Biomedical Signal Processing and Control 42 (2018) 63–72

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Robust insulin estimation under glycemic variability using Bayesian filtering and Gaussian process models

Luis Omar Avila[a,*], Mariano De Paula[b], Ernesto Carlos Martinez[c], Marcelo Luis Errecalde[a]

[a] LIDIC, Universidad Nacional de San Luis, Av. Ejército de los Andes 950–1[o] Piso, D5700HHW San Luis, Argentina
[b] INTELYMEC group, Centro de Investigaciones en Física e Ingeniería del Centro CIFICEN – UNICEN – CICpBA – CONICET, Argentina
[c] INGAR, CONICET-UTN, Avellaneda 3657, S3002 GJC, Santa Fe, Argentina

ARTICLE INFO

Article history:
Received 14 August 2017
Received in revised form 1 December 2017
Accepted 27 January 2018

Keywords:
Plasma insulin estimation
Stochastic model
Glycemic variability
Bayesian filtering
Gaussian processes

ABSTRACT

The ultimate goal of an artificial pancreas (AP) is finding the optimal insulin rates that can effectively reduce high blood glucose (BG) levels in type 1 diabetic patients. To achieve this, most autonomous closed-loop strategies continuously compute the optimal insulin bolus to be administrated on the basis of the estimated plasma concentrations for glucose and insulin. Unlike subcutaneous glucose levels which can be measured in real-time, unavailability of insulin sensors makes it essential the use of mathematical models so as to fully estimate plasma insulin concentrations. For model-based estimation, GP-Bayesian filters have been recently proposed to incorporate probabilistic non-parametric Gaussian process (GP) models of dynamic systems into Kalman filtering techniques. As a result, model uncertainty can explicitly be incorporated into the prediction step and in the filtering processes, which is usually not the case for more traditional filtering strategies that resort to parametric models for state estimation. More specifically, the question arises as to whether glycemic variability is properly taken into account in model formulations and whether it would compromise proper estimation of plasma insulin concentration. To tackle this, a stochastic glycemic model including variability was incorporated into different parametric and nonparametric filtering techniques to provide an estimate of the plasma insulin levels. In particular, we compared density representation against using knowledge about the parameterization of the transition dynamics and the observation function. We found that, as glycemic variability increases, filtering techniques based on parametric models rapidly degrades their performance as a consequence of large nonlinearities. Results show that Bayes' filtering techniques increase predictability of the patient state, and thus, boost safety and performance in the AP control and monitoring tasks.

1. Introduction

With existing sensing and pump technologies, widespread acceptance and use of an AP is steadily increasing and hopefully it will soon take part of routine clinical care [1]. The key control goal of an AP is the real-time calculation of the optimal insulin rates to be infused in type 1 diabetic patients so as to mimic the body's natural regulatory mechanism, i.e. BG levels between 70 and 140 [mg/dl]. To this aim, a number of control and monitoring strategies [2–6] has been proposed to compute optimal exogenous insulin infusion profiles on the basis of plasma glucose and plasma insulin estimation. For BG determination this is achieved by a continuous glucose monitor (CGM) that senses the glucose concentration in the interstitial area, and later on by considering the dynamics between this determination and the actual plasmatic concentration [7]. However, unlike plasma glucose which can be measured in real-time, the lack of specific sensors to determine plasma insulin levels makes the use of mathematical or inductive models for inferring insulin concentration the alternative of choice.

For obtaining plasma insulin estimations, a minimal parametric glucose-insulin dynamic model can be used in an open-loop configuration [8]. The main limitation of this state estimation strategy is that available CGM data are not taken into account to adapt the parameters of the model employed, despite these adjustments are mandatory in a dynamic system exhibiting significant levels of variability and complex regulatory behavior. To improve state

* Corresponding author.
E-mail addresses: loavila@unsl.edu.ar (L.O. Avila), mariano.depaula@fio.unicen.edu.ar (M. De Paula), ecmarti@santafe-conicet.gov.ar (E.C. Martinez), merreca@unsl.edu.ar (M.L. Errecalde).

https://doi.org/10.1016/j.bspc.2018.01.019
1746-8094/

estimation of a diabetic patient, plasma insulin concentration has recently been estimated from BG data using Bayesian filtering techniques which allow improving the estimation of glycemic conditions in real-time [9,10]. Filtering techniques are based on the proper combination of a dynamic model of the system and a state observer and have enjoyed remarkable success in the estimation of hidden states for different types of biomedical systems [11,12].

A number of Bayesian filtering techniques for nonlinear dynamic systems have been proposed and extensively studied [13,14]. The key issue in a Bayesian filter operation is the propagation of a Gaussian density function through the system dynamics. In the Extended Kalman Filter (EKF) the state distribution is represented by a Gaussian, which is then fully propagated through the first-order Taylor series expansion, that is, linearization of a nonlinear system dynamics [15]. In turn, the Unscented Kalman Filter (UKF) addresses state estimation by using a deterministic sampling approach, where the probability distribution of states is represented using a set of sample points. Noteworthy, the foregoing Kalman filters for state estimation are based on known parametric models of the state transition and measurement functions. However, for most nonlinear systems accurate parametric models are never readily available to describe all the (hidden) aspects of their dynamics. A feasible solution when facing nonlinear dynamics is the use of an approximated model in a nonparametric approach based on Gaussian processes (GP) models [16]. The so-called GP-Bayesian filters incorporate probabilistic non-parametric GPs models for states into the design EKF and UKF techniques[17]. In this manner, model uncertainty can explicitly be incorporated into the state prediction and the filtering steps.

Poor predictability of the glucose-insulin dynamics in a diabetic patient is a key issue that any control and monitoring strategy, to be implemented in an AP, should be able to address. Therefore, it is of significance and concern whether excessive variability might affect the estimation of plasma insulin concentration and, in consequence, compromise safety and performance of an AP operation. In this work, a stochastic version of the well-known Hovorka glucose-insulin model [18] was incorporated into parametric and nonparametric Bayesian filtering techniques to provide a real-time estimate of the plasma insulin concentration. Better understanding the effect of BG variability on the error between a given model, describing glucose-insulin interactions, and the real and complex physiologic system is of great significance for accelerating the acceptance of an AP.

This article is structured as follows. Section 2 introduces a stochastic model for describing the glucose-insulin dynamics in diabetic patients. Section 3 provides an overview of Bayesian filtering and GP regression models used to capture the underlying latent function for state transitions. In Section 4 different Bayesian filtering techniques are presented. In particular, we evaluate whether the filter propagates the full densities on the system dynamics (EKF and GP-ADF) or resorts to a sampling approach (UKF and GP-UKF). Also, it is analyzed whether the filter has full knowledge of the parameterization of the transition and measurement functions (EKF and UKF) or it uses a Gaussian approximation (GP-UKF and GP-ADF). In Section 5, results obtained for plasma insulin estimation in a simulation environment are shown and discussed. Finally, in Section 6 some remarks and future research efforts are outlined.

