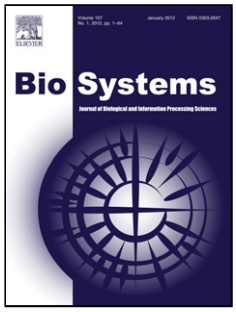


Title: Estimation of plasma insulin concentration under glycemic variability using nonlinear filtering techniques

Authors: Luis Omar Avila, Mariano De Paula, Carlos Roberto Sanchez-Reinoso

PII: S0303-2647(17)30236-8
DOI: https://doi.org/10.1016/j.biosystems.2018.06.003
Reference: BIO 3856

To appear in: *BioSystems*

Received date: 12-6-2017
Revised date: 26-3-2018
Accepted date: 18-6-2018

Please cite this article as: Avila LO, De Paula M, Sanchez-Reinoso CR, Estimation of plasma insulin concentration under glycemic variability using nonlinear filtering techniques, *BioSystems* (2018), https://doi.org/10.1016/j.biosystems.2018.06.003

# Estimation of plasma insulin concentration under glycemic variability using nonlinear filtering techniques

Luis Omar Avila[1], Mariano De Paula[2] and Carlos Roberto Sanchez-Reinoso[3]

[1]Corresponding author: Luis Omar Avila
Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC) - Laboratorio de Mecatrónica (LABME), CONICET-UNSL, Av. Ejército de los Andes 950, D5700BPB San Luis, Argentina.
Phone: +54-0266-4420823
Email: loavila@unsl.edu.ar
Orcid ID: 0000-0003-0321-068X

[2] Mariano De Paula
INTELYMEC group, Centro de Investigaciones en Física e Ingeniería del Centro CIFICEN – UNICEN – CICpBA – CONICET, Av. Del Valle 5537, B7400JWI Olavarría, Argentina.
Phone: +54-02284-451055.
Email: mariano.depaula@fio.unicen.edu.ar
Orcid ID: 0000-0001-7582-9188

[3]Carlos Roberto Sanchez-Reinoso
Centro De Investigación y Desarrollo en Modelado, Simulación y Optimización de Sistemas Electrónicos (CIDMOS), Universidad Nacional de Catamarca, Av. Maximio Victoria 55, K4700 Catamarca, Argentina.
Phone: +54-0383-4427807
Email: csanchezreinoso@santafe-conicet.gov.ar

**Abstract**

The ultimate goal of an artificial pancreas is finding the optimal insulin rates that can effectively reduce high blood glucose (BG) levels in type 1 diabetic patients. To achieve this, most closed-loop control strategies need to compute the optimal insulin action on the basis of precedent glucose and insulin levels. Unlike glucose levels which can be measured in real-time, unavailability of insulin sensors makes it essential the use of mathematical models to estimate plasma insulin concentrations. Between others, filtering techniques based on a generalization of the Kalman filter

(KF) have been the most widely applied in the estimation of hidden states in nonlinear dynamic systems. Nevertheless, poor predictability of BG levels is a key issue since the glucose-insulin dynamics presents great inter- and intra-patient variability. Here, the question arises as to whether glycemic variability is not properly taken into account in models formulations and whether or it would compromise proper estimation of plasma insulin concentration. In order to tackle this point, a deterministic model describing glucose-insulin interaction plus a stochastic process to account for BG fluctuations were incorporated into the extended (EKF), cubature (CKF) and unscented (UKF) configurations of the Kalman filter to provide an estimate of the plasma insulin concentration. We found that for low glycemic variability, insulin state estimation can be attained with acceptable accuracy; however, as glycemic variability rises, Kalman filters rapidly degrade their performance as a consequence of large nonlinearities.

***Keywords-*** insulin estimation, Kalman filter, glycemic variability, stochastic model, artificial pancreas.

## 1. Introduction

Type 1 diabetes is caused by selective destruction of the pancreas β-cells, which results in deficient insulin production for glucose regulation. In order to restore safe glucose levels in diabetic patients exogenous insulin delivery is required. Due to the increased scientific and industrial effort focused on the automatic glycemic regulation, the commercial development of an artificial pancreas (AP) continues to progress rapidly and it hopefully will soon take part of clinical care (Maahs et al., 2016). The ultimate goal of an AP is finding the optimal insulin rates that can effectively maintain normal BG levels, and infusing it in such way that glycemic profiles can mimic the body's natural regulatory mechanism (Ali and Padhi, 2011). To achieve this, most closed-loop control strategies compute the optimal insulin bolus to be administrated on the basis of plasma glucose and plasma insulin concentration (De Paula et al., 2015; Ghosh and Maka, 2014; Hughes et al., 2011). This can partly be achieved by the availability of continuous glucose monitors that measure glucose levels in the interstitial fluid and considering the dynamics between this evaluation and plasmatic concentrations (Sparacino et al., 2010). However, the development of more intelligent, efficient and robust control strategies requires also the quantification of plasma insulin levels.

However, unlike glucose levels which can be determined in real-time using commercial devices, the lack of insulin sensors makes it mandatory the use of mathematical or experimental models to estimate the current plasmatic insulin concentrations (Zisser et al., 2008). To perform plasma estimation, some approaches -see for example Neatpisarnvanit and Boston (2002)- make use of a mathematical model describing a glucose-insulin correlation to determine plasmatic insulin levels in an open-loop strategy. The main limitation is that current sensor measurements are not used to adjust the states or the parameters of the model, even if these adjustments could be necessary especially in processes with high variability (Patek et al., 2012). More recently, plasma insulin concentration has been estimated in real-time from BG data using filtering techniques (de Pereda et al., 2016; Eberle and Ament, 2011). These techniques are based on a generalization of the Kalman filter to nonlinear systems and have been widely used in the study of different dynamic systems (Nguyen et al., 2017; Sepasi et al., 2014; Turksoy et al., 2017). Briefly, filtering techniques are based on the conjunction of a partially observed model of the dynamic system and a state observer. In such situation, the observer first predicts the states variables and once new data is acquired the model state estimations are updated using a weighted process (Bishop and Welch, 2001).

Nevertheless, poor predictability of BG dynamics is a key issue that any filtering technique will have deal with. This is because the glucose-insulin dynamics shows great intra and inter-variability among different times and patients according to the carbohydrate content of meals, exercise level, age and stress (Krüger et al., 2003). This circumstance leads us to consider whether excessive variability might affect the proper estimation of plasma insulin concentration and, in consequence, compromise safety due to an inappropriate insulin control action. Even when a number of works have addressed the issue of plasma insulin and plasma glucose estimation taking into account uncertainty in the model parameters (Makroglou et al., 2006; Steil et al., 2005; Wang et al., 2014), the feasibility of insulin estimation under excessive but factual BG variability has not been sufficiently studied. De Pereda et al. (2016) studied the performance of a linearized version of the Kalman filter in which different selected time-variant model parameters have been considered as extended states. Møller (2008) has previously used the extended version of the Kalman filter in order to perform the modelling of the glucose-insulin dynamics using stochastic differential equations. Basically, this filter creates the possibility of working with nonlinear systems. However, when the state transition and observation models are highly nonlinear, the filter can give particularly poor performance.

