



Research article

An OGI model for personalized estimation of glucose and insulin concentration in plasma

Weijie Wang¹, Shaoping Wang^{1,2*}, Yixuan Geng¹, Yajing Qiao¹, Teresa Wu³

¹ School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China

² Beijing Advanced Innovation Center for Big Data-based Precision Medicine, Beihang University, Beijing 100191, China

³ School of Computing, Informatics, and Decision Systems Engineering, Arizona State University and College of Medicine, Mayo Clinic, Tempe AZ 85281, the USA

* **Correspondence:** Email: shaopingwang@vip.sina.com/shaopingwang@buaa.edu.cn.

Abstract: Plasma glucose concentration (PGC) and plasma insulin concentration (PIC) are two essential metrics for diabetic regulation, but difficult to be measured directly. Often, PGC and PIC are estimated from continuous glucose monitoring and insulin delivery data. Nevertheless, the inter-individual variability and external disturbance (e.g. carbohydrate intake) bring challenges for accurate estimations. This study is to estimate PGC and PIC adaptively by identifying personalized parameters and external disturbances. An observable glucose-insulin (OGI) dynamic model is established to describe insulin absorption, glucose regulation, and glucose transport. The model parameters and disturbances can be extended to observable state variables and be identified dynamically by Bayesian filtering estimators. Two basic Gaussian noise based Bayesian filtering estimators, extended Kalman filtering (EKF) and unscented Kalman filtering (UKF), are implemented. Recognizing the prevalence of non-Gaussian noise, in this study, two new filtering estimators: particle filtering with Gaussian noise (PFG), and particle filtering with mixed non-Gaussian noise (PFM) are designed and implemented. The proposed OGI model in conjunction with the estimators is evaluated using the data from 30 in-silico subjects and 10 human participants. For in-silico subjects, the OGI with PFM estimator has the ability to estimate PIC and PGC adaptively, achieving RMSE of PIC 9.49 ± 3.81 mU/L, and PGC 0.89 ± 0.19 mmol/L. For human, the OGI with PFM has the promise to identify disturbances ($95.46\% \pm 0.65\%$ accurate rate of meal identification). OGI model provides a way to fully personalize the parameters and external disturbances in real time, and has potential clinical utility for artificial pancreas.

Keywords: artificial pancreas; model personalization; Bayesian filter; plasma glucose concentration; plasma insulin concentration

1. Introduction

Type 1 diabetes is a chronic autoimmune disease that insulin production ceases because an autoimmune attack destroys the insulin-producing cells (β -cells) in the pancreas [1]. Insulin deficiency can induce high blood glucose (BG) and cause various complications (i.e., retinopathy, angiocardopathy, and neuropathy). Generally, people with type 1 diabetes rely on exogenous insulin infusion to manage blood glucose levels. The time and amount of exogenous insulin infusion directly determine the controlling effectiveness of glucose level [2]. Given that the insulin demand always changes along with the changes of individual condition, it is difficult for patients to have appropriate and timely insulin infusion [3].

Recent technology advanced in the continuous glucose monitoring (CGM) [4,5] and insulin pump [6] has resulted in the development of artificial pancreas (AP) for diabetes care [7,8]. AP also refers to a closed-loop glucose regulation system that can offer glucose-responsive insulin delivery automatically, by integrating the CGM system and insulin pump [9,10].

At this stage, the most widely used CGMs are to measure interstitial glucose concentration [11]. It is worth noting that the decline of BG caused by insulin and the raise of BG caused by carbohydrate intake both occur in plasma. Obtaining the plasma glucose concentration (PGC) is helpful for AP to control glucose changes [12]. Similarly, given the safety and convenience, insulin pump normally infuses exogenous insulin in subcutaneous tissue [13]. Indeed, the infused insulin cannot be absorbed rapidly, remaining at the subcutaneous tissue before entering circulation [12], especially after insulin bolus. The amount of previously infused insulin that is present in plasma or the subcutaneous space, referred to as the insulin on board (IOB), must be quantified to prevent overdosing [14]. Therefore, it is necessary to estimate PIC and PGC from the sampled glucose and insulin infusion data in subcutaneous tissue [15].

Recognizing that (1) the inherent noise error of sensor and accuracy of mathematical models have a significant impact on the estimation accuracy, (2) the manual input of carbohydrate intake information can introduce another source of error from patient's inputs [16], filtering technology are adopted to estimate the hidden states [17] (i.e. the glucose and insulin concentration in plasma and the carbohydrate intake). For example, De Pereda et al. estimated the plasma insulin concentration by combining the classical Hovorka glucose-insulin dynamic model with extended Kalman filtering (EKF) [18]. Turksoy et al. detected the carbohydrate intake using unscented Kalman filtering (UKF), and the effectiveness was verified using the UVa/Padova simulation platform [19]. Charalampidis et al. introduced particle filter technology to the estimation of plasma glucose concentration for ICU subjects, overcoming the uncertainty from glucose measurement error [20]. Considering the individual variability, Hajizadeh et al. [21] estimated PIC by estimating the extended parameters related to insulin absorption based on UKF. While promising, the efforts reviewed above still face the challenges from the model parameter uncertainty due to the inter-individual variability. Specifically, since only part of the model parameters extendable to observable state variables can be identified dynamically [22], research to-date has mainly focused on personalizing part of the model parameters related to insulin absorption. The inter-individual variability presents in whole metabolism [23,24]. Although many models have been proposed and analyzed [25–28], we contend that (1) there is a need to establish a glucose-insulin dynamic model of which the parameters and disturbances can be extended to observable variables; and (2) an appropriate Bayesian filtering estimator needs to be developed to personalize the parameters and

disturbances.