2. Stochastic model of the glucose-insulin dynamics

In this section, the reference deterministic model of the glucose-insulin dynamics based on the work of Hovorka et al. [18] is first presented. Later on, a stochastic diffusion process to model glycemic variability in synthetic diabetic patients is discussed. The Hovorka glucose-insulin model is a nonlinear compartmental numerical model with two inputs, insulin and glucose intake, and one output, glycemia. Particularly, the insulin-glucose interaction is nonlinear and it is given as

$$\begin{aligned}\frac{dQ_1(t)}{dx} &= -\left[\frac{F_0}{V_G G(t)} + x_1(t)\right] Q_1(t) + k_{12}Q_2(t) - F_R + U_G(t) + EGP \quad [1 - x_3(t)] \\ \frac{dQ_2(t)}{dx} &= x_1(t)Q_1(t) - [k_{12} + x_2(t)] \quad Q_2(t)\end{aligned} \tag{1}$$

where $Q_1(t), Q_2(t)$ represent the amounts of glucose in the accessible and non-accessible compartments, respectively, whereas k_{12} is the transfer rate constant and EGP is the parameter for endogenous glucose production; F_0 is a parameter that represents the total non-insulin dependent glucose flux and F_R represents renal glucose clearance above the glucose concentration threshold of 9 [mmol/l]. In turn, $U_G(t)$ is carbohydrate absorption rate, $x_1(t)$ is the remote effect of insulin on the rate of glucose transport, while $x_2(t)$ and $x_3(t)$ account for the elimination and endogenous glucose production, respectively. For space consideration, the remaining subsystems are not given, but they are fully described in [18]. BG measurements are then given as

$$G(t) = \frac{Q_1(t)}{V_G} \tag{2}$$

where V_G is glucose distribution volume. Table 1 shows the parameters used in the glucose-insulin interaction model of a type 1 diabetic patient, as given in [18]. The current patient state can be summarized in the vector $\mathbf{x}(t)$ in Eq. (3), where the entries are different state variables.

$$\mathbf{x}(t) = [S_1(t), S_2(t), I(t), x_1(t), x_2(t), x_3(t), Q_1(t), Q_2(t), G(t)] \tag{3}$$

As CGM determines BG levels in the interstitial fluid and the glucose exchange across the capillary walls occurs, by a simple but not instantaneous diffusion across a concentration gradient with a time-lag τ_{lag}, we have that the interstitial concentration is given as

$$\frac{dIG(t)}{dx} = \frac{1}{\tau_{lag}}(G(t) - IG(t)) \tag{4}$$

Finally, the obtained glucose profile is multiplied by a random time-varying calibration error $\xi(k)$ and later corrupted by an additive noise sequence sampled from a zero mean white Gaussian noise process $v(k)$, that is:

$$CGM(t) = \left(1 + \xi(t)\right) IG(t) + v(t) \tag{5}$$

For the development of efficient control and monitoring strategies of an AP, the deterministic glucose-insulin model might be enhanced by taking into account the variable behavior of patient metabolism [19–21]. An effective, yet simple alternative way to describe such fluctuating behavior is modeling temporal variability through a stochastic diffusion process. Ito [22] provided an alternative to ordinary numerical rules of calculus by defining a particular kind of uncertainty representation based on the Wiener diffusion process. Accordingly, the system transition function is described as a controlled Ito's diffusion process of the form

$$\dot{\mathbf{x}}(t) = \mathrm{a}\left(\mathbf{x}(t), u(t)\right) + \sigma d\omega \tag{6}$$

Table 1
Set of model parameters.

Parameter	Value	Unit
k_{12}	0.066	[min−1]
V_G	0.16*BW	[l]
EGP	0.0161	[mmol/min]
F_0	0.8507	[mmol/min]
F_R	$0.003(G(t) - 9)V_G$	[mmol/min]
τ_{lag}	5	[min]
$\xi(k)$	2	%

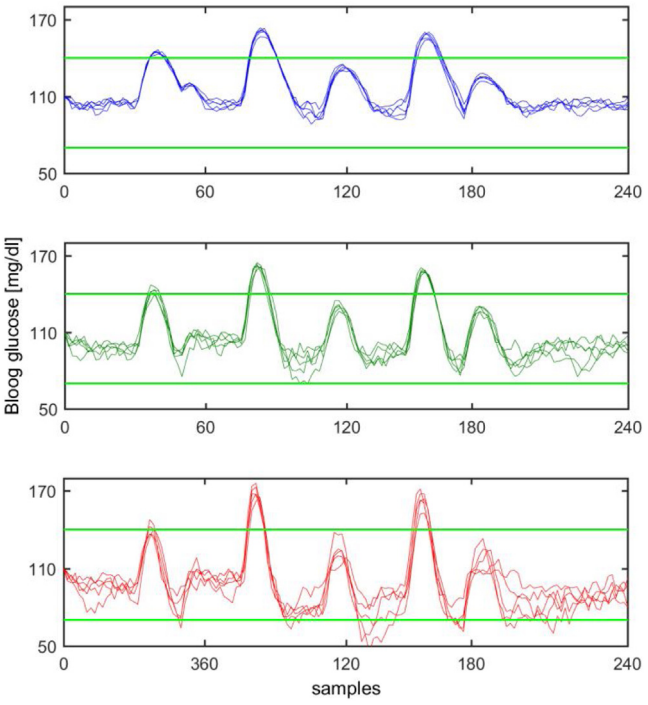

Fig. 1. BG profiles of diabetic patients using different scales of variability, $\sigma = 0.10$ (blue), 0.25 (green) and 0.50 (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where $a(\mathbf{x}(t), u(t))$ represents the glucose-insulin transition function in our implementation, $d\omega \in \mathbb{R}^{n_u}$ is the increment of a Wiener process and σ denotes its scaling parameter. Notice that, since noise and control signals acts over the same state space $\mathbb{R}^{n_u}$, any glycemic state can be achieved by the effect of either inherent glycemic variability or by exogenous actions (i.e. insulin boluses and/or carbohydrate intakes). In this way, to account for glycemic variability the Ito's process parameter is included into the model formulation such that it affects the amount of glucose in the accessible compartment

$$\frac{dQ_1(t)}{dx} = -\left[\frac{F_0}{V_G G(t)} + x_1(t)\right] Q_1(t) + k_{12} Q_2(t) - F_R + U_G(t) + EGP \quad [1 - x_3(t)] + \sigma \, d\omega \tag{7}$$

To represent different levels of variability in glycemic behavior, in Fig. 1 three implementations of the presented Hovorka's stochastic model are obtained for different scales of the Ito noise parameter σ. This mathematical model ensures a cohort of *in silico* diabetic patients that accounts sufficiently well for the observed inter and intra glycemic variability. To simulate an increase in the glycemic variability, the noise scale parameter is progressively augmented from $\sigma = 0.10$ (blue) to 0.25 (green) and 0.50 (red), whereas all the remaining parameters are not modified. It is clear that as the value of the parameter σ grows the magnitude of fluctuations in the BG profiles also increases.

3. Bayesian filtering with GP models

If the parametric models used to predict the latent states differ from the true underlying process, common filtering techniques may provide poor estimates. GP regression can be applied directly to the problem of learning the prediction and observation models required by the Bayesian filter approach in a nonparametric manner. GP models often make better use of available information since they can capture the uncertainty in the latent function

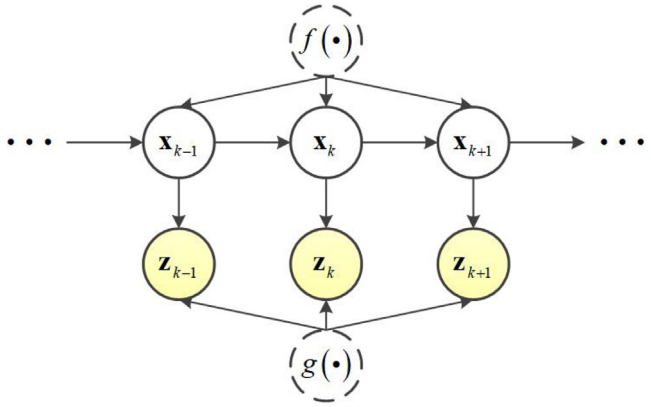

Fig. 2. Graphical model of a nonlinear dynamic system: the yellow nodes $\mathbf{z}_k$ are observed variables whereas the other nodes correspond to latent variables $\mathbf{x}_k$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

for state transitions, and so they can lead to very data-efficient state estimation despite the presence of unmodeled dynamics.