In this work, an Ito´s diffusion process plus the Hovorka's glucose-insulin model is proposed and incorporated into the extended (EKF), cubature (CKF) and unscented (UKF) configurations of the

Kalman filter techniques to provide an estimate of the plasma insulin concentration. Introducing a stochastic model allows us to evaluate the effect of patient variability in diabetes control, while considering other sources of uncertainty in the estimation of glycemic values -such as sensor errors for example. This also allows us performing more realistic simulations of abnormal metabolic conditions while capturing a heterogeneous cohort of in-silico individuals. In such a manner, we can assess the performance of the different filtering techniques in a wide range of patient inter and intra variability situations. Then, a comparative study between these techniques is accomplished. Eventually, quantifying the error between a given dynamic model and the real and complex physiological system is useful to improve AP control and monitoring strategies.

The paper has been organized as follow. In Section 2 the widely accepted Hovorka's glucose–insulin model is presented and then a stochastic formulation, based this original Hovorka model, using the Ito diffusion process is presented. In Section 3 the observability of the stochastic model formulation is proven, which is mandatory for a successful estimation of plasma insulin concentration under glycemic variability. Then, different Kalman filter formulations for insulin estimation under glycemic variability are implemented. In Section 4, *in silico* different implementations of the glucose-insulin stochastic process are evaluated for different degree of glycemic variability in simulated patients and the obtained result are shown and discussed. Finally, in Section 5 the main contributions of the paper and future works are discussed.

## 2. Stochastic model of the glucose-insulin dynamics

In this section, we briefly present the reference deterministic model of the glucose-insulin interactions based on the work of Hovorka et al. (2004). Later, we present a stochastic diffusion process to model glycemic variability in synthetic diabetic patients.

### 2.1. Glucose-insulin system

The Hovorka glucose-insulin model is a nonlinear compartment numerical model with two inputs (insulin and glucose intake) and one output (glycemia). The model consists of three subsystems, namely the insulin absorption, the insulin action and the insulin-glucose dynamics subsystem. The carbohydrate absorption rate $U_G(t)$ [mmol/min] is described, as given in de Pereda et al. (2016), by the following equation

$$U_G(t) = \frac{D_G A_G t e^{-t/t_G}}{t_G^2} \quad (1)$$

where $D_G$ [mmol] is the carbohydrate content of the meal ingested (considering a molar mass of 180 [g/mol] for glucose molecule), $A_G$ is the carbohydrate bioavailability and $t_G$ [min] is the time-of-maximum appearance of glucose in the accessible compartment.

Three equations are introduced to describe the insulin absorption subsystem

$$\begin{aligned} \frac{dS_1(t)}{dt} &= u(t) - \frac{S_1(t)}{t_I} \\ \frac{dS_2(t)}{dt} &= \frac{S_1(t)}{t_I} - \frac{S_2(t)}{t_I} \\ \frac{dI(t)}{dt} &= \frac{S_2(t)}{t_I V_I} - k_I I(t) \end{aligned} \quad (2)$$

where $u(t)$ [mU/min] represents the insulin administration. Signals $S_1(t)$ and $S_2(t)$ are state variables describing absorption of subcutaneously administered insulin, $t_I$ is the time to maximum insulin absorption, $I(t)$ is the plasma insulin concentration subsystem output, $V_I$ is the insulin distribution volume and $k_I$ is the fractional elimination rate.

Another three equations are used to describe the insulin action on glucose kinetics

$$\begin{aligned} \frac{dx_1(t)}{dt} &= k_{b1} I(t) - k_{a1} x_1(t) \\ \frac{dx_2(t)}{dt} &= k_{b2} I(t) - k_{a2} x_2(t) \\ \frac{dx_3(t)}{dt} &= k_{b3} I(t) - k_{a3} x_3(t) \end{aligned} \quad (3)$$

where $x_1(t)$ [min$^{-1}$] is the rate of remote effect of insulin on glucose transport, while $x_2(t)$ [min$^{-1}$] and $x_3(t)$ [-] account for the elimination and endogenous glucose production respectively. Dynamics of these effects is given by the constants: $k_{a1}$, $k_{a2}$, $k_{a3}$ (deactivation rate constants), $k_{b1}$, $k_{b2}$, $k_{b3}$ (activation rate constants).

The last subsystem is nonlinear and it describes the insulin-glucose interaction dynamics

$$\frac{dQ_1(t)}{dt} = -\left[\frac{F_0}{V_G G(t)} + x_1(t)\right] Q_1(t) + k_{12} Q_2(t) - F_R + U_G(t) + EGP[1 - x_3(t)] \quad (4)$$

$$\frac{dQ_2(t)}{dt} = x_1(t) Q_1(t) - [k_{12} + x_2(t)] Q_2(t) \quad (5)$$

where $Q_1(t)$, $Q_2(t)$ represent the masses of glucose in the accessible and non-accessible compartments respectively, $k_{12}$ 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.

$$F_0 = \begin{cases} F_0 & G(t) \geq 4.5 \text{ mmol/l} \\ F_0 G(t)/4.5 & \text{otherwise} \end{cases} \quad (6)$$

In turn, $F_R$ represents renal glucose clearance above the glucose concentration threshold of 9 [mmol/l] and is given as

$$F_R = \begin{cases} 0.003(G(t) - 9)V_G & G(t) \geq 9 \text{ mmol/l} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Finally, BG levels are given as

$$G(t) = \frac{Q_1(t)}{V_G} \quad (8)$$

where $V_G$ is glucose distribution volume. Since glycemic sensors operate in the interstitial space, they have access to the interstitial glucose ($IG$) rather than to plasmatic levels. Glucose exchange across the capillary walls occurs by a simple but not instantaneous diffusion across a concentration gradient, such that

$$\frac{dIG(t)}{dt} = \frac{1}{\tau}(G(t) - IG(t)) \quad (9)$$

In Fig. 1, is given an implementation of the system of differential equations presented before in order to model the course of the blood glucose absorption rate, particularly when a person eats the meal protocol given in Table 2.

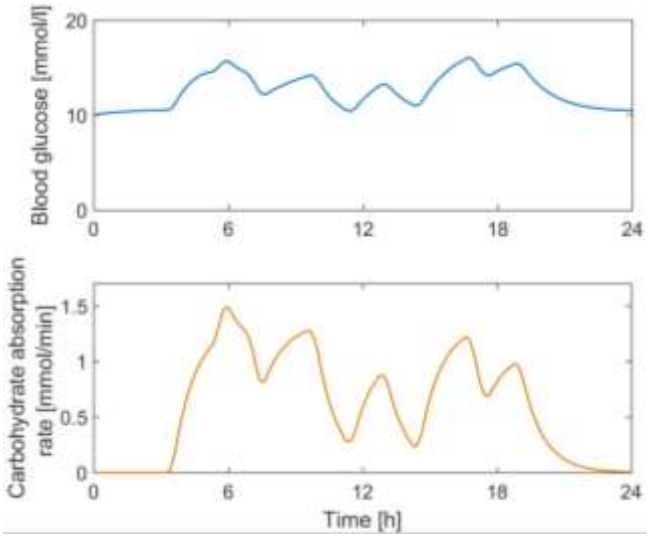


Figure 1: Blood glucose and carbohydrate absorption rate $U_G(t)$ for the carbohydrate intake protocol given in Table 2.