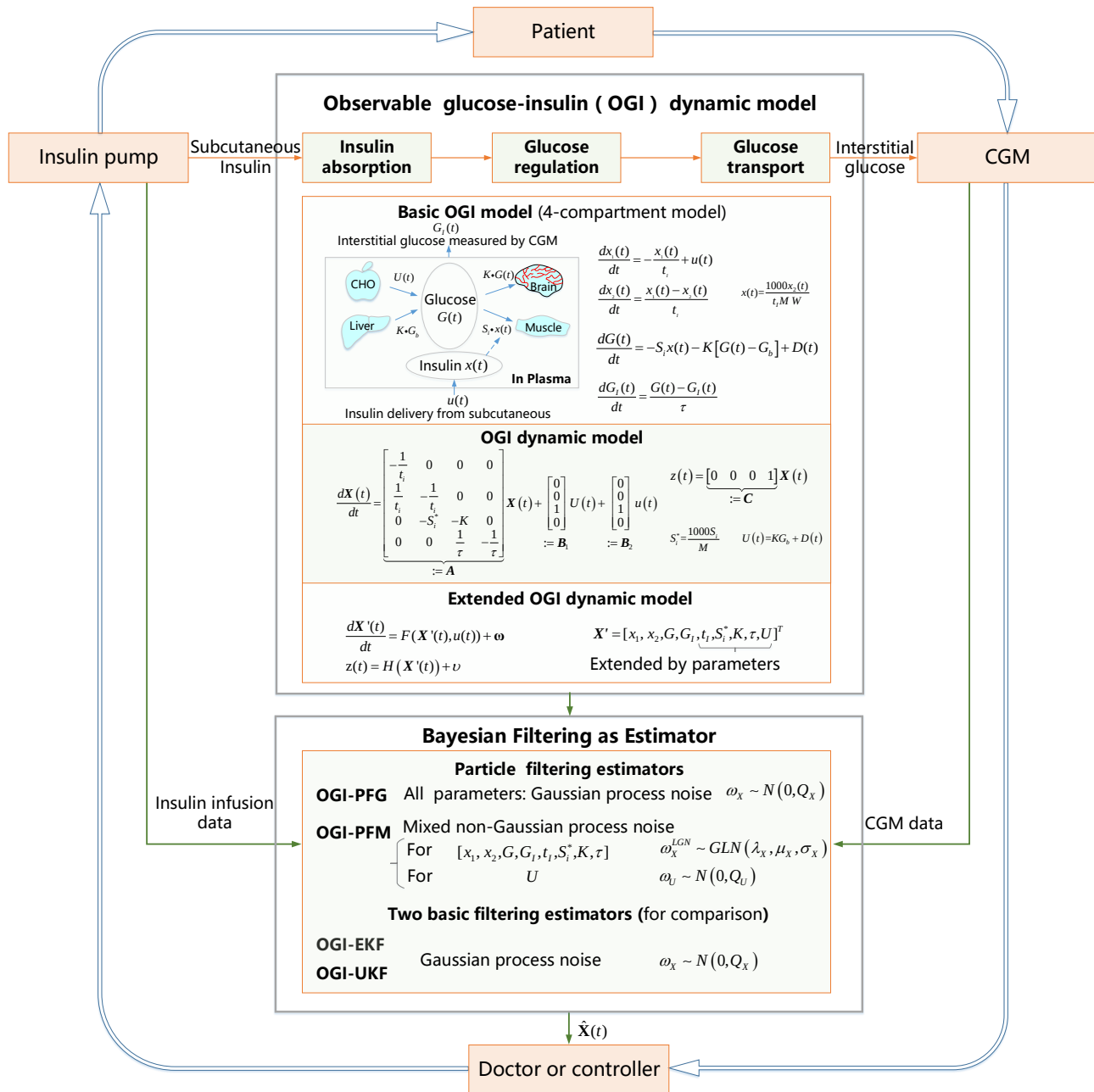


Figure 1. The OGI dynamic model for PGC and PIC estimation.

In this research, a new observable glucose-insulin (OGI) dynamic model is proposed. According to the physiological mechanism of glucose metabolism, the OGI model describes the integrated glycemic regulation with three processes: insulin absorption, glucose regulation and glucose transport. In the OGI model, the model parameters can be extended to the observable variables and mapped by CGM data. Next, the Bayesian filtering technology is implemented to personalize the model parameters and external disturbances for PIC and PGC estimation in real time. EKF and UKF are first explored. Since both the filtering techniques are built upon Gaussian propagation process [22], the applicability of these

filters may be questionable given that some state variables are non-negative due to the physiological constraints. In view of this, an emerging filtering technology, particle filtering (PF) [29] is introduced. PF obtains numerical results by Monte Carlo simulation, having advantages in dealing with nonlinear engineering problems and non-Gaussian assumption [30, 31]. In this study, the applicability of PF and its extensions to OGI is investigated comparing to EKF and UKF.

The rest of this paper is organized as follows. In Section 2, the observable glucose-insulin (OGI) dynamic model is developed and the observability of OGI dynamic model is verified. In section 3, four Bayesian filtering estimators, based on EKF, UKF, the PF with Gaussian noises (PFG), and the PF with mixed non-Gaussian noises (PFM), are designed to estimate the extended state variables (i.e. the model parameters and external disturbances) dynamically and to estimate the plasma glucose and insulin concentration in real time. Section 4 introduces the experiment for validation. The results and discussions are illustrated in section 5 and 6. The conclusion is drawn in section 7.

2. Modelling

2.1. Basic OGI model

As seen in Figure 1, basic OGI model considers three processes: (1) the insulin absorption from subcutaneous infusion to plasma, (2) the glucose regulation under the influence of PIC and external disturbance, and (3) the glucose transport from plasma to the subcutaneous interstitial, where the glucose concentration can be measured by CGM.

2.1.1. Insulin absorption

The insulin absorption sub-model describes the insulin absorption process from subcutaneous infusion to plasma. As given by [23], a two-compartment model is established to describe this process:

$$\begin{aligned}\frac{dx_1(t)}{dt} &= -\frac{x_1(t)}{t_I} + u(t) \\ \frac{dx_2(t)}{dt} &= \frac{x_1(t) - x_2(t)}{t_I} \\ x(t) &= \frac{1000x_2(t)}{t_I M W}\end{aligned}\quad (2.1)$$

where $u(t)$ represents the exogenous insulin delivery rate (U/min, 1U/min=1/60 U/h) at time t (immediate insulin bolus is modeled as a short burst insulin infusion), $x_1(t)$ and $x_2(t)$ represent the amount of effective insulin in the first and second insulin absorption compartment respectively (U), $x(t)$ is the plasma insulin concentration (PIC) (mU/L), the number 1000 is derived from the unit conversion (1 U/L=1000 mU/L), t_I is the time-to-maximum of effective insulin concentration (min), M is the metabolic clearance rate (L/kg/min) of effective insulin, and W is the measured body weight (kg).