3.1. Overview of Bayes' filters

Stochastic processes are usually described by expressions with noise in both the transition dynamics and the measurement functions as

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) + w_k \quad w_k \sim N(0, \Sigma_w) \tag{8}$$

$$\mathbf{z}_k = g(\mathbf{x}_k) + v_k \quad v_k \sim N(0, \Sigma_v) \tag{9}$$

where $f(\mathbf{x}_{k-1})$ is the transition function and $g(\mathbf{x}_k)$ is the measurement function. In the classical Bayesian filtering approach [23], the latent state $\mathbf{x}_k \in \mathbb{R}^D$ is assumed to follow a first-order Markov process, whereas $\mathbf{z}_k \in \mathbb{R}^M$ is the current sensor measurements. Here, w_k and v_k are the process and observation white Gaussian additive noises. In Fig. 2, a graphical model of the considered nonlinear dynamic system with hidden states and their observations is depicted. In the figure, the causal dependencies between variables are represented by arrows. The dashed nodes represent the functions f and g, which can either be observed or latent depending on the model used.

Bayesian filters recursively estimate posterior distributions for the system state $\mathbf{x}_k$, conditioned on all sensor information collected so far, and propagates the posterior in time according to

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \propto p(\mathbf{z}_k|\mathbf{x}_k) \int p(\mathbf{x}_k|\mathbf{x}_{k-1})\, p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k})\, d\mathbf{x}_{k-1} \tag{10}$$

where conditional distributions $p(\boldsymbol{x}_k|\boldsymbol{z}_{1:k}) \approx N\left(\boldsymbol{x}_k|\mu_k, \Sigma_k^{\boldsymbol{x}}\right)$ are Gaussian approximations of the latent state. At this point, it is worth highlighting the difference between the prediction step (moving from $\mathbf{x}_{k-1}$ to $\mathbf{x}_k$) and the filtering step (going from $\mathbf{z}_k$ to $\mathbf{x}_k$). Hence to predict, we recursively determine the distribution $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ based on the information of the previous filter results $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$. In such a case, the Bayes rule yields to

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})\, p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})\, d\mathbf{x}_{k-1} \tag{11}$$

The filter update step determines the distribution $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ of the hidden state $\mathbf{x}_k$ based on collected observations from all previous and the current time periods. Thus, the Bayes' rule gives rise to

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k)\, p(\mathbf{x}_k|\mathbf{z}_{1:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{1:k-1})} \tag{12}$$

where the likelihood $p(\mathbf{z}_k|\mathbf{x}_k)$ is defined according to Eq. (9) and the prior belief $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$is the result of the prediction step in Eq. (11). Expressions (11) and (12) usually does not admit a closed-form solution and thus cannot be computed exactly and require resorting to approximate methods.

3.2. Gaussian process regression

When working with dynamic systems, observation data can be collected in training sets $\{X, Y\}$, where $X: \{\mathbf{x}_i \in \mathbb{R}^d / i = 1, 2, \ldots, n\}$is the set of input vectors,$Y: \{y_i \in \mathbb{R} / i = 1, 2, \ldots, n\}$ is the set of corresponding observations and n is the number of samples. The goal is to infer an inductive model h of the latent process that gives rise to the observed noisy data and assuming that the observations are generated according to$y_i = h(\mathbf{x}_i) + \varepsilon$, with$\varepsilon \sim N(0, \sigma_\varepsilon^2)$. Similarly to a Gaussian distribution, which is fully specified by a mean vector and a covariance matrix, a GP is specified by a mean function $\mu(\bullet)$and a covariance function $\Sigma(\bullet, \bullet)$, also known as the kernel.

Considering that the function $h(\mathbf{x})$is GP distributed $h(\mathbf{x}) \sim GP_h(\mu, \Sigma)$, we are interested in predicting its value for an arbitrary input vector x_*. The predictive (marginal) distribution is also Gaussian distributed with mean and variance given by

$$\mu_h(\mathbf{x}_*) = \Sigma(\mathbf{x}_*, X) + (\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1} Y \tag{13}$$

$$\sigma_h(\mathbf{x}_*) = \Sigma(\mathbf{x}_*, \mathbf{x}_*) + \Sigma(\mathbf{x}_*, X)(\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1} \Sigma(X, \mathbf{x}_*) \tag{14}$$

where $\mathbf{K}$ is the kernel matrix with $K_{ij} = \Sigma\left(\mathbf{x}^i, \mathbf{x}^j\right) \forall \mathbf{x} \in X$. A common kernel function, and which is going to be used hereafter for simulation experiments, is the squared exponential (SE)

$$\Sigma_{SE}\left(\mathbf{x}^i, \mathbf{x}^j\right) = \zeta^2 \exp\left[-\frac{1}{2}\left(\mathbf{x}^i - \mathbf{x}^j\right)^T \boldsymbol{\Lambda}\left(\mathbf{x}^i - \mathbf{x}^j\right)\right] \tag{15}$$

where the matrix $\boldsymbol{\Lambda} = diag\left(\left[\ell_1^2, \ell_2^2, \ldots, \ell_{(n_x)}^2\right]\right)$ contains the characteristic length scales and ζ^2 describes the variance of the inductive model h. To fit the hyperparameters of the covariance function to data, the evidence maximization approach is typically used (see [24] for details).

It is worth noting that when the test inputs are noisy$\boldsymbol{x}_* \sim N(\mu, \Sigma)$, so the mapping $\mathbf{x}_{k-1} \to \mathbf{x}_k$ is not fully known, we need to take the input uncertainty into account by Bayesian averaging according to the GP distribution. Thus, the prediction regression problem corresponds to finding the predictive distribution

$$p(h(\mathbf{x}_*)|\mu_*, \Sigma_*) = \int p(h(\mathbf{x}_*)|\mathbf{x}_*) p(\mathbf{x}_*|\mu, \Sigma) d\mathbf{x}_* \tag{16}$$

This distribution can be approximated by a Gaussian with the exact mean and the exact covariance matrix. For details refer to results given in [25].

3.3. Prediction and observation models with GPs

Non-parametric probabilistic GP models can describe distributions over all functions that plausibly explain insulin and glycemic fluctuations in different diabetic patients. The training data for each GP function is a set of input-output pairs. The prediction model maps the current latent state $\mathbf{x}_k$ to the state change or state transition $\Delta\mathbf{x}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$. The next state is found by adding the change to the previous state. In turn, the observation model maps from the latent state $\mathbf{x}_k$ to the observation $\mathbf{z}_k$. The appropriate form of the prediction and observation training data sets is

$$D_f = \langle X, X' \rangle \tag{17}$$

$$D_g = \langle X, Z \rangle \tag{18}$$

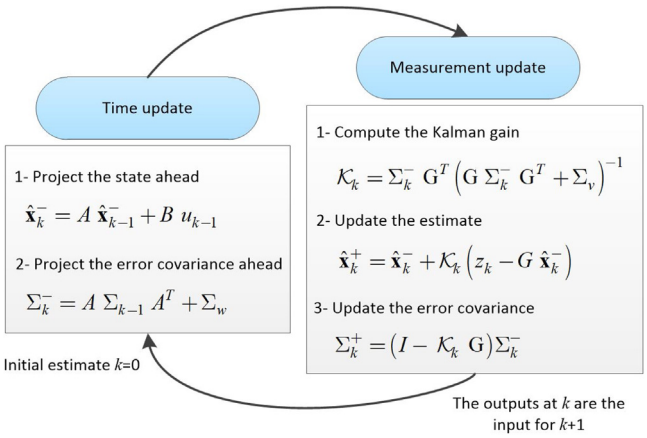

Fig. 3. The Kalman Filter algorithm: the time update projects the state estimation forwards in time while the availability of a new measurement allows updating this projected estimation.