According to the glucose-insulin dynamics presented, the system state can be summarized such that the Hovorka model can be integrated into a stochastic system of differential equations

$$\boldsymbol{x}(t) = [S_1, S_2, I, x_1, x_2, x_3, Q_1, Q_2, IG] \tag{10}$$

Later, to define which output signals $\boldsymbol{z}(t)$ can be measured from the states, the general output equation is given as

$$\boldsymbol{z}(t) = h(\boldsymbol{x}(t)) \tag{11}$$

Hence the complete state space model formulation incorporates $\boldsymbol{z}(t)$ and, as only interstitial glucose $IG(t)$ levels can be measured, the corresponding output equation results to be linear

$$\boldsymbol{z}(t) = H\boldsymbol{x}(t) \quad \text{where } H = [0,0,0,0,0,0,0,0,1] \tag{12}$$

In Table 1 are given the value of the parameter for the Hovorka model, taken from (Hovorka et al., 2004), in order to represent glycemic patterns in type 1 diabetic patients.

Table 1: Set of model parameters.

| Parameter | Value | Unit |
|---|---|---|

| | | |
|---|---|---|
| BW | 75 | [Kg] |
| $G_b$ | 10 | [mmol/l] |
| $A_G$ | 0.8 | [-] |
| $t_G$ | 40 | [min] |
| $t_I$ | 55 | [min] |
| $k_I$ | 0.138 | [$min^{-1}$] |
| $V_I$ | 0.12*BW | [l] |
| $\tau$ | 15 | [min] |
| $k_{12}$ | 0.066 | [$min^{-1}$] |
| $V_G$ | 0.16*BW | [l] |
| $EGP$ | 0.0161 | [mmol/min] |
| $F_0$ | 0.8507 | [mmol/min] |
| $k_{a1}$ | 0.006 | [$min^{-1}$] |
| $k_{a2}$ | 0.06 | [$min^{-1}$] |
| $k_{a3}$ | 0.03 | [$min^{-1}$] |
| $k_{b1}$ | $k_{a1}$*51.2*10−4 | [$min^{-2}mU^{-1}l$] |
| $k_{b2}$ | $k_{a2}$*8.2*10−4 | [$min^{-2}mU^{-1}l$] |
| $k_{b3}$ | $k_{a3}$*520*10−4 | [$min^{-1}mU^{-1}l$] |

### 2.2. Glucose-insulin stochastic dynamic model

A number of works (see for example Borg et al. (2010), Marling et al. (2011), Siegelaar et al. (2010)) have described the considerable variability of BG levels that exist among diabetic patients. Thus, for the development of control and monitoring strategies for an artificial pancreas, the deterministic glucose-insulin dynamic model might be enhanced by taking into account the variable behavior of metabolism in diabetic patients. One way to regard such behavior is modeling temporal variability, in a certain way. In this manner, the sources of uncertainty that are temporally dependent can be taken into account, in the model formulation using stochastic processes.

A stochastic process is a variable that evolves with time in a probabilistic way and which has not temporal derivatives in the conventional sense, and in consequence, cannot be manipulated numerically using conventional methods. The Ito´s diffusion process is a particular stochastic

process which can be differentiated and integrated by means of the Ito's Lemma (Ito, 1975). Therefore, to model the BG variable behavior in synthetic diabetic subjects we employ an Ito´s diffusion process superimposed on the Hovorka´s deterministic model. This allows us performing more realistic simulations of metabolic conditions in diabetic patients while capturing a heterogeneous cohort of in-silico subjects that accounts sufficiently well for the observed inter and intra variability.

Ito (Ito, 1975) provided an alternative to ordinary numerical rules of calculus by defining a particular kind of uncertainty representation based on the Wiener diffusion process as a building block. Therefore, the glucose-insulin dynamic function can be described as a controlled Ito's diffusion process of the form

$$dq(t) = f(\boldsymbol{x}(t), \boldsymbol{u}(t))dt + \sigma d\omega \tag{13}$$

where $f(\boldsymbol{x}(t), \boldsymbol{u}(t))$ describes the glucose-insulin transition function, $d\omega$ is the increment of a *Wiener* type process and $\sigma$ denote Brownian noise and it is a scaling parameter, respectively. Lastly, in order to account for glycemic variability, the noise parameter is included such that it affects the mass of glucose in the accessible compartment as given by

$$dQ_1(t) = \left(-\left[\frac{F_0}{V_G G(t)} + x_1(t)\right] Q_1(t) + k_{12} Q_2(t) - F_R + R_a(t) + EGP[1 - x_3(t)]\right) dt + \sigma d\omega \tag{14}$$

## 3. Plasma insulin estimation

Firstly, we discuss the state observability of the presented stochastic dynamic model. Then, the Extended Kalman Filter (EKF), the Cubature Kalman Filter (CKF) and the Unscented Kalman Filter (UKF) formulations are presented for the plasma insulin estimation using the, previously presented, stochastic dynamic model.

### 3.1. State observability

To build an efficient estimator, first, it has to be proven the feasibility to uniquely determine the system state $\boldsymbol{x}(t)$ from the observed output $\boldsymbol{z}(t)$. This analysis is usually performed using the chain of Lie derivatives (Kocarev et al., 1998)

$$\Theta(\boldsymbol{x}) = \begin{bmatrix} L_f^0 h(\boldsymbol{x}) \\ L_f^1 h(\boldsymbol{x}) \\ \vdots \\ L_f^{r-1} h(\boldsymbol{x}) \end{bmatrix} \quad with \begin{cases} L_f^0 h(\boldsymbol{x}) = h(\boldsymbol{x}) \\ L_f^j h(\boldsymbol{x}) = \frac{\partial L_f^{j-1} h(x)}{\partial \boldsymbol{x}} \end{cases} \quad for j = 1, \dots, r-1 \tag{15}$$

where $r \leq n$ denotes the observation relative degree and $n$ is the system order. Afterwards, it has to be verified whether the Jacobian matrix (Eq. (16)) has at least the rank of the system order $n$ to ensure that the system is observable

$$J(\boldsymbol{x}) = \frac{\partial \Theta(x)}{\partial \boldsymbol{x}} \tag{16}$$

The observability of the Hovorka model with the state vector presented in Eq. (10) has already been proven using the Lie derivatives in the work of Szalay et al. (2012). The most important constraints that need to be taken into consideration when constructing a state estimator for the Hovorka model are given below

- State variables $G(t)$, $Q_1(t)$, and $Q_2(t)$ are observable
- State variables $I(t)$, $S_1(t)$, and $S_2(t)$ are observable if at least one of $x_1(t)$, $x_2(t)$, or $x_3(t)$ is observable
- $x_1(t)$ is observable if $x_1(t) \neq 0$, which always apply as long as there is injected insulin flow
- $x_2(t)$ is observable if $x_2(t) \neq 0$ and $I(t)$ is observable, which always apply as long as there is injected insulin flow
- $x_3(t)$ is observable if $x_3(t) \leq 1$

If those conditions are fulfilled, then the Jacobian matrix results to have full rank and hence all the states are observable, in other words, it is possible to estimate plasma insulin from glucose measurements.