2.1.2. Glucose regulation

The glucose regulation sub-model describes the kinetics of PGC regulation. Similar to the Bergman model [32, 33], this regulation is described by a single-compartment model, where the rate of PGC changes equals the production of glucose minus the consumption of glucose. As shown in Figure 1, the glucose production involves the endogenous glucose production (by liver) and the digestion of carbohydrate. The glucose consumption involves the consumption independent of insulin (by brain),

and the consumption stimulated by insulin (by the liver and muscles). Here, supported by [23, 34, 35], it is assumed that the insulin-dependent glucose utilization does not depend on PGC. That is

$$\frac{dG(t)}{dt} = -S_i x(t) - K[G(t) - G_b] + D(t) \quad (2.2)$$

where $G(t)$ is the PGC (mmol/L), S_i is the effect of insulin sensitivity (mmol/L/min per mU/L), G_b is the basal glucose level (mmol/L) and K is the glucose self-regulation fractional rate (min^{-1}) which has an effect of glucose self-regulation towards G_b . $D(t)$ is a lumped signal to describe the external disturbance to glucose changes, such as carbohydrate intake and exercise (mmol/L/min).

2.1.3. Glucose transport

The CGM measures the glucose level in interstitial fluid. Lumping the dynamics introduced by the CGM systems, the plasma glucose converts into interstitial glucose across a concentration gradient with time lag τ (min). The interstitial glucose concentration is given as

$$\frac{dG_I(t)}{dt} = \frac{G(t) - G_I(t)}{\tau} \quad (2.3)$$

where $G_I(t)$ is the interstitial glucose concentration which can be measured by CGM (mmol/L).

2.2. Observable Glucose-insulin dynamic model

The inter-individual variability induces the variability of model parameters, named parameter uncertainty, which has a direct effect on PIG and PGC estimation. To identify parameters by filtering techniques, the model parameters should be extendable to observable state variables, as discussed in [18].

It is the purpose of this study to achieve a deeper personalization, not limited to personalize the parameters related to insulin absorption. We merged and simplified the parameters to form an OGI dynamic model. To summarize, the OGI dynamic model can be expressed as

$$\begin{aligned} \frac{d}{dt} \mathbf{X}(t) &= \underbrace{\begin{bmatrix} -\frac{1}{t_I} & 0 & 0 & 0 \\ \frac{1}{t_I} & -\frac{1}{S_i^*} & 0 & 0 \\ 0 & -\frac{S_i^*}{t_I W} & -K & 0 \\ 0 & 0 & \frac{1}{\tau} & -\frac{1}{\tau} \end{bmatrix}}_{:=\mathbf{A}} \mathbf{X}(t) + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}}_{:=\mathbf{B}_1} U(t) + \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{:=\mathbf{B}_2} u(t) \\ z(t) &= \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}}_{:=\mathbf{C}} \mathbf{X}(t) \end{aligned} \quad (2.4)$$

where $\mathbf{X} = [x_1, x_2, G, G_I]^T$. The $U(t)$ represents the compound uncertainty of the system, composed by K, G_b and $D(t)$, $U(t) = KG_b + D(t)$. $S_i^* = \frac{1000S_i}{M}$ represents a parameter related to insulin sensitivity. $z(t)$ is the measurable output of system (interstitial glucose concentration $G_I(t)$).

On this basis, the parameters $[t_I, S_i^*, K, \tau]$ and compound uncertainty U are generally regarded as the extended variables, $\frac{d\theta}{dt} = 0$, $\theta \in \Theta$, $\Theta = [t_I, S_i^*, K, \tau, U]$. Here, the extended OGI dynamic model can be formed as:

$$\begin{aligned} \frac{d\mathbf{X}'(t)}{dt} &= \mathbf{F}(\mathbf{X}'(t), u(t)) \\ z(t) &= \mathbf{H}(\mathbf{X}'(t)) \end{aligned} \quad (2.5)$$

where X' is the vector of state variables, $X' = [x_1, x_2, G, G_I, t_I, S_i^*, K, \tau, U]^T$.

To construct the estimator (observer), observability ensures that the state variable can be observed and estimated [22], i.e., the state variable $X'(t)$ can be uniquely mapped to the output $z(t)$.

For a nonlinear system, the proof of observability is usually performed by calculating the observability matrix derived from the chain of Lie derivatives [36, 37]. The observability matrix of nonlinear system is

$$\psi(t) = \begin{bmatrix} \frac{\partial}{\partial X'} L_F^0 (H(X'(t))) \\ \frac{\partial}{\partial X'} L_F^1 (H(X'(t))) \\ \dots \\ \frac{\partial}{\partial X'} L_F^{n-1} (H(X'(t))) \end{bmatrix} \quad (2.6)$$

where n is the order of the system, $L_F^k [H(X'(t))] = \frac{\partial}{\partial X'} L_F^{k-1} [H(X'(t))] F$, and $L_F^0 [H(X'(t))] = H(X'(t))$. For the extended OGI dynamic model, the rank of observability matrix is $\text{Rank}(\psi(t)) = 9$, and it is concluded that all the state variables are observable. That is, all the state variables $[x_1, x_2, G, G_I, t_I, S_i^*, K, \tau, U]$ can be estimated from glucose measurement.