where X is a matrix containing ground truth states, X' = [Δx_1, Δx_2, . . ., Δx_k] is a matrix containing transitions made from those states and Z is the matrix of observations made when the system is at the corresponding states X. The resulting GP prediction and observation models are, respectively,

$$p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) \approx GP_f(\mu(\boldsymbol{x}_{k-1}, D_f), \Sigma(\boldsymbol{x}_{k-1}, D_f)) \tag{19}$$

$$p(\boldsymbol{z}_k|\boldsymbol{x}_k) \approx GP_g\left(\mu\left(\boldsymbol{x}_{k-1}, D_g\right), \Sigma\left(\boldsymbol{x}_{k-1}, D_g\right)\right) \tag{20}$$

4. Bayes' filters implementations

The crucial issue for accurate estimation is the availability of an appropriate parametric model although, for most nonlinear cases, an accurate parametric model is rarely available. An alternative, when facing with non-linear systems, is the use of an approximated model by means of a nonparametric method. Here, different parametric and nonparametric filtering techniques are considered. Firstly, we made a distinction between those methods that use a full density representation of the underlying transition and observation functions (EKF and GP-ADF) from those that employ a sampling strategy to estimate the posterior distribution (UKF and GP-UKF). Also, methods that assume complete knowledge of the parameterization of the transition and the observation functions (EKF and UKF) are distinguished from those that use a GP model for such functions (GP-ADF and GP-UKF).

4.1. Nonlinear Kalman filters

The KF recursively estimates the current system state by using a form of feedback control, where the filter first estimates the system state at some time and then obtains feedback in the form of noisy measurements [26]. The time update equations are responsible for projecting forward the current state and its covariance, so as to obtain the a priori estimate $\hat{\boldsymbol{x}}k-$ for the state in the next time step. The measurement update equations are responsible for incorporating the new measurement into the a priori estimate to obtain an improved a posteriori estimate of the state distribution $\hat{\boldsymbol{x}}k+$. Indeed the algorithm resembles the predictor-corrector method shown in Fig. 3.

A KF that linearizes about the current mean and covariance is referred to as an EKF [26]. Similarly to a Taylor series, it is possible to linearize around the current state estimation using the partial derivatives of the transition and measurement functions, and thus, computing estimates even in the face of non-linear dynamics. Given

the transition and measurement functions in Eqs. (8) and (9), the corresponding partial derivatives can be computed as follows

$$F[k-1] = \frac{\partial f}{\partial \mathbf{x}_{k-1}}\Big|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1}|k-1}$$
$$G[k] = \frac{\partial g}{\partial \mathbf{x}_k}\Big|_{\mathbf{x}_k=\hat{\mathbf{x}}_k|k-1} \quad (21)$$

However, if the model dynamics is strongly nonlinear, the EKF will face severe performance degradation as errors are amplified forward in time [27]. This is the result of propagating the covariance through a linearized model of a nonlinear dynamic system which leads to numerical instability.

Instead of a linearization using the Jacobian matrices, the UKF uses a deterministic sampling approach to capture the mean and covariance estimates with a minimal set of weighted sigma points [28]. The nonlinear transformation of these points is intended to be an estimation of the posterior distribution. This transformation is known as the *unscented transform*. The basic premise behind the UKF is that it results easier to approximate a Gaussian distribution than to approximate an arbitrary nonlinear function. The number of sigma points is $r = 2n_x + 1$, where n_x represents the number of state variables. Considering the nonlinear process represented by Eqs. (8) and (9), the sigma points X and their corresponding weights W are given by

$$\begin{aligned} X^0 &= \hat{\boldsymbol{x}} & W_0^{\mu} &= \lambda/(n_x+\lambda) \\ X^i &= \hat{\boldsymbol{x}} + \left(\sqrt{(n_x+\lambda)+P_x}\right)_i \; i = 1,\ldots,n_x & W_0^{\Sigma} &= \lambda/(n_x+\lambda) + \left(1-\alpha^2+\beta\right) \\ X^i &= \hat{\boldsymbol{x}} - \left(\sqrt{(n_x+\lambda)+P_x}\right)_{i-n_x} \; i = n_x+1,\ldots,2n_x & W_i^{\mu} &= W_i^{\Sigma} = 1/2(n_x+\lambda) \end{aligned} \quad (22)$$

where W^{μ} is used to reconstruct the predicted mean and W^{Σ} used for the predicted covariance. Moreover, $\lambda = \alpha^2(1+\kappa) - n_x$ is the scaling parameter; α has a small positive value which represents the spread of sigma points around $\hat{\boldsymbol{x}}$; κ is the second calibration parameter and finally β is a scalar parameter used to incorporate any extra prior knowledge of the distribution of $\hat{\boldsymbol{x}}$. This technique removes the requirement to explicitly calculate the Jacobian matrix. For complex functions, computing the Jacobian matrix can be a difficult task either because it requires complicated function derivatives, if it is done analytically, or it is computationally expensive if done numerically. Therefore the UKF tends to be more efficient than the EKF.

4.2. Gaussian process unscented Kalman filter (GP-UKF)

By incorporating probabilistic non-parametric GP models into the EKF and UKF techniques, model uncertainty can explicitly be incorporated into the prediction and the filtering steps [17]. The main idea of introducing the UKF is to replace the linearization in the EKF for a more accurate method based on the unscented transform. Briefly, the unscented transform performs the approximation of the underlying function by extracting the sigma points X from the Gaussian estimate

$$X_k^i = (\mu_{k-1} \quad \mu_{k-1} + \lambda\sqrt{\Sigma_{k-1}} \quad \mu_{k-1} - \lambda\sqrt{\Sigma_{k-1}}) \quad (23)$$

Where the sigma points are located at the mean μ and symmetrically along the main axes of the covariance function (two per dimension). The $r = 2n_x + 1$ sigma points are then propagated forward using the prediction model GP_f

$$\widehat{X}_k^i = X_{k-1}^i + GP_f(\mu(X_{k-1}^i, D_f)) \quad \text{for } i = 0,\ldots,2n \quad (24)$$

Later on, a new set of sigma points is computed from the previous estimate, where

$$\breve{X}_k^i = (\widehat{\mu}_{k-1} \quad \widehat{\mu}_{k-1} + \lambda\sqrt{\widehat{\Sigma}_{k-1}} \quad \widehat{\mu}_{k-1} - \lambda\sqrt{\widehat{\Sigma}_{k-1}}) \quad (25)$$

Then, the observation model Gp_g is used to predict an observation for each of these sigma points

$$\breve{Z}_k^i = GP_g\left(\breve{X}_{k-1}^i, D_g\right) \quad \text{for } i = 0,\ldots,2n \quad (26)$$

This resulting set of sigma points now captures the overall uncertainty following the state observation step. A complete description of the GP-UKF approach can be found in [17].

4.3. Gaussian process assumed density filter (GP-ADF)

In contrast to finite-sample approximations (UKF, GP-UKF) of the prior and the predictive distribution, GP-ADF propagates full densities by using the properties of GPs [29]. The resulting predictive distribution is described by a Gaussian distribution with exact mean and covariance, and is computed analytically [25]. In contrast to EKF and UKF, the GP-ADF explicitly takes into account nonlinearities by using GP models (linearization is not required) while propagating full Gaussian densities (sigma points are not used).