### 3.2. Kalman state estimators

The most simple way to obtain an estimate for plasma insulin would be to open Hovorka's model, take out the submodel of the plasma insulin compartment, and drive it with the plasma glucose measurement $G(k)$. In such a case, plasma insulin is estimated using a submodel of the insulin compartment as an open loop estimator. Another manner is to obtain an estimate of the plasma insulin level using the idea presented in Simon (2006), this is, building a closed-loop estimator that uses a complete in-silico glucose-insulin model in parallel to the observed in-silico process, as presented in Fig. 2.

For computation reasons the described dynamic model (Section 2) must be integrated every sampling interval $k$ and the numerical Euler integration method with integration step size $\Delta T$ can be carried out with initial condition $\boldsymbol{x}_{k-1}$. In this way, the following KF update equations will have the index $k$ referring to a certain time instant $k\Delta T$. Thus, for each $k$-th time step the state prediction is obtained and then it is corrected incorporating the current measurement.

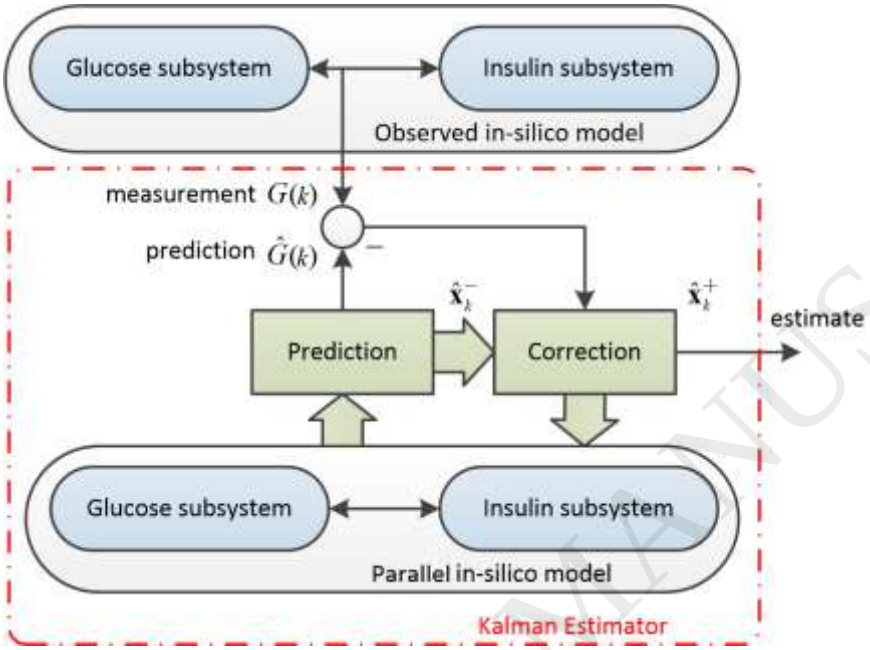


Figure 2: State estimation using a parallel model. For each time step the prediction given the parallel model is corrected incorporating the current measurement.

Processes with high variability, as the one studied in this work, can be represented by systems with noise in their transition and measurement functions. Commonly, in the standard discrete form these systems are expressed as

$$\begin{aligned} \boldsymbol{x}_k &= \mathrm{f}(\boldsymbol{x}_{k-1}) + w_{k-1} \quad & w_{k-1} &\sim N(0, Q_x) \\ \boldsymbol{z}_k &= \mathrm{h}(\boldsymbol{x}_k) + v_k \quad & v_k &\sim N(0, R_z) \end{aligned} \tag{17}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ is the system state and the function f represents the system dynamics in discrete form. In turn, h denotes the measurement function and $\boldsymbol{z}_k \in \mathbb{R}^m$ is the current measurement. Additionally, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ represent the process and the observation noises, while $Q_{\boldsymbol{x}}$ is the process noise covariance and $R_{\boldsymbol{z}}$ denote the measurement noise variance.

The estimation is carried out in equidistant time steps $t = k\Delta T$ with a typical time step size of $\Delta T = 6\ min$. At the beginning the filtering process is fed with the initial conditions of $\hat{\boldsymbol{x}}_0$ and the initial

covariance matrix of the estimation error $\hat{P}_0$. Subsequently, the estimates of the states and parameters in $\hat{\boldsymbol{x}}_k$ and the covariance matrix of the estimation error $\hat{P}_k$ have to be propagated through time.

The initial state may assume a specific value, but because this value is generally not known a priori, the initial state is considered to be a random vector with a Gaussian distribution. Therefore, we assume that the mean of the system state equals the initial state. For plasma glucose the basal value $G_b$ is chosen, see Table 1, while all other estimates are started with zero. Therefore, for the initial state we have that

$$\hat{\boldsymbol{x}}_0^+ = E\{\boldsymbol{x}(0)\} \tag{18}$$

In turn, the initial covariance matrix of the estimation error is initialized as diagonal matrix

$$\hat{P}_0^+ = E\{(\boldsymbol{x}(0) - \hat{\boldsymbol{x}}_0^+)(\boldsymbol{x}(0) - \hat{\boldsymbol{x}}_0^+)^T\} \tag{19}$$

where $E\{.\}$ is the expectation operator. For each time step from $t_{k-1}$ to $t_k$, the propagation of those two initial quantities is done in two steps. First the filter estimates the current state using a feedback loop configuration where the filter first estimates the system state at some time and then obtains feedback in the form of noisy measurements (Bishop and Welch, 2001). The time update equations are responsible for projecting forward the current state and error covariance estimates to obtain the a priori estimates $\hat{\boldsymbol{x}}_k^-$ for the next time step. Therefore, the prediction step updates $\hat{\boldsymbol{x}}_{k-1}^+$ and $\hat{P}_{k-1}^+$ to $\hat{\boldsymbol{x}}_k^-$ and $\hat{P}_k^-$ by evaluating the state equations of the model.

In turn, the measurement update equations are responsible for the feedback -i.e. for incorporating a new measurement into the a priori estimate to obtain an improved a posteriori estimate $\hat{\boldsymbol{x}}_k^+$. Indeed the final estimation algorithm resembles that of a predictor-corrector algorithm for solving numerical problems as seen in Fig. 3. In the correction step, the state and covariance estimations are adjusted to $\hat{\boldsymbol{x}}_k^+$ and $\hat{P}_k^+$ using the output equations of the model and the current measurement $\boldsymbol{z}_k$.

The main challenge when dealing with nonlinear systems is not providing a proper prediction of the state $\boldsymbol{x_k}$ but providing an appropriate estimate of the state covariance $\hat{P}$. Assuming a Gaussian distribution around the state $\boldsymbol{x_k}$, the distribution will remain Gaussian for subsequent steps given a linear system. However, this is not completely true when dealing with nonlinear systems. To overcome nonlinearities, different filtering techniques have been proposed based on the original

Kalman filter approach (Kalman, 1960). Between others, in following we evaluate the estimation performance of the Extended Kalman Filter (EKF), the Cubature Kalman Filter (CKF) and the Unscented Kalman Filter (UKF).