Since the CGM data are sampled discretely, a discrete OGI dynamic model is rewritten using the first forward difference derivative approximation. To be consistent with practice, unmodeled dynamics is introduced (refer to process noise). Besides, considering the parameter uncertainty, the extended states are also incorporated in process noise [21]. To characterize the uncertainty of glucose measurement, the measurement noise is added. The discrete OGI dynamic model is

$$\begin{aligned} x_1(k+1) &= \left(-\frac{x_1(k)}{t_I(k)} + u(k) \right) T + x_1(k) + \omega_{x1}(k) \\ x_2(k+1) &= \left(\frac{x_1(k) - x_2(k)}{t_I(k)} \right) T + x_2(k) + \omega_{x2}(k) \\ G(k+1) &= \left(\frac{-S_i^*(k)x_2(k)}{t_I(k)W} + U(k) - K(k)G(k) \right) T + G(k) + \omega_G(k) \\ G_I(k+1) &= \left(\frac{G(k) - G_I(k)}{\tau(k)} \right) T + G_I(k) + \omega_{G_I}(k) \\ t_I(k+1) &= t_I(k) + \omega_{t_I}(k) \\ S_i^*(k+1) &= S_i^*(k) + \omega_{S_i^*}(k) \\ K(k+1) &= K(k) + \omega_K(k) \\ \tau(k+1) &= \tau(k) + \omega_\tau(k) \\ U(k+1) &= U(k) + \omega_U(k) \\ z(k+1) &= G_I(k+1) + v(k) \end{aligned} \quad (2.7)$$

where T is the sampling period. ω represents the process noise vector, $\omega = [\omega_{x1}, \omega_{x2}, \omega_G, \omega_{G_I}, \omega_{t_I}, \omega_{S_i^*}, \omega_K, \omega_\tau, \omega_U]^T$, and v represents the measurement noise.

3. Bayesian filtering as estimator

Bayesian filtering technology incorporates the Bayesian inference with first-order hidden Markov process [38]. The posterior distributions of the state variables are recursively conditioned on sensor information collected up till the present moment. Based on the Bayesian inference, the posterior distribution X_k can be obtained by

$$p(X_k | z_{1:k}) \propto p(z_k | X_k) p(X_k | z_{1:k-1}) \quad (3.1)$$

where $p(z_k|X_k)$ is the likelihood probability at the k^{th} sampling time, $p(X_k|z_{1:k-1})$ represents the prediction-step probability determined by the previous results $p(X_{k-1}|z_{1:k-1})$, that is

$$p(X_k|z_{1:k-1}) = \int p(X_k|X_{k-1})p(X_{k-1}|z_{1:k-1})dX_{k-1} \quad (3.2)$$

Differing in the specific way to update $p(X_k|z_{1:k})$ and $p(X_k|z_{1:k-1})$, four filtering estimators are designed in the following parts, including extended Kalman filtering (EKF), unscented Kalman filtering (UKF), particle filtering with Gaussian noises (PFG), and PF with mixed non-Gaussian noises (PFM). The EKF and UKF are developed for comparison with PF.

3.1. EKF and UKF

EKF and UKF are based on the Kalman filtering principle. Both of them adopt the Gaussian propagation to obtain the posterior solution, i.e., the process noise and measurement noise are assumed to be Gaussian [22] ($\omega \sim N(0, Q)$ and $\nu \sim N(0, R)$). In EKF, the nonlinear model is approximated with first-order Taylor expansion at the last estimated point, then the standard Kalman filtering formulas are used to calculate the posterior distribution [18]. The details of the EKF algorithm are given in Appendix. In UKF, unscented transformation (UT) is adopted. It uses a deterministic sampling approach to estimate the mean and covariance of variables with a minimal set of sample points, also called sigma points [39]. These sigma points are propagated through the true nonlinear system, meanwhile the weighted mean and the covariance are calculated. The details of the UKF algorithm are given in Appendix. The determination of the noise covariance matrix in EKF and UKF is given in section 4.

3.2. Particle filtering technology

3.2.1. PFG-particle filtering for Gaussian noise

Particle filtering (PF) approximates the posterior distribution of state variables based on Monte Carlo statistical method [31]. For each $X \in \mathbf{X}$, the probability $p(X_k|z_{1:k})$ can be approximated by N random particles $X_k^{(i)}$ and their associated weights $W_{X_k}^{(i)}$ in a weighted summation, such as

$$p(X_k|z_{1:k}) = \sum_{i=1}^N W_{X_k}^{(i)} X_k^{(i)} \quad (3.3)$$

As the measured data are updated, the weights are adjusted in real time. In each update step, sequential importance sampling (SIS) is incorporated to achieve the iteration [31]. The weight of the i^{th} particle at the k^{th} step can be calculated by

$$W_{X_k}^i \propto W_{X_{k-1}}^i \frac{p(z_k|X_k^i)p(X_k^i|X_{k-1}^i)}{q(X_k^i|X_{k-1}^i, z_k)} \quad (3.4)$$

where $W_{X_{k-1}}^i$ is the weight of the i^{th} particle at the last step, $p(z_k|X_k^i)$ is the likelihood probability related to the measurement function and measurement noise ν , $p(X_k^i|X_{k-1}^i)$ is the prior probability related to the state transition function and process noise ω_X , $\omega_X \in \omega$, $q(X_k^i|X_{k-1}^i, z_k)$ is the sampling probability

related to the distribution of particle samples. Generally, the particles are sampled from the prior distribution, and the weight of the i^{th} sampled particle is

$$W_{X_k}^i \propto W_{X_{k-1}}^i p(z_k | X_k^i) \quad (3.5)$$

Nevertheless, the particles tend to degenerate in the recursive estimation, and few particles survive if the case comes to the worst. In view of this, a resampling procedure is conducted in regions of higher sampling importance [40]. In other words, particles with higher probabilities generate more new particles.

For a fair comparison, Gaussian noise is taken in PF, i.e., $\omega \sim N(0, Q)$ and $v \sim N(0, R)$. This filtering is termed PFG (the details of PFG algorithm is given in Appendix).

3.2.2. PFM - particle filtering for mixed non-Gaussian noise

The state variables $[x_1, x_2, G, G_I, t_I, S_i^*, K, \tau, U]$ consist of (1) the state variables with non-negativity constraint, that is, $[x_1, x_2, G, G_I]$, which are transformed from glycemic model variables, and the state variables $[t_I, S_i^*, K, \tau]$, which are extended from model parameters with non-negative physiological meaning; (2) the compound uncertainty U can be either positive or negative (specifically, positive D indicates the glycemic effect induced by carbohydrate disturbance, while negative D indicates the hypoglycemic effect induced by exercise disturbance).