By using the result of the preceding filter step $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ as a prior distribution on $\mathbf{x}_{k-1}$, a new prediction through the transition model GP_f is obtained. Therefore, the values μ_k^x and Σ_k^x for the predictive distribution are obtained and thus the conditional distribution for the system state upon an observation $p(\boldsymbol{x}_k|\boldsymbol{z}_{1:k-1}) \propto N\left(\mu_k^x, C_k^x\right)$ is approximated, where the upper-script x indicate a one-step ahead prediction in the latent space. For the filter update at time k, the preceding prediction result $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ serves as the prior on $\mathbf{x}_k$, and it is combined with the recent observation $\mathbf{z}_k$ to determine the filter update step using Eq. (12). To accomplish this, the joint distribution is required

$$p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{z}_{1:k-1}) = p(\mathbf{z}_k|\mathbf{x}_k)\, p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) \quad (27)$$

The measurement model Gp_g yields the Gaussian likelihood $p(\mathbf{z}_k|\mathbf{x}_k)$, which when combined with the Gaussian prior $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$ allow us obtaining the Gaussian predictive distribution $p(\boldsymbol{z}_k|\boldsymbol{z}_{1:k-1}) \propto N\left(\mu_k^z, \Sigma_k^z\right)$, where the upper-script z indicate a one-step ahead prediction in the measurement space. To approximate the joint distribution with a Gaussian, we compute the joint covariance

$$\Sigma = \begin{bmatrix} \Sigma_k^x & \Sigma_{xy} \\ \Sigma_{xy}^T & \Sigma_k^z \end{bmatrix} \quad (28)$$

Finally, the resulting joint Gaussian distribution in Eq. (27) leads to the current filter update as $p(\boldsymbol{x}_k|\boldsymbol{z}_{1:k}) = N\left(\boldsymbol{x}_k|\hat{\mu}_k, \hat{\Sigma}_k\right)$, where the mean and variance are given by

$$\begin{aligned} \hat{\mu}_k &= \mu_k^x + \Sigma_{xy}\left(\Sigma_k^z\right)^{-1}\left(\mathbf{z}_k - \mu_k^z\right) \\ \hat{\Sigma}_k &= \Sigma_k^x - \Sigma_{xy}\left(\Sigma_k^z\right)^{-1}\Sigma_{xy}^T \end{aligned} \quad (29)$$

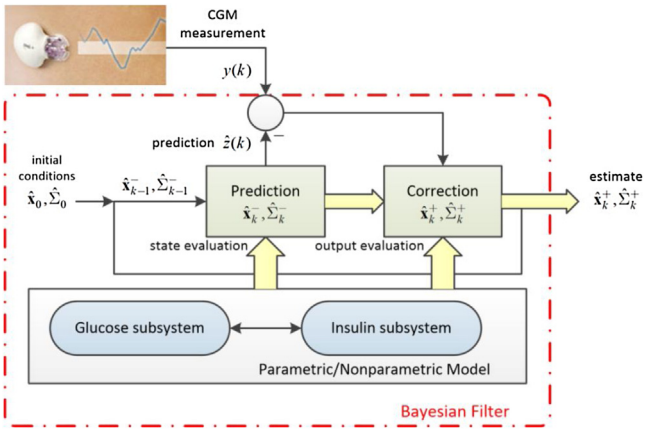

Fig. 4. Filtering process: for each time step the prediction given the transition model is corrected incorporating the current measurement.

Table 2
Carbohydrate intake protocol.

Carbohydrate content D_G [g]	47	16	63	31	63	31
Meal times [h]	3.00	5.00	7.30	11.00	14.30	17.00

5. Plasma insulin estimation

The simulation setup for the filtering process, working in parallel with the glycemic stochastic model, is presented in Fig. 4. Thus, at each time step the state prediction (including plasma insulin concentration) is first obtained and later on this estimate is corrected by incorporating the most recent CGM measurement.

5.1. Simulation modeling setup

In order to facilitate analysis of the results obtained, all simulations are based on an insulin injection protocol of only one subcutaneous injection, of a bolus of $u = 3$ U of insulin, at 8 a.m. Moreover, it is assumed that the basal insulin infusion rate is equal to 1 [pmol/kg/min] as only short-acting insulin is administrated, i.e. a small time-delay between the insulin infusion and the change in the levels of plasma insulin is ensured. A protocol of multiple carbohydrate intakes, given in Table 2, is adopted.

The algorithm presented in Fig. 5 outlines the simulation setup for the proposed filtering processes. The UKF and GP-UKF specific parameters were set to default values $\alpha = 1$, $\beta = 2$ and $\kappa = 0$. In line 1, the latent state is initialized given the prior distribution $p(\mathbf{x}_0) = N(\mu_0, \Sigma_0)$. The prior variance was set to $\Sigma_0 = 0.5^2$, whereas both the system noise and the measurement noise were set to $\Sigma_w = 0.1^2$ and $\Sigma_v = 0.1^2$. With this initialization, the uncertainty is initially fairly high, but the system and measurement noises are fairly small considering the amplitudes of the Ito's parameter for the glycemic transition function in Eq. (7) plus the CGM error for the measurement function in Eq. (5).

In total, 500 points were drawn by interacting with the stochastic Hovorka model to generate the training set for the transition model GP_f. For the measurement model Gp_g the targets of the GP_f were used as training inputs and when mapping them through CGM the corresponding training data were obtained. Hyperparameter optimization for all GP models fitted is done off-line using ground truth of the hidden states, i.e. BG states without the influence of variability and sensor errors. Additionally, for GP training the squared-exponential with automatic relevance determination (SE-ARD) covariance function plus a noise term was used. In line 10, the procedure is iterated for 240 time steps and each filter performance is evaluated. Considering that each CGM data is available every 6 [min], the glycemic profiles for a period of 24 [h] are readily obtained. The performances of the EKF, the UKF, the GP-UKF and the GP-ADF filters for estimating the plasma insulin state are based on averages over 100 test runs.

Table 3
Summary of errors for insulin estimation with different input states (100 runs, in μU/ml).

State	MAHA	RMSE	NLL
$\mathbf{x}(t) = [G(t)]$	24.61 ± 5.5	28.67 ± 8.6	4.18 ± 0.9
$\mathbf{x}(t) = [G(t), I(t-1)]$	13.60 ± 2.6	16.11 ± 2.5	-0.71 ± 0.5
$\mathbf{x}(t) = [G(t), U_G(t)]$	12.50 ± 2.4	15.60 ± 2.0	-0.74 ± 0.5

5.2. State space specification for GP models

The GP transition and observation models are learned using set of data points which are obtained by interacting with the stochastic Hovorka model. Nevertheless, when using GPs a word of caution is in order: training points that properly map the input and output space and at the same time ensures a broad enough coverage of the state space of the system operation regimen must be selected. To this aim, the Ito's noise parameter is set to $\sigma = 0.10$ for obtaining the training data sets. This allowed us to adequately characterize acceptable glycemic variability and, at the same time, be able to detect excessive BG fluctuations by evaluating any increase in the GP predicted covariances.

To achieve proper insulin estimations, an accurate mapping between the input and output space should be guaranteed. In Fig. 6, results obtained for predicting the latent state $I(t)$ using GP regression with different settings for the state vector in the input space are shown. In Table 3 the corresponding prediction errors are given. In Fig. 6a), only the current BG value $G(t)$ is used as state variable in order to predict the change in the plasma insulin level $I(t)$. It seems that this measurement is not informative enough to properly estimate the predictive distribution for the insulin state and in consequence leads to large errors in both the mean and variance. Also, in panel b) of Fig. 6, the input state vector is defined by the current BG value $G(t)$ and the current carbohydrate absorption rate $U_G(t)$ (refer to Eq. (1)). Based on the input variables, it was found that hidden state estimation provides more accurate predictions of the plasma insulin level $I(t)$. This situation can be rapidly inferred comparing Fig. 6a against b.

The choice of such a more informative input is quite obvious since glycemic fluctuations mostly depend on the amount of carbohydrates consumed. Notwithstanding, in real-life conditions this implies to keep track of every meal consumed and may result in a burden for the diabetic patient in control and monitoring implementations for an AP. Moreover, there always exists some risk that the meal will not be eventually consumed [30]. Finally, in the panel c) of Fig. 6, results obtained when formulating the state variable $\mathbf{x}(t)$ as the current BG value and the previous plasma insulin prediction, i.e. $\mathbf{x}(t) = [G(t), \quad I(t-1)]$. This is tricky, since errors in the estimated state distribution might be propagated forward and led to increased filtering inaccuracy. However, given the case of satisfactory recursive estimations, this choice would alleviate the burden of being constantly aware of carbohydrates intake, which is highly sought-after for a real implementation of an AP.