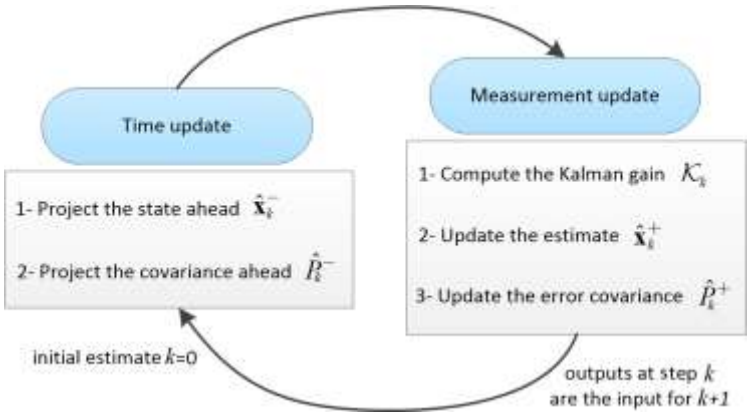


Figure 3: Kalman filter algorithm. The time update projects the current estimate ahead in time. The measurement update adjusts the projected estimate by an actual measurement at that time.

3.2.1 Extended Kalman Filter (EKF)

A Kalman filter that linearizes around the current mean and covariance is referred to as an EKF (Bishop and Welch, 2001). Similarly to a Taylor series, we can linearize around the current estimate using the partial derivatives of the transition and measurement functions and thus computing estimates even in the face of non-linear correspondences. The EKF has been the most studied filter in the presence of non-linear systems. However, when the state transition and observation models –i.e. the prediction and update functions are highly non-linear, the EKF can experience poor performance (Julier and Uhlmann, 1997). This is the result of propagating the covariance through linearization of the underlying non-linear dynamic model.

3.2.2 Unscented Kalman Filter (UKF)

In the unscented approach, the probability density is approximated by a deterministic sampling of sigma-points which represent the underlying distribution as a Gaussian. The nonlinear transformation of these points is intended to be an estimation of the posterior distribution, the moments of which can then be derived from the transformed samples. This transformation is known as the unscented transform (Wan and Merwe, 2000). The UKF tends to be more robust and more accurate than the EKF in its estimation of error in all the directions (Gustafsson and Hendeby, 2012). This technique also removes the requirement to explicitly calculate the Jacobian matrix,

which for complex functions can be a difficult task –either for requiring complicated function derivatives if done analytically or being computationally costly if done numerically.

3.2.3 Cubature Kalman Filter (CKF)

The heart of the CKF is a spherical-radial cubature rule, which makes it possible to numerically compute multivariate moment integrals encountered in the nonlinear Bayesian filter (Arasaratnam and Haykin, 2009). Specifically, the algorithm derives a third-degree spherical-radial cubature rule that provides a set of cubature points scaling linearly with the state-vector dimension. In consequence, the CKF may provide a systematic solution for high dimensional nonlinear filtering problems, and actually is numerically more stable than the UKF.

## 4. Results

### 4.1 Experimental setup

A common problem when implementing in-silico glucose-insulin models is that they are often developed assuming perfect representation of the metabolic process, in particular of the parameter values. However, high variability in this process leads to uncertainty in both the states and the parameters estimation. Thus, the stochastic formulation presented in Eq. (14) models this uncertain glycemic behavior. To show this stochastic behavior, due to the intra-inter patient variability, in Fig. 4 a number of simulations using the presented Hovorka stochastic model are shown using different scales of BG fluctuations. This 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, regarding Eq. (14), the noise scale parameter is progressively augmented from $\sigma$ = 0.10 (blue lines at top panel in Fig. 4) to 0.25 (green lines at central panel in Fig. 4) and 0.50 (red lines at bottom panel in Fig. 4), whereas all the remaining parameters are not modified. Notice that as the value of the parameter $\sigma$ grows also does the number and magnitude of sudden highs and lows in BG levels increase.

To generate the samples measurements of blood glucose we implement the stochastic system presented in Section 2 (Eq. 14) in a computational way. Thus, we take samples every T time

instants (integration step size) during periods of 24 hours. We use T = 6 minutes, therefore we have 240 samples/day.

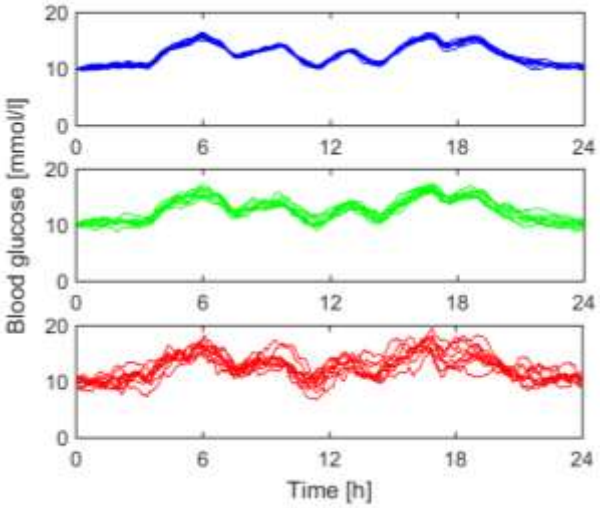


Figure 4: Implementations of the glucose stochastic process of a diabetic patient using different scales of variability, $\sigma$ = 0.10 (blue), 0.25 (green) and 0.50 (red).

In all the simulations, we followed a protocol with only one subcutaneous insulin infusion of 3 U of insulin injected at 5hs. It is assumed that the basal insulin infusion rate is equal to 1 [pmol/kg/min] as only short-acting insulin is administrated. This allowed us obtaining a clear graphical description of the plasma insulin estimated profile and how it is affected by large glycemic excursions. Besides this, we maintained the protocol of carbohydrate intake given in Table 2.

Table 2 Carbohydrate intake protocol.

| | | | | | | |
|---|---|---|---|---|---|---|
| Carbohydrate content (g) | 31 | 10 | 31 | 10 | 31 | 10 |
| Meal times (h) | 3.00 | 5.00 | 7.30 | 11.00 | 14.30 | 17.00 |

The parameters of the dynamic model presented in Section 2 are considered time invariant, thus to simplify the notation and we write them independent of the time.

For the Kalman Filter algorithms, we needed to determine the R and Q matrices. Since we know the variance of the Gaussian white noise we are injecting into the signals, we set the off-diagonal entries of R to 0 and set the diagonal entries to be the value of the noise variance value. Thus according to the previous work of (de Pereda et al., 2016), where the initial transition and

measurement covariance matrices were heuristically adjusted, we take the same tested initialization of these matrices, such that $Q_x = \text{diag}[20,20,0.01,10^{-6}, 10^{-8}, 10^{-5}, 1,1,1]$ and $R_z = 8$.

Random variables were adjusted to represent the transition noise $w_k = 0.01 * I_{9\times9}$ and measurement noise $v_k = 0.025$. 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. (14).