As for the state variables with non-negativity constraint, $[x_1, x_2, G, G_I, t_I, S_i^*, K, \tau]$, the shifted log-normal distribution [41] with three parameters is introduced to model the process noise to ensure the non-negativity of state variables. For each process noise of the non-negative state variable ω_X^{GLN} , it is assumed $\omega_X^{GLN} \sim GLN(\lambda_X, \mu_X, \sigma_X)$ with PDF

$$f_{GLN}(\omega_X^{GLN}) = \frac{\exp\left\{-\frac{1}{2\sigma_X^2}[\ln(\omega_X^{GLN} + \lambda_X) - \mu_X]^2\right\}}{(\omega_X^{GLN} + \lambda_X) \sqrt{2\pi}\sigma_X} \quad (3.6)$$

where λ_X is the position parameter, μ_X is the shape parameter, and σ_X is the scale parameter. The position parameter λ_X shifts the log-normal distribution along X-axis to ensure that the expectation of process noise is 0. Thus, log-normal distribution is chosen as the prior distribution of each non-negativity parameter, just like the assumptions in [23]. Specifically, the hyper-parameters $\lambda_X^k, \mu_X^k, \sigma_X^k$ at the k^{th} step can be derived as

$$\begin{aligned} \lambda_X^k &= F(X_{k-1}) \\ \mu_X^k &= \ln(F(X_{k-1})) - \frac{1}{2} \ln\left(1 + \frac{\text{var}(F(X_{k-1}) + \omega_X)}{F(X_{k-1})^2}\right) \\ \sigma_X^k &= \sqrt{\ln\left(1 + \frac{\text{var}(F(X_{k-1}) + \omega_X)}{F(X_{k-1})^2}\right)} \end{aligned} \quad (3.7)$$

As for the compound uncertainty U , without the limitation of non-negativity, its process noise is still set to Gaussian. This filtering mixed non-Gaussian noise is termed PFM (the details of PFM algorithm is given in Appendix).

4. Experiments

The objective of this study is to develop an observable glucose-insulin model of which the parameters can be personalized, to estimate plasma glucose concentration (PGC) and plasma insulin

concentration (PIC) dynamically. Taking a progressive research approach, three sets of experiments are conducted: (1) Experiment I is to prove that the proposed model has the ability to estimate PIC and PGC. (2) Built upon the success from the first experiment, Experiment II is to validate the capability of our model in identifying physiological parameters, in this study, 4 parameters, $[t_I, S_i^*, K, \tau]$. Considering the difficulty to continuously monitor PIC and PGC values on human participants, 30 in-silico subjects with type 1 diabetes produced by UVa/Padova simulator [42] are used both in Experiment I and Experiment II. Other than the 4 physiological parameters, we have a compound uncertainty parameter U , related to the disturbance, in the OGI model. (3) Since U is a time-varying parameter changed in responding to meal intakes, we design Experiment III to demonstrate that parameter U is identifiable as an indicator for meal identification. Recognizing that the meal intake of in-silico subjects only have intra-day variability, not inter-day variability, we have extended Experiment III to include 10 human subjects.

Table 1. The schedule of carbohydrate intake for all in-silico subjects.

	Breakfast	Lunch	Snack	Dinner	Snack
Time	7:00	12:00	16:00	18:00	23:00
Carbohydrate	45(g)	70(g)	20(g)	80(g)	20(g)

The UVa/Padova simulator is a substitute for preclinical testing in diabetes care accepted by FDA (Food and Drug Administration) [43]. The model embedded in UVa/Padova simulator is different from the OGI model, which consists of 12 nonlinear differential equations, 18 algebraic equations and 35 parameters [42]. In general, this model is not usable for the quantification of specific metabolic relationships, but is suitable for system simulation [44,45]. It contains 30 in-silico subjects (10 adults, 10 adolescents, 10 children), having been widely used in glycemic modeling and simulation [46–48]. In the UVa/Padova simulator, a simulation scenario was developed. The experimental duration was set to one week (7 days), and the insulin injection rate of each subject was set to the default value, determined automatically by the simulator in terms of the subject's basal injection rate and the specific ratio of carbohydrate to insulin. The schedule of daily carbohydrate intake is given in Table 1. In total, 1050 meals were produced in the experiment.

Table 2. The information of 10 human participants.

ID of participant	1	2	3	4	5
Gender	F	F	M	F	M
Age (year)	25	32	18	23	26
Weight (kg)	56	60	70	44	57
Duration of diabetes (year)	14	20	10	8	10
ID of participant	6	7	8	9	10
Gender	F	F	F	M	M
Age (year)	22	31	37	26	29
Weight (kg)	50	80	75	60	52
Duration of diabetes (year)	8	14	23	13	14

This study was approved by the institutional review boards (IRB), and all the participants signed

informed consent forms prior to participation. The study was based on the group of 10 participants with type 1 diabetes (see Table 2 for details). This dataset was used in [24]. The insulin and glucose data of each participant were monitored over one week. All the participants used insulin pumps (MiniMed 712, Medtronic) for insulin Lispro infusion, and they did not take any drugs other than exogenous insulin infusion. The experiment of human subjects is an ambulatory study where subjects who are already insulin pump users participated, came in to get a CGM, then went home to carry on with their life. The daily routines and carbohydrate intakes of subjects are online reported to respective physicians in real time via the Internet. Taking meal intake as an example, the types, weights, and cooking methods of the ingredients for each meal were reported to their physicians, and then carbohydrate intake was calculated by the hospital's nutritionist. A total of 206 meals were recorded, and the glucose data were sampled every 15 min by the glucose sensor, FreeStyle Libre CGM system (Abbott Diabetes Care). The injection time and doses of insulin were recorded by the insulin pump. The timestep of updating model is 15 min ($T=15\text{min}$).