5.3. Numerical experiments

Numerical simulations were carried out so as to evaluate the performance of the implemented filtering techniques. Thus, the error between the true BG states and the filtered distribution were computed using the root mean square error (RMSE), the Mahalanobis distance (MAHA) and the negative log-likelihood (NLL)

Algorithm - Simulation setup for insulin estimation

Input. $\Sigma_w, \Sigma_v, \alpha, \beta, \kappa, D_G, u, \sigma$

initialize $\mathbf{x}_0(t) \sim \mathcal{N}\left(\mu_0, \Sigma_0\right)$
for $j = 1$ **to** 500 **do** ▷ generate training sets
  sample $\mathbf{x}_j(t) = \mathrm{a}\left(\mathbf{x}_{j-1}(t), u(t)\right) + \sigma d\omega$ with Eq.(9) using $\sigma = 0.1$ ▷ training inputs $\mathcal{GP}_f$
  observe $f\left(\mathbf{x}_j(t)\right) + w_j$ with Eq.(11) ▷ training targets $\mathcal{GP}_f$, trainig inputs $\mathcal{GP}_g$
  observe $g\left(f\left(\mathbf{x}_j(t)\right)\right) + v_j$ with Eq.(12) ▷ training targets $\mathcal{GP}_g$
end for
train $\mathcal{GP}_f$ using $\left\langle \mathbf{x}(t), f\left(\mathbf{x}(t)\right) + w \right\rangle_{j=1}^{500}$
train $\mathcal{GP}_g$ using $\left\langle f\left(\mathbf{x}(t)\right) + w, g\left(f\left(\mathbf{x}(t)\right)\right) + v \right\rangle_{j=1}^{500}$
for $k = 1$ **to** 100 **do** ▷ filter evaluations
  for $i = 1$ **to** 240 **do** ▷ filter experiments
    sample latent state $\mathbf{x}_i(t) = \mathrm{a}\left(\mathbf{x}_{i-1}(t), u(t)\right) + \sigma d\omega$ with Eq.(9) ▷ glucose-insulin model
    measure $\mathbf{z}_i = g\left(\mathbf{x}_i\right) + v_i$ with Eq.(12) ▷ succesor state
    compute $p\left(\mathbf{x}_i \,|\mathbf{z}_{1:i}\right)$ with Eq.(15) ▷ filter distribution
  end for
  for all filters **do**
    compute performance metrics
  end for
end for

Fig. 5. Experimental setup for the filtering process.

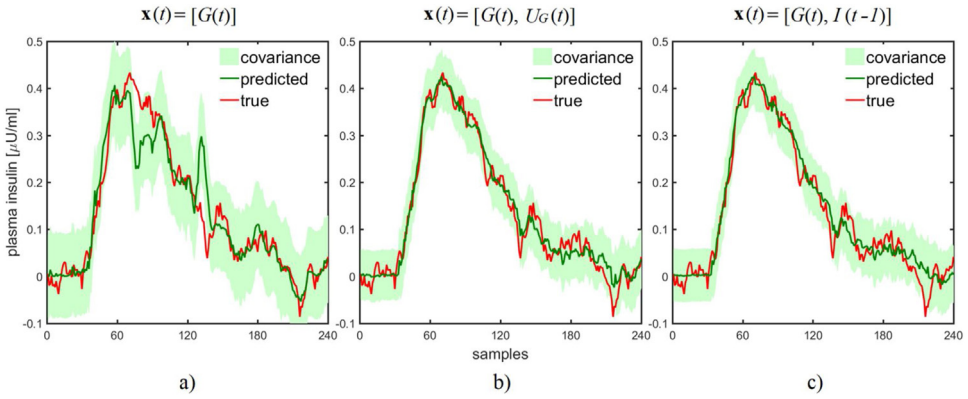

Fig. 6. Prediction errors for the hidden state $I(t)$ using different variables as the input states.

metric. The aim of including MAHA and NLL is to account for prediction uncertainty by measuring the coherence of the filtering distributions whilst penalizing the volume of the posterior covariance matrix. For all measures, lower values indicate better performance. All the simulations were executed with GNU Octave (version 4.2.1) on an Intel(R) Core i7-7700HQ 4-Core 2.8 GHz processor.

During the simulation experiments, at time t, the current input state used for the parametric filtering techniques is the one given in Eq. (3), whereas the input state for the nonparametric filters is given as $\mathrm{x}(t) = [G(\mathrm{t}), I(t-1)]$, where $G(t)$ is the current glucose level and $I(t-1)$ is the plasma insulin concentration at the previous time step. Eventually, as the only measurement available is the one provided by the CGM, the filter distribution of the current plasma insulin level $I(t)$ needs to be estimated.

Table 4
Summary of the errors for plasma insulin estimation (100 runs).

Variability: low ($\sigma = 0.10$)	MAHA	RMSE	NLL
EKF	23.52 ± 1.8	17.40 ± 0.3	1.68 ± 0.1
UKF	21.48 ± 1.7	12.48 ± 0.3	1.55 ± 0.1
GP-UKF	17.38 ± 0.6	8.49 ± 0.1	-0.10 ± 0.0
GP-ADF	10.48 ± 0.4	5.50 ± 0.1	-0.29 ± 0.0

To simulate an increase in the glycemic variability, the noise scale parameter was progressively augmented from $\sigma = 0.10$ to 0.25 and 0.50, while performance measures are given in Tables 4–6 respectively. Furthermore, a comparison between the run-time per iteration of the presented filtering approaches is provided in Table 7. From Figs. 7–9, results obtained when estimating plasma insulin profiles for different levels of glycemic variability are plot-

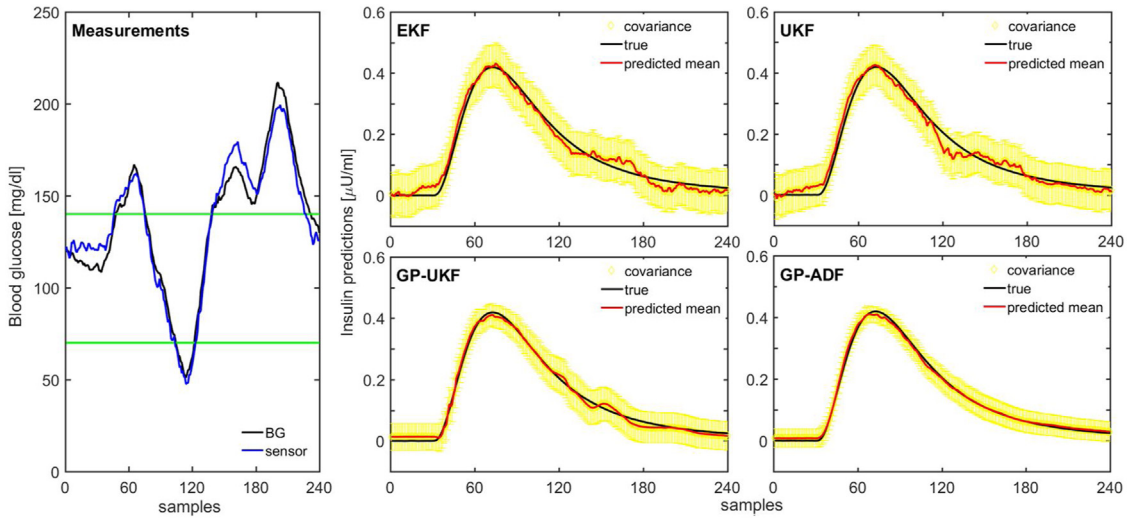

Fig. 7. CGM measurements and insulin estimation for low glycemic variability ($\sigma = 0.10$).