In the prediction step the nonlinear model state equations have to be evaluated. For the unscented transformation (Simon, 2006), Eq. (9) have to be integrated numerically. In this step the state noise covariance matrix Q can be used as design parameter. Increasing values result in quick decreasing values in a smooth estimation behavior.

As initial conditions (Fig. 2) the best possible guess for the states and their covariances should be used. For plasma glucose its basal value is chosen (Table 1) whereas all other estimates are started with zero. The UKF is initialized in the same way as the EKF, using the same values for the state vector and error covariance matrix upon startup. Cubature Kalman filter (CKF) is a special case of UKF with α = 1, β = 0, and κ = 0 – the mean weight becomes zero with these choices.

**4.2 Simulation Experiments**

Numerical simulations were computed so as to evaluate the performance of the EKF, UKF and CKF techniques by means of the root mean square error (RMSE) and the negative log-likelihood (NLL) metric. These metrics allow an assessment of the effectiveness of the uncertainty propagation algorithm through the nonlinear system for each of the filtering techniques. The aim of including 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 using GNU Octave (version 4.2.1) on an Intel(R) Core i7-7700hq Pentium(R) processor. For each degree of glycemic variability: low, medium, high with $\sigma$ = 0.10, 0.25 and 0.50, respectively; 100 runs were simulated and numerical results are shown in Tables 3 and 4.

From Fig. 5 to 7 we plot sensor measurements and the corresponding plasma insulin estimation for different amounts of glycemic variability. Notice that each filter computes the plasma insulin state estimate $I$ and besides the measured states $IG$ are shown. Additionally, the state covariance $\hat{P}$ estimated is plotted using yellow lines, in order to illustrate the reliability of the estimates. Also measurements of true insulin values (black line) are hidden to the estimator and they are only considered as a reference to describe the divergence of the insulin estimate. At first glance, we can observe that for small values of the Ito noise parameter, i.e. small BG fluctuations, insulin state estimation can be attained with acceptable accuracy regardless of the filtering approach employed. However, as glycemic variability is incremented through the noise parameter, the filters degrade their performance. This is clear since overestimation of the insulin peaks and marked fluctuations occurs. Especially, when a peak in the insulin estimated level occurs, the standard deviation considerably deviates emphasizing that the reliability of the estimate is quite limited.

One hundred runs of stochastic simulations for each filter per variability scenario were carried out and numerical results are shown in Table 3. Compared with EKF and UKF, the CKF performance is observed to be relatively consistent and reliable when comparing the estimation error, over a wide range of variability. Consistently, the NLL value says that the CKF presents less variance in the filtering process.

From Table 3, we have that for low variability the EKF and UKF show comparable performance to predict mean values and no significant difference between them can be found. The UKF is known to provide better and more reliable performance than the EKF, however the advantage in using the UKF is insignificant in terms of the RMSE. However, looking deeply at the estimated insulin plots in Figs. 5 to 7, we can observe that the EKF estimates seem to diverge at the end of the profile. Apparently, the first order linearization provides insufficient precision when there are significant nonlinearities in the glycemic system. For medium and large variability, the UKF outperforms the EKF in terms of the estimation precision for the present stochastic system.

In the presence of high variability, the performance of all the filters deteriorates considerably. Unlike for the previous variability cases, the distribution is noticeably distorted as a larger amount of variability in the glycemic dynamics is introduced. This is caused by the excessive nonlinearity introduced by Ito's parameter $\sigma$ in the stochastic system, which make the filtering process very unstable. In particular, an unstable behavior is observed for the EKF which is attributed to outlier estimates since the filter does not converge under large variability.

From Figs, 5 to 7, we can observe that the UKF performs reasonably well in terms of the RMSE but poorly in the NLL measure. This indicates that the mean of the UKF's filter estimations is on average close to the true insulin state, but the filter variances are overestimated leading to incoherent filter distributions. Remember that the RMSE value only evaluates how close the mean of the filtered state distribution is to the true state. NLL values reflect the coherence, that is, the accuracy in both the mean and the covariance of the filter distribution.

Table 3 Summary of errors for insulin estimation (100 runs in μU/ml)

| | EKF | | UKF | | CKF | |
|---|---|---|---|---|---|---|
| Variability | RMSE | NLL | RMSE | NLL | RMSE | NLL |
| Low ($\sigma = 0.1$) | 3.87±1.7 | 2.87± 0.6 | 3.72±1.4 | 3.00±0.9 | 2.79±1.0 | 2.14±0.6 |
| Medium ($\sigma = 0.25$) | 6.15±1.8 | 3.91±1.3 | 5.02±1.4 | 4.81±1.1 | 4.28±1.6 | 3.98±0.7 |
| High ($\sigma = 0.5$) | 10.05±4.5 | 7.88±1.8 | 8.84±1.8 | 7.99±1.5 | 7.34±1.7 | 6.122±1.9 |

In Table 4 a comparison between the computational times for the different filtering techniques is given. In terms of time consumptions, the EKF cost is slightly larger than the UKF's, and the EKF and UKF techniques are known to have nearly equivalent complexity (Thrun et al., 2005). Although the UKF algorithm requires more function evaluations in each time step due to sigma-points propagation, it requires a few less CPU time than the EKF in our implementation. This is probably due to the fact that the EKF requires costly matrix inversion operations in the linearization process due to the calculation of derivatives or Jacobian. The CKF algorithm is considered to have least numerical costs, as supported by the results. Noteworthy, we must consider that BG measurements are obtained every six minutes and thus neither of the filtering approaches may seem to have any problem to perform in real-time.

Based on the numerical results, the CKF is preferred over the EKF and UKF for the current application. The CKF shows satisfactory performance with excellent computational efficiency, as long as a reasonable amount of variability is presented. Detailed performance comparisons between the EKF and UKF revealed that the UKF outperforms the EKF in terms of filter stability and estimate accuracy at the price of minor increase in computational effort.

Table 4 Overall time consumption for insulin estimation (1 run in seconds)

| Variability | Model | EKF | UKF | CKF |
|---|---|---|---|---|
| Low ($\sigma = 0.1$) | 0.043 | 0.885 | 0.803 | 0.344 |

| | | | | |
|---|---|---|---|---|
| Medium ($\sigma = 0.25$) | 0.072 | 0.861 | 0.706 | 0.412 |
| High ($\sigma = 0.5$) | 0.11 | 0.912 | 0.871 | 0.502 |

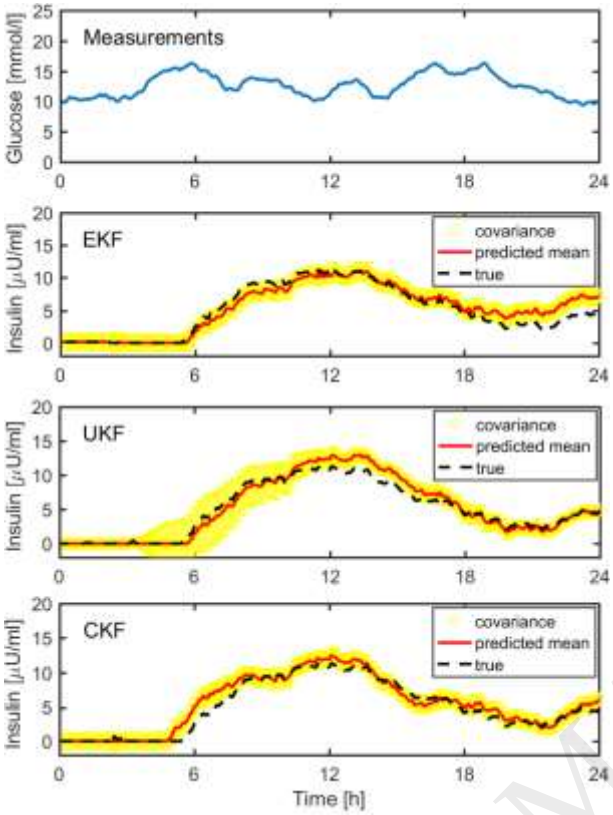


Figure 5: Glucose sensor measurements and plasma insulin estimation for low glycemic variability.