In this study, the variances of state variables were heuristically adopted to $\mathbf{Q} = \text{diag}[10^{-2}, 10^{-2}, 1, 1, 10^{-6}, 10^{-6}, 10^{-6}, 1, 1]$. The variance of measurement noise was set to $R = 0.45$ [18]. The initial state variables $[x_1, x_2, G, G_I]$ were set $[u_1 \cdot t_I, u_1 \cdot t_I, z_1, z_1]$, where u_1 was the first value of insulin delivery rate (IDR) data and z_1 was the first value of CGM data. The initial values of $[t_I, K]$, were set as $[42, 0.004]$ according to the population values in [23]. The initial values of S_i^* is 294, calculated by $S_i^* = \frac{1000S_i}{M}$, where initial value of S_i is 0.005 [23]. The initial value of τ was set to 16 min [17]. The number of particles in PFG and PFM was set to 1000 after balancing the relationship between the accuracy and the amount of calculation (details please see appendix). It is worth noting that the sensor noise produced by UVa/Padova simulator is close to that in real CGM (FreeStyle Libre, Abbott Diabetes Care), which is complex [49]. For simplicity, this noise is usually considered to be Gaussian approximately, such as in [39] and [46].

To evaluate the performances of OGI model on estimating PGC and PIC, root mean square error (RMSE) is used to measure the accuracy:

$$RMSE = \sqrt{\frac{\sum_{i=1}^m (y_j - \hat{y}_j)^2}{m}} \quad (4.1)$$

One of the outputs is the estimated U , which can be used to identify the state of carbohydrate disturbances [46,47]. Similar to [16], the CGM data, the forward difference of disturbance $D_U(k)$, and previous identification results are used to identify a new carbohydrate intake. Specifically, an active identification of new carbohydrate intake, three conditions must be met at the same time: (1) $D_U(k)$ exceeds the threshold two times successively; (2) measured CGM data goes beyond a threshold (5.56 mmol/L [47]); (3) it has been more than 120 minutes since the last active identification. The indicator 'Flag' takes 1 when carbohydrate intake is identified, otherwise it takes 0. The specific formula of 'Flag' can be given as:

$$Flag(k) = 1 \quad \text{if} \begin{cases} D_U(k), D_U(k-1) \geq \text{Threshold} \\ CGM(k) \geq 5.56 \text{ mmol/L} \\ Flag(k-1, k-2, k-3, k-4) = 0 \end{cases} \quad (4.2)$$

$$Flag(k) = 0 \quad \text{otherwise}$$

Here the ability to identify carbohydrate disturbance within two hours is also evaluated.

For each carbohydrate disturbance, the 120 min counted from the beginning of the meal is defined as an identification test window T_{carbo} , as defined in literature [16,47]. The definitions of true-positive (TP) event, false-negative (FN), false-positive (FP) event, and true-negative (TN) event are given below.

$$\begin{aligned} \text{TP} : \text{Flag}(k) &= 1, \text{ while } k \in T_{carbo} \\ \text{FN} : \text{Flag}(k) &= 0, \text{ while } k \in T_{carbo} \\ \text{FP} : \text{Flag}(k) &= 1, \text{ while } k \notin T_{carbo} \\ \text{TN} : \text{Flag}(k) &= 0, \text{ while } k \notin T_{carbo} \end{aligned}$$

On this basis, four metrics: accuracy rate (AR), precision rate (PR), false-positive rate (FPR) and recall rate (RR) are derived as:

$$\begin{aligned} \text{AR} &= \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}} \\ \text{PR} &= \frac{N_{TP}}{N_{TP} + N_{FP}} \\ \text{FPR} &= \frac{N_{FP}}{N_{TP} + N_{FP}} \\ \text{RR} &= \frac{N_{TP}}{N_{TP} + N_{FN}} \end{aligned} \quad (4.3)$$

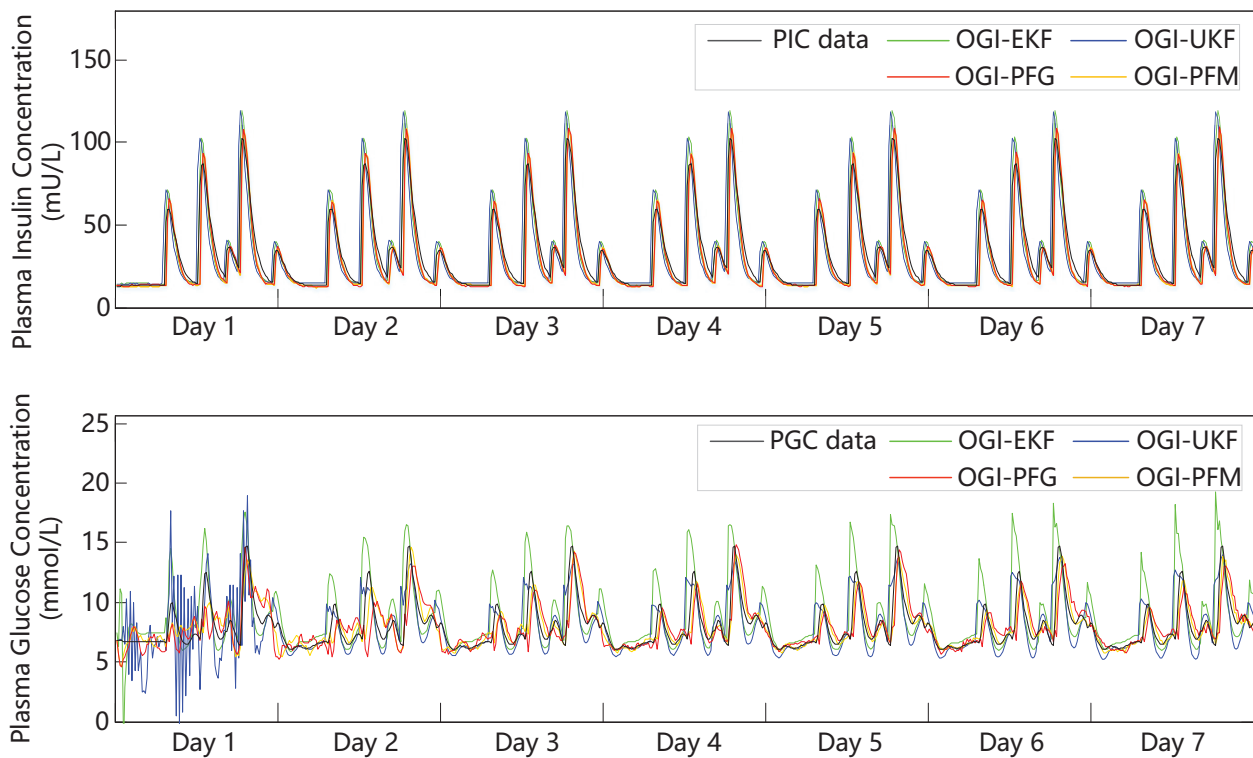


Figure 2. An example of PGC and PIC estimation based on OGI model with different Bayesian filtering estimators (OGI-EKF, OGI-UKF, OGI-PFG, and OGI-PFM). The sampling period is 15min.