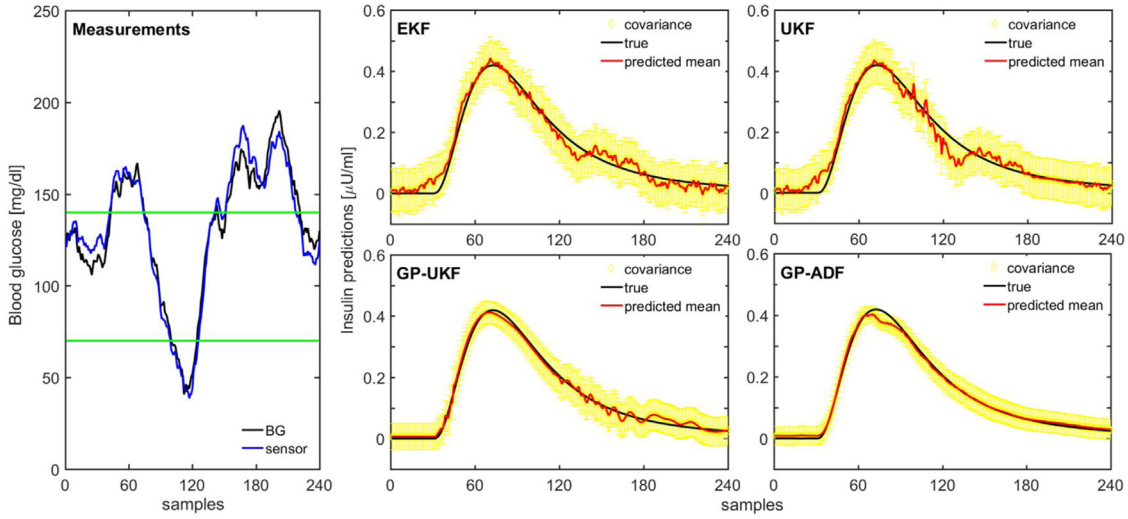

Fig. 8. CGM measurements and insulin estimation for medium glycemic variability ($\sigma = 0.25$).

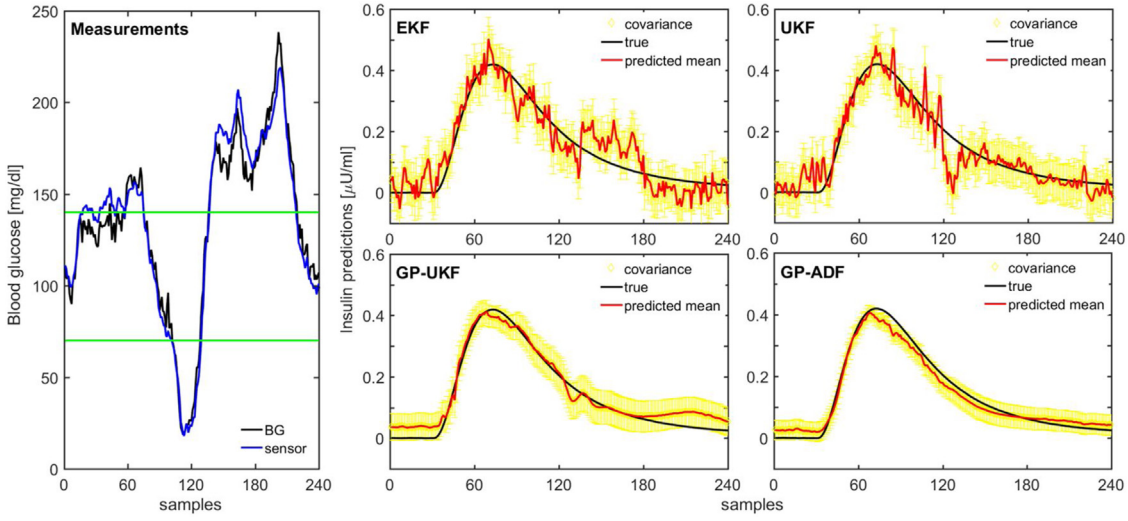

Fig. 9. CGM measurements and insulin estimation for high glycemic variability ($\sigma = 0.50$).

Table 5
Summary of the errors for plasma insulin estimation (100 runs).

Variability: medium ($\sigma = 0.25$)	MAHA	RMSE	NLL
EKF	27.37 ± 2.9	18.29 ± 0.6	1.82 ± 0.2
UKF	25.81 ± 2.0	13.06 ± 0.5	1.76 ± 0.2
GP-UKF	15.26 ± 1.8	7.70 ± 0.5	0.39 ± 0.1
GP-ADF	12.73 ± 2.6	7.11 ± 0.5	-0.12 ± 0.0

Table 6
Summary of the errors for plasma insulin estimation (100 runs).

Variability: high ($\sigma = 0.50$)	MAHA	RMSE	NLL
EKF	61.62 ± 5.5	25.65 ± 0.9	2.15 ± 0.1
UKF	34.60 ± 6.0	20.07 ± 1.3	1.93 ± 0.1
GP-UKF	22.47 ± 5.7	10.30 ± 0.9	0.44 ± 0.0
GP-ADF	15.94 ± 4.1	10.37 ± 0.9	0.27 ± 0.0

Table 7
Overall time consumption for insulin estimation (1 run, in seconds).

Variability	Model	EKF	UKF	GP-UKF	GP-ADF
Low ($\sigma = 0.1$)	0.479	0.214	0.321	4.772	2.442
Medium ($\sigma = 0.25$)	0.573	0.212	0.318	4.886	2.459
High ($\sigma = 0.5$)	0.701	0.219	0.319	4.850	2.470

ted. Note that in the left panels, plasmatic BG levels and the CGM measurements are shown. In the remaining panels, the estimated state covariance Σ_k is plotted (yellow lines) for each filter implementation in order to illustrate the reliability of the resulting estimates. Notice that measurements of true insulin values (black lines) are hidden to the estimator and they are only considered as benchmarks to highlight the difference with the estimated insulin concentration (red line).

For small values of the Ito noise parameter, i.e. low glycemic variability, insulin state estimation can be attained with acceptable accuracy regardless of the filtering approach employed. However, filters based on GP models generally exhibit consistently better performance than their parametric counterparts. It can also be observed that the performance of GP filters is particularly good even for estimating plasma insulin under large increments of glycemic variability. Moreover, because of the appropriate treatment of such uncertainties, the predictions of GP-based filters rarely are inconsistent and systematic error-prone. On the other hand, parametric filters rapidly degrades they performance in presence of large variability, as a consequence of poor treatment of system nonlinearities and model prediction errors.

More specifically, the GP-ADF outperforms GP-UKF by a small margin. In turn, it can be seen that the UKF technique performs slightly better than the EKF when comparing the error metrics reported in Tables 4–6, which can be justified by its better handling of nonlinearities in the glucose dynamics. The EKF requires intricate mathematical computations in order to linearize the system distribution, which is considered its main shortcoming especially when the system is definitively nonlinear. Finally, UKF based techniques also degrade when the functions, which are used for mapping the sigma points, are highly nonlinear with high variances.

In terms of time consumptions (see Table 7 for details) in our implementations, the UKF cost is slightly larger than the corresponding EKF's. Notice that even though the UKF algorithm requires more function evaluations in each time step due to sigma-points propagation, the EKF requires costly matrix inversion operations in the linearization process. Filters based on GP models are more than ten times slower than the parametric ones since prediction uncertainty needs to be calculated for each point per step. But considering that BG measurements are obtained every five to six minutes, neither of the filtering approaches might experience any problem to perform in real-time.

