_2) &\doteq \int \nu_g(\theta_1, i_1, \theta_2, i_2) dS_{I,1} dS_{I,2} \\ &= \frac{p_\rho(\rho_1) p_\rho(\rho_2) \chi_{S_I^D}^\rho(\rho_1, \rho_2) \chi_{\Omega^\rho}^\rho(\rho_1, \rho_2)}{c} \end{aligned}$$

where p_ρ is given by eq.(21) and

$$\begin{cases} \chi_{S_I^D}^\rho(\rho_1, \rho_2) = 1 \quad if \quad S_{I,1}^D = S_{I,2}^D \\ \chi_{S_I^D}^\rho(\rho_1, \rho_2) = 0 \quad otherwise \end{cases}$$

while $\chi_{\Omega^\rho}^\rho$ is the indicator function of the set Ω^ρ defined as

$$\Omega^\rho = \{\rho_1, \rho_2 | S_{G,1} = S_{G,2}, G_{0,1} = G_{0,2}\}$$

Now, the problem of generating a random walk converging in distribution to $\nu_{S_I^D}^\rho(\rho_1, \rho_2)$ is similar to the previous problem where $\pi_{S_I}(\theta_1, \theta_2)$ was involved. The corresponding pseudo-code is omitted for reason of space.

REFERENCES

[1] E. R. Carson, C. Cobelli, and L. Finkelstein, *The Mathematical Modeling of Metabolic and Endocrine Systems*. New York: Wiley, 1983.

[2] J. Jacquez, *Compartmental analysis in biology and medicine*. Ann Arbor: University of Michigan Press, 1985.

[3] L. Ljung, *System Identification - Theory For the User*. Prentice Hall, 1999.

[4] T. Soderstrom and P. Stoica, *System Identification*. Prentice Hall, 1989.

[5] R. Bergman, C. Bowden, and C. Cobelli, *Carbohydrate Metabolism*. Wiley, New York, 1981, ch. The minimal model approach to quantification of factors controlling glucose disposal in man.

[6] R. Bergman, Y. Ider, C. Bowden, and C. Cobelli, "Quantitative estimation of insulin sensitivity," *Am. J. Physiol. 236 (Endocrinol. Metab. Gastrointest. Physiol.)*, vol. 5, pp. E667–E677, 1979.

[7] R. Bergman, "Minimal model: perspective from 2005," *Hormone research*, vol. 64, pp. 8–15, 2006.

[8] S. Haffner, R. D. Jr, L. Mykkanen, R. Tracy, B. Howard, M. Rewers, J. Selby, P. Savage, and M. Saad, "Insulin sensitivity in subjects with type 2 diabetes. relationship to cardiovascular risk factors: the insulin resistance atherosclerosis study," *Diabetes Care*, vol. 22, pp. 562–568, 1999.

[9] L. Mykkanen, D. Zaccaro, C. Hales, A. Festa, and S. Haffner, "The relation of proinsulin and insulin to insulin sensitivity and acute insulin response in subjects with newly diagnosed type ii diabetes: the insulin resistance atherosclerosis study," *Diabetologia*, vol. 42, pp. 1060–1066, 1999.

[10] J. Raffael, D. Robbins, J. Norris, E. Boerwinkle, R. Defronzo, S. Elbein, W. Fujimoto, C. Hanis, S. Kahn, M. Permutt, K. Chiu, J. Cruz, D. Ehrmann, R. Robertson, J. Rotter, and J. Buse, "The gennid study: a resource for mapping the genes that cause niddm," *Diabetes care*, vol. 19, pp. 864–872, 1996.

[11] W. Gilks, S. Richardson, and D. Spiegelhalter, *Markov chain Monte Carlo in Practice*. London: Chapman and Hall, 1996.

[12] W. Hastings, "Monte carlo sampling methods using markov chain and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.

[13] G. Pillonetto, A.Caumo, G. Sparacino, and C. Cobelli, "A new dynamic index of insulin sensitivity," *IEEE Trans. on Biomedical Engineering*, vol. 53, pp. 369–379, 2006.

[14] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 1991.

[15] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*. Springer-Verlag, 1993.

[16] G. Pillonetto, G. Sparacino, and C. Cobelli, "Numerical non-identifiability regions of the minimal model of glucose kinetics: superiority of bayesian estimation," *Mathematical Biosciences*, vol. 184(1), pp. 53–67, 2003.

[17] G. Pillonetto, P. Magni, R. Bellazzi, G. Sparacino, and C. Cobelli, "Minimal model si=0 problem in niddm subjects: nonzero bayesian estimates with credible confidence intervals," *Am. J. Physiol. 282 (Endocrinol. Metab.)*, vol. 3, pp. E564–E573, 2002.

[18] P. Vicini and C. Cobelli, "The iterative two-stage population approach to ivgtt minimal modeling: improved precision with reduced sampling," *Am. J. Physiol. 280 (Endocrinol. Metab.)*, 2001.