5. Results

The performances of OGI model on estimating PGC, PIC and physiological states are evaluated with four Bayesian filtering estimators: EKF, UKF, PFG, and PFM.

5.1. Experiment I: PIC and PGC estimation

Figure 2 shows an example of PIC and PGC estimation for an in-silico subject based on OGI-EKF, OGI-UKF, OGI-PFG, and OGI-PGM respectively. As seen, the results show that the OGI model can estimate the insulin and glucose concentration in plasma with any one type of Bayesian filtering estimator.

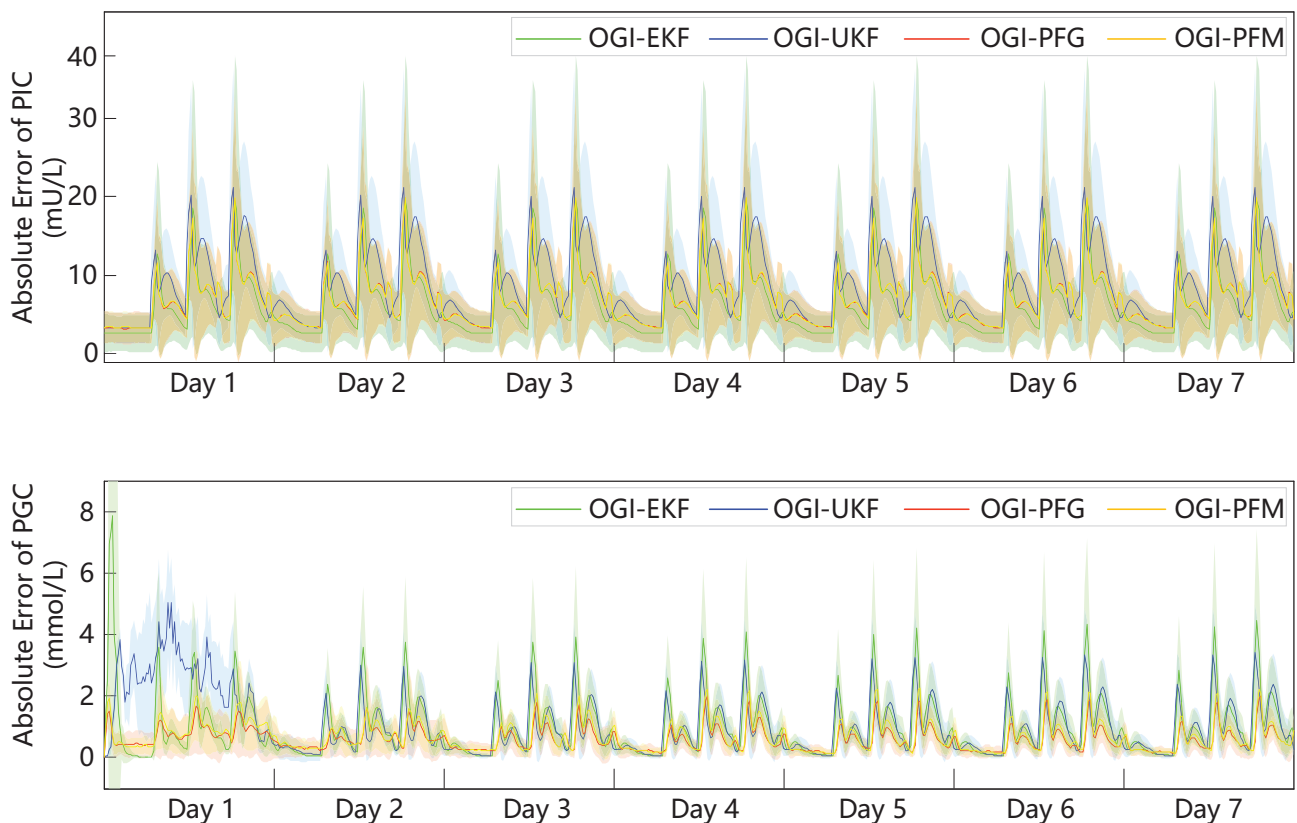


Figure 3. Comparison of different Bayesian filtering estimators on PIC and PGC estimation for in-silico subjects. (The solid lines are the mean values, and the upper and lower boundaries of shadow areas represent the upper and lower standard deviations.)

To quantify the performances, Figure 3 illustrates the absolute errors of different Bayesian filtering estimators on PGC and PIC estimation for all in-silico subjects. The top figure shows the absolute errors of PIC estimation. It is observed that the absolute errors of OGI model stay lower than 20 mU/L for all four estimators. The error profiles of OGI-PFG and OGI-PFM are fairly close to each other, and both are under OGI-EKF and OGI-UKF. The bottom figure illustrates the estimation results of PGC. It is worth noting that similar to the results in [50], the absolute errors in the first few days are larger than those of the other days, especially for OGI-EKF and OGI-UKF. This is because the state variables require a period of training time to converge. Besides, the errors of estimated PIC and PGC

appear periodic changes. For PGC estimation, OGI-EKF and OGI-UKF have larger absolute errors than OGI-PFG and OGI-PFM. The results confirm the advantages of PF technology (PFG and PFM) for PIC and PGC estimation.

Table 3. RMSEs of estimated PGC (mmol/L) and PIG (mU/L) for different groups of in-silico subjects (mean±standard deviation).