6. Concluding remarks

New technologies such as the AP are promising for type 1 diabetes treatment. However, most autonomous control and monitoring strategies aimed at managing glycemic values require of real-time estimations of plasma glucose and plasma insulin concentrations so as to increase the efficacy and safety of control strategies. Unfortunately, unlike BG levels which can be determined in real-time using CGM, plasma insulin determination needs of filtering techniques so as to estimate its current concentrations.

A key factor for the filtering process accuracy is the availability of an appropriate glucose-insulin model. Filtering techniques based on parametric models are solely applicable when the models are fully known or readily available. If only samples of the underlying function are available, GP models can be employed to advantage. Particularly, Bayesian filtering techniques with GP prediction and observation models allow for robust estimations as they coherently represent uncertainties about the system and measurement functions at locations that have not been encountered in the data collection phase.

It is imperative to study how variability and uncertainty in the determination of BG levels may affect the quality of the resulting insulin concentration estimations. To achieve realistic simulations of metabolic conditions, a stochastic version of the Hovorka glucose-insulin model was presented. Later on, this model was incorporated into nonlinear parametric and nonparametric filtering techniques so as to estimate plasma insulin concentration under glycemic variability. Since uncertainty can explicitly be incorporated into the prediction and the filtering steps through GP regression, it was found that GP filters performed significantly better than their parametric counterparts. In this work we used an in-silico model based on the well-known Hovorka glucose-insulin interaction model, thus the next step must consider the evaluation of the presented techniques on a real human dataset.

Competing interests

None.

References

[1] D.M. Maahs, B.A. Buckingham, J.R. Castle, A. Cinar, E.R. Damiano, E. Dassau, J.H. DeVries, F.J. Doyle, S.C. Griffen, A. Haidar, Outcome measures for artificial pancreas clinical trials: a consensus report, Diabetes Care 39 (2016) 1175–1179.
[2] M. De Paula, L.O. Ávila, E.C. Martínez, Controlling blood glucose variability under uncertainty using reinforcement learning and Gaussian processes, Appl. Soft Comput. 35 (2015) 310–332.
[3] C.S. Hughes, S.D. Patek, M. Breton, B.P. Kovatchev, Anticipating the next meal using meal behavioral profiles: a hybrid model-based stochastic predictive control algorithm for T1DM, Comput. Methods Programs Biomed. 102 (2011) 138–148.
[4] L. Magni, D.M. Raimondo, L. Bossi, C. Dalla Man, G. De Nicolao, B. Kovatchev, C. Cobelli, Model predictive control of type 1 diabetes: an in silico trial, J. Diabetes Sci. Technol. 1 (2007) 804–812.
[5] L. Avila, E. Martínez, An active inference approach to on-line agent monitoring in safety–critical systems, Adv. Eng. Inf. 29 (2015) 1083–1095.
[6] L. Ávila, E. Martínez, A grid-based tool for optimal performance monitoring of a glycemic regulator, Optim. Control Appl. Methods 37 (2016) 1103–1121.
[7] G. Sparacino, A. Facchinetti, C. Cobelli, Smart continuous glucose monitoring sensors: on-line signal processing issues, Sensors 10 (2010) 6751–6772.
[8] C. Neatpisarnvanit, J.R. Boston, Estimation of plasma insulin from plasma glucose, IEEE Trans. Biomed. Eng. 49 (2002) 1253–1259.
[9] D. de Pereda, S. Romero-Vivo, B. Ricarte, P. Rossetti, F.J. Ampudia-Blasco, J. Bondia, Real-time estimation of plasma insulin concentration from continuous glucose monitor measurements, Comput. Methods Biomech. Biomed. Eng. 19 (2016) 934–942.
[10] C. Eberle, C. Ament, The unscented Kalman filter estimates the plasma insulin from glucose measurement, Biosystems 103 (2011) 67–72.

[11] A.C. Charalampidis, K. Pontikis, G.D. Mitsis, G. Dimitriadis, V. Lampadiari, V.Z. Marmarelis, A. Armaganidis, G.P. Papavassilopoulos, Calibration of a microdialysis sensor and recursive glucose level estimation in ICU patients using Kalman and particle filtering, Biomed. Signal Process. Control 27 (2016) 155–163.
[12] S. Aslan, Comparison of the hemodynamic filtering methods and particle filter with extended Kalman filter approximated proposal function as an efficient hemodynamic state estimation method, Biomed. Signal Process. Control 25 (2016) 99–107.
[13] Z. Mahmoudi, K. Nørgaard, N.K. Poulsen, H. Madsen, J.B. Jørgensen, Fault and meal detection by redundant continuous glucose monitors and the unscented Kalman filter, Biomed. Signal Process. Control 38 (2017) 86–99.
[14] F. Cameron, G. Niemeyer, B.W. Bequette, Extended multiple model prediction with application to blood glucose regulation, J. Process Control 22 (2012) 1422–1432.
[15] F.L. Lewis, Optimal Estimation: with an Introduction to Stochastic Control Theory, Wiley, New York, 1986.
[16] E. Kowsari, B. Safarinejadian, Applying GP-EKF and GP-SCKF for non-linear state estimation and fault detection in a continuous stirred-tank reactor system, Trans. Inst. Meas. Control (2016).
[17] J. Ko, D. Fox, GP-BayesFilters. Bayesian filtering using Gaussian process prediction and observation models, 2008 IEEE/RSJ Int. Conf. Intell. Robot. Syst., IEEE (2008) 3471–3476.
[18] R. Hovorka, V. Canonico, L.J. Chassin, U. Haueter, M. Massi-Benedetti, M.O. Federici, T.R. Pieber, H.C. Schaller, L. Schaupp, T. Vering, Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes, Physiol. Meas. 25 (2004) 905–920.
[19] R. Borg, J.C. Kuenen, B. Carstensen, H. Zheng, D.M. Nathan, R.J. Heine, J. Nerup, K. Borch-Johnsen, D.R. Witte, A.S. Group, Real-life glycaemic profiles in non-diabetic individuals with low fasting glucose and normal HbA1c: the A1C-Derived Average Glucose (ADAG) study, Diabetologia 53 (2010) 1608–1611.
[20] C.R. Marling, J.H. Shubrook, S.J. Vernier, M.T. Wiley, F.L. Schwartz, Characterizing blood glucose variability using new metrics with continuous glucose monitoring data, J. Diabetes Sci. Technol. 5 (2011) 871–878.
[21] S.E. Siegelaar, F. Holleman, J.B.L. Hoekstra, J.H. DeVries, Glucose variability; does it matter? Endocr. Rev. 31 (2010) 171–182.
[22] K. Ito, Stochastic differentials, Appl. Math. Optim. 1 (1975) 374–381.
[23] A. Doucet, S. Godsill, C. Andrieu, On sequential Monte Carlo sampling methods for Bayesian filtering, Stat. Comput. (2000).
[24] C. Williams, C. Rasmussen, Gaussian processes for machine learning, MIT Press, 2006.
[25] J. Candela, A. Girard, J. Larsen, Propagation of uncertainty in Bayesian kernel models-application to multiple-step ahead forecasting, In Acoustics, Speech, and Signal Processing, Proceedings (ICASSP'03), IEEE International Conference on, vol. 2, II-701 (2003).
[26] G. Bishop, G. Welch, An introduction to the kalman filter, Proc SIGGRAPH, Course (2001).
[27] S.J. Julier, J.K. Uhlmann, New extension of the Kalman filter to nonlinear systems, in: AeroSense'97, International Society for Optics and Photonics, 1997, pp. 182–193.
[28] E. Wan, R. Van Der Merwe, The unscented Kalman filter for nonlinear estimation, in: As-spcc. Ieee 2000, 2000.
[29] P. Maybeck, Stochastic Models, Estimation, and Control, Academic press, 1982.
[30] B.W. Bequette, Challenges and recent progress in the development of a closed-loop artificial pancreas, Annu. Rev. Control. 36 (2012) 255–266.