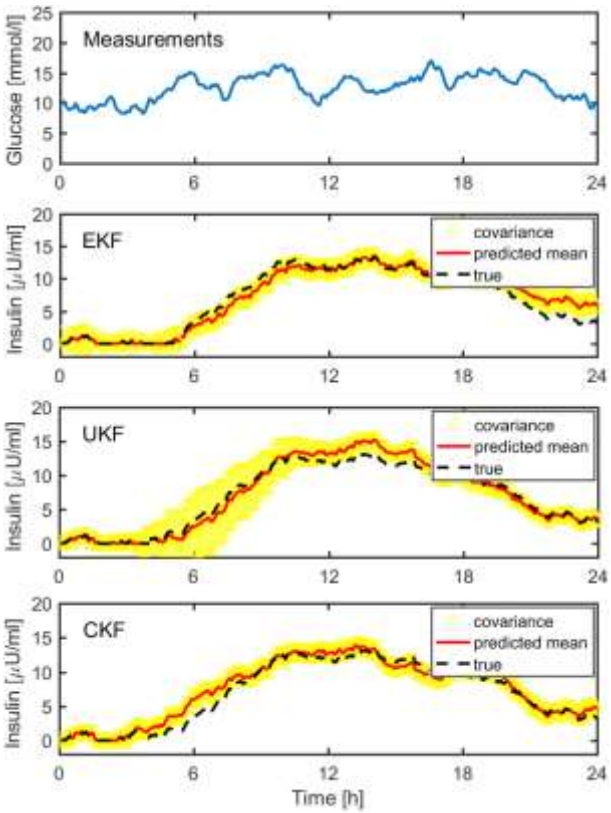


Figure 6: Glucose sensor measurements and plasma insulin estimation for medium glycemic variability.

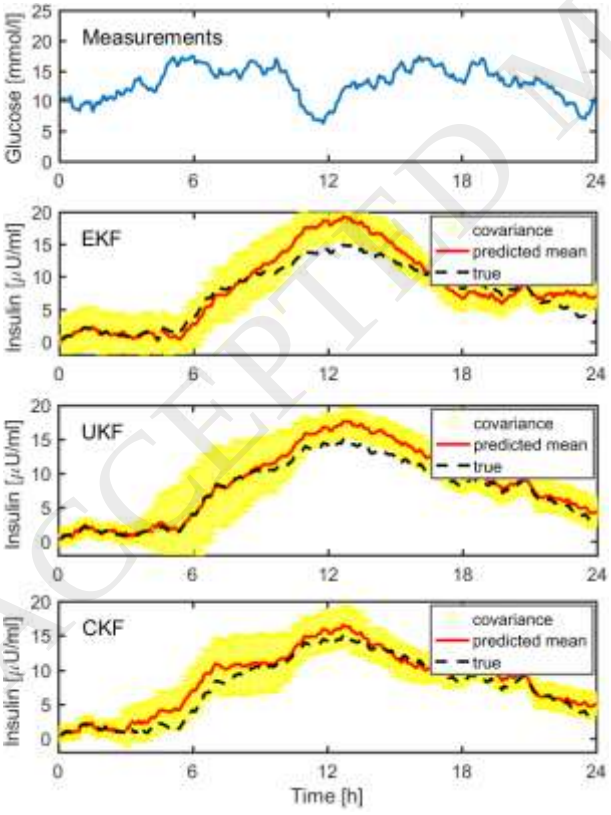


Figure 7: Glucose sensor measurements and plasma insulin estimation for high glycemic variability.

## 5. Concluding remarks

New diabetes technologies such as the artificial pancreas are promising for management of BG levels. However, most autonomous glycemic controllers need of real-time estimations of both plasma glucose and plasma insulin concentrations to increase the efficiency of control strategies. However, unlike BG levels which can be determined in real-time by means of continuous glucose monitors, plasma insulin still requires mathematical models to estimate current concentration.

A determinant factor for the estimation quality using filtering techniques is the availability of a truthful glucose-insulin model. Therefore, the dynamic model must be fixed for every particular patient, in order to consider individual metabolic behaviors and resulting deviations. Moreover, it is also important to understand how fluctuating BG levels affect the quality of the resulting insulin estimations. In order to achieve more realistic simulations of metabolic conditions, we formulated a stochastic version of the Hovorka glucose-insulin model and then we incorporated the stochastic model into different nonlinear versions of the Kalman filter, in order to estimate plasma insulin concentration under glycemic variability.

Notice that, the use of a stochastic model (Ito's model) does not require of adjusting the model parameters to a singular case. This allows us performing more realistic simulations of abnormal metabolic conditions while capturing a heterogeneous cohort of patients. In this way, we assessed the performance of the different filtering techniques in a wide range of patient inter and intra variability situations

For small BG fluctuations, insulin state estimation can be precisely attained regardless of the filtering approach employed. Unlike the conventional EKF with a first-order accuracy, the UKF guarantees that -at least- the second order accuracy of the posterior distribution is captured. The last made possible to achieve better estimations of plasma insulin concentration. However, as the amount of glycemic variability grows, both filters rapidly degrade their performance. In turn, the CKF algorithm provides a new direction to perform the nonlinear transformation, as the third-degree spherical-radial cubature rule is employed. Results show that the CKF technique not only presented lower errors compared to true insulin levels but also small values of standard deviation, and thus better management of state uncertainties. Because of the lower computational requirements and small estimation error, the CKF resulted to be the best option for this specific implementation.

Our aim is to evaluate filtering techniques feasible to be used in the development of robust control strategies to guarantying patient safety under glycemic variability. A limitation of this work is the necessity of an accurate mathematical parametric model, which is inaccessible as describing all the (hidden) aspects of the metabolic process is quite unattainable. A suitable solution and future work would be the use of an approximated model in a nonparametric filtering approach. On the other hand, despite continuous subcutaneous insulin infusion would be more appropriate in the context of an artificial pancreas, a single bolus of insulin is administered for a better appreciation of the insulin estimated state. Besides continuous insulin administration, the next step also considers the evaluation of the filtering techniques on a real human dataset. The last will certainly make possible to open a new avenue to implement an AP that can certainly face inter- and intra-patient variability. Finally, the issue of how to exploit the dynamics and capture correlations via transition and projection property via each version of Kalman Filter is an open avenue for subsequent immediate works.

Conflicts of interest: none

**References**

Ali, S., Padhi, R., 2011. Optimal blood glucose regulation of diabetic patients using single network adaptive critics. Optim. Control Appl. Methods 32, 196–214.

Arasaratnam, I., Haykin, S., 2009. Cubature kalman filters. IEEE Trans. Automat. Contr. 54, 1254–1269.

Bishop, G., Welch, G., 2001. An introduction to the kalman filter. Proc SIGGRAPH, Course.

Borg, R., Kuenen, J.C., Carstensen, B., Zheng, H., Nathan, D.M., Heine, R.J., Nerup, J., Borch-Johnsen, K., Witte, D.R., Group, A.S., 2010. Real-life glycaemic profiles in non-diabetic individuals with low fasting glucose and normal HbA1c: the A1C-Derived Average Glucose (ADAG) study. Diabetologia 53, 1608–1611.

De Paula, M., Ávila, L.O., Martínez, E.C., 2015. Controlling blood glucose variability under uncertainty using reinforcement learning and Gaussian processes. Appl. Soft Comput. 35, 310–332.

de Pereda, D., Romero-Vivo, S., Ricarte, B., Rossetti, P., Ampudia-Blasco, F.J., Bondia, J., 2016. Real-time estimation of plasma insulin concentration from continuous glucose monitor measurements. Comput. Methods Biomech. Biomed. Engin. 19, 934–942.

Eberle, C., Ament, C., 2011. The unscented Kalman filter estimates the plasma insulin from glucose measurement. Biosystems 103, 67–72.

Ghosh, S., Maka, S., 2014. A constrained sub-optimal controller for glucose regulation in type 1 diabetes mellitus. Optim. Control Appl. Methods 35, 191–203.

Gustafsson, F., Hendeby, G., 2012. Some relations between extended and unscented Kalman filters. IEEE Trans. Signal Process. 60, 545–555.

Hovorka, R., Canonico, V., Chassin, L.J., Haueter, U., Massi-Benedetti, M., Federici, M.O., Pieber, T.R., Schaller, H.C., Schaupp, L., Vering, T., 2004. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. Physiol. Meas. 25, 905–920.

Hughes, C.S., Patek, S.D., Breton, M., Kovatchev, B.P., 2011. Anticipating the next meal using meal behavioral profiles: A hybrid model-based stochastic predictive control algorithm for T1DM. Comput. Methods Programs Biomed. 102, 138–148.

Ito, K., 1975. Stochastic differentials. Appl. Math. Optim. 1, 374–381.

Julier, S.J., Uhlmann, J.K., 1997. New extension of the Kalman filter to nonlinear systems, in: AeroSense'97. International Society for Optics and Photonics, pp. 182–193.

Kalman, R., 1960. A new approach to linear filtering and prediction problems. J. basic Eng. 82, 35–45.

Kocarev, L., Parlitz, U., Hu, B., 1998. Lie derivatives and dynamical systems. Chaos, Solitons & Fractals 9, 1359–1366.

Krüger, L., Slabber, M., Joubert, G., Venter, C., 2003. Intra-and inter-individual variation in blood glucose response to white bread and glucose in patients with type 2 diabetes mellitus. South African J. 93, 18–29.

Maahs, D.M., Buckingham, B.A., Castle, J.R., Cinar, A., Damiano, E.R., Dassau, E., DeVries, J.H., Doyle, F.J., Griffen, S.C., Haidar, A., 2016. Outcome measures for artificial pancreas clinical trials: a consensus report. Diabetes Care 39, 1175–1179.

Makroglou, A., Li, J., Kuang, Y., 2006. Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: an overview. Appl. Numer. Math. 56, 559–573.

Marling, C.R., Shubrook, J.H., Vernier, S.J., Wiley, M.T., Schwartz, F.L., 2011. Characterizing Blood Glucose Variability Using New Metrics with Continuous Glucose Monitoring Data. J. Diabetes Sci. Technol. 5, 871–878.

Møller, J.B., 2008. Stochastic Models and Methods to characterize the Glucose/Insulin System. Technical University of Denmark, Lyngby, Denmark.

Neatpisarnvanit, C., Boston, J.R., 2002. Estimation of plasma insulin from plasma glucose. IEEE Trans. Biomed. Eng. 49, 1253–1259.

Nguyen, N.A.T., Yang, H.-J., Kim, S., 2017. HOKF: High Order Kalman Filter for Epilepsy Forecasting Modeling. Biosystems.

Patek, S.D., Magni, L., Dassau, E., Karvetski, C., Toffanin, C., De Nicolao, G., Del Favero, S., Breton, M., Dalla Man, C., Renard, E., 2012. Modular closed-loop control of diabetes. IEEE Trans. Biomed. Eng. 59, 2986–2999.

Sepasi, S., Ghorbani, R., Liaw, B.Y., 2014. A novel on-board state-of-charge estimation method for aged Li-ion batteries based on model adaptive extended Kalman filter. J. Power Sources 245, 337–344.

Siegelaar, S.E., Holleman, F., Hoekstra, J.B.L., DeVries, J.H., 2010. Glucose Variability; Does It Matter? Endocr. Rev. 31, 171–182.

Simon, D., 2006. Optimal state estimation: Kalman, H infinity, and nonlinear approaches. John Wiley & Sons.

Sparacino, G., Facchinetti, A., Cobelli, C., 2010. "Smart" continuous glucose monitoring sensors: on-line signal processing issues. Sensors 10, 6751–6772.

Steil, G.M., Clark, B., Kanderian, S., Rebrin, K., 2005. Modeling Insulin Action for Development of a Closed-Loop Artificial Pancreas. Diabetes Technol. Ther. 7, 94–108.

Szalay, P., Sas, P.I., Barkai, L., Kovács, L., 2012. Nonlinear analysis of Type 1 Diabetes Models by Differential Geometric Approach. IFAC Proc. Vol. 45, 55–60.

Thrun, S., Burgard, W., Fox, D., 2005. Probabilistic robotics. MIT press.

Turksoy, K., Hajizadeh, I., Samadi, S., Feng, J., Sevil, M., Park, M., Quinn, L., Littlejohn, E., Cinar, A., 2017. Real-time insulin bolusing for unannounced meals with artificial pancreas. Control Eng. Pract. 59, 159–164.

Wan, E., Merwe, R. Van Der, 2000. The unscented Kalman filter for nonlinear estimation. 2000. AS-

SPCC. IEEE 2000.

Wang, Q., Molenaar, P., Harsh, S., Freeman, K., Xie, J., Gold, C., Rovine, M., Ulbrecht, J., 2014. Personalized State-space Modeling of Glucose Dynamics for Type 1 Diabetes Using Continuously Monitored Glucose, Insulin Dose, and Meal Intake. J. Diabetes Sci. Technol. 8, 331–345.

Zisser, H., Robinson, L., Bevier, W., Dassau, E., Ellingsen, C., Doyle III, F.J., Jovanovic, L., 2008. Bolus calculator: a review of four "smart" insulin pumps. Diabetes Technol. Ther. 10, 441–444.