Model	Project	Adults	Adolescents	Children	All
OGI-EKF	PGC	1.08±0.47	1.59±1.00	1.74±0.63	1.47±0.76
	PIC	8.69±3.03	11.94±5.71	13.78±2.61	11.47±4.43
OGI-UKF	PGC	0.94±0.22	1.32±0.58	1.62±0.53	1.29±0.53
	PIC	7.91±3.40	10.94±6.50	12.62±3.70	10.49±4.99
OGI-PFG	PGC	0.75±0.16	1.05±0.27	0.95±0.15	0.92±0.23
	PIC	7.08±2.83	9.56±4.73	11.82±1.99	9.49±3.81
OGI-PFM	PGC	0.73±0.10	0.96±0.17	0.99±0.18	0.89±0.19
	PIC	7.08±2.83	9.56±4.42	11.82±2.01	9.49±3.81

For different groups of in-silico subjects (adults, adolescents, and children), the RMSEs of estimated PGC and PIC on each age group are given in Table 3. The results of each group are not only confirmed the advantages of PFG and PFM for the estimation again, but also interestingly found that OGI model performs better for the adult group.

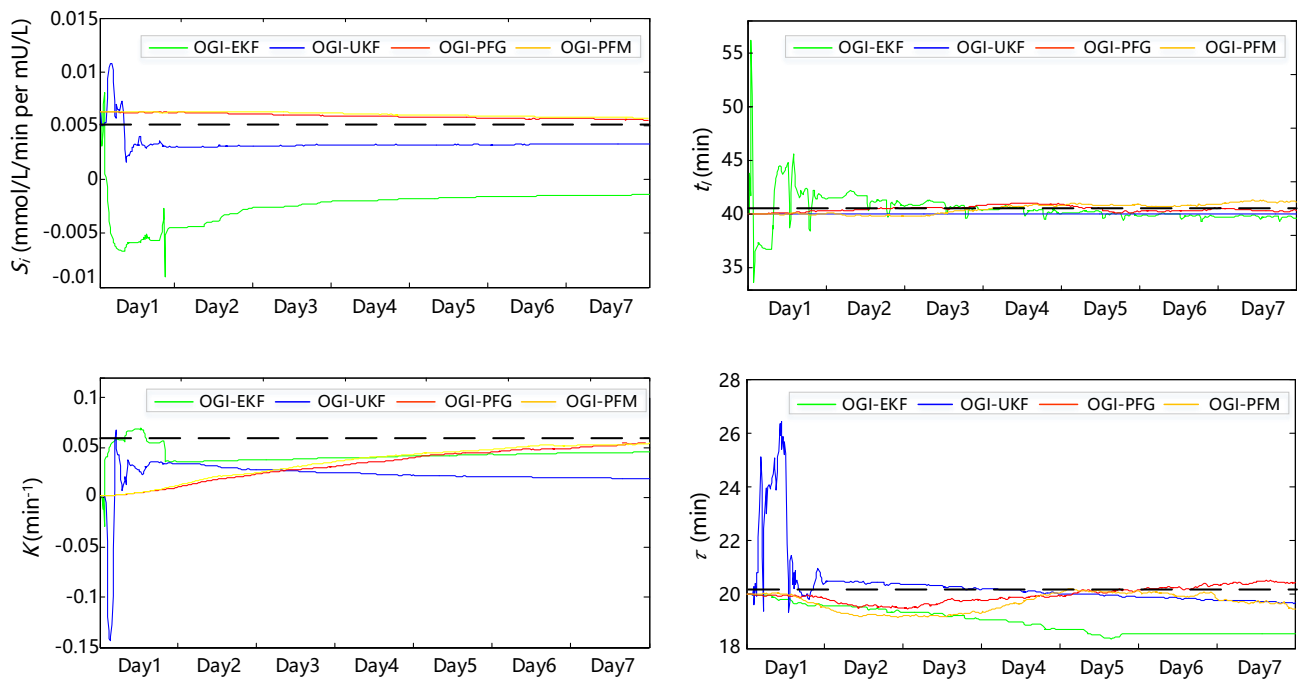


Figure 4. An example of parameter evolution over 7 days. The sampling period is 15min.

5.2. Experiment II: Identifying physiological parameters

Due to the difference in model structure, the parameters in UVa/Padova simulator are different from those in OGI model. As human subject based study, obtaining the ground truth from in-silico subjects

is challenging. Instead of providing ground truth as a prior, we decide to take a sensitivity analysis approach, that is, we interrogate the model performances on the data generated by giving parameter settings. A glucose dataset is generated by giving model parameters, meal intake and insulin delivery data. Glucose data and insulin delivery data, as the input, are used to estimate parameters. An example of parameter evolution over 7 days is shown in Figure 4. The black dashed line is the true value of the parameter.

It can be seen that the estimated values converge to the true values, especially after the warm-up period. Although the changes of all parameters estimations under the four filters are smoothed over time, it is observed that the results from OGI-UKF and OGI-EKF fluctuate greatly in the early stage, especially on the first day due to the warm-up period. This phenomenon is consistent with the results in [50]. In contrast, the performance of OGI-PFG and OGI-PFM in this warm-up period is more stable.

The parameter estimations for in-silico subjects with mean and standard deviation are shown in Table 4. The results are average estimation across 30 in-silico subjects throughout 7 days.

The parameter estimations for in-silico subjects with population mean and standard deviation are shown in Table 4. It is worth noting that the in-silico subject are closer to the real environment, the nonlinearity and uncertainty are stronger, and the performances of different filters have remarkable differences. This further reflects the advantages of the designed particle filter.

Table 4. Parameter identification for 30 in-silico subjects over 7 days (mean $\pm$ standard deviation).

Model	t_I	$S_i(\times 10^{-2})$	$K(\times 10^{-2})$	τ
OGI-EKF	41.46 $\pm$ 7.70	0.23 $\pm$ 0.83	4.58 $\pm$ 2.93	21.85 $\pm$ 14.44
OGI-UKF	42.22 $\pm$ 7.30	0.63 $\pm$ 0.60	4.79 $\pm$ 4.36	32.67 $\pm$ 12.89
OGI-PFG	41.83 $\pm$ 6.64	0.39 $\pm$ 0.24	6.36 $\pm$ 3.76	29.19 $\pm$ 8.05
OGI-PFM	41.83 $\pm$ 6.63	0.41 $\pm$ 0.22	6.41 $\pm$ 3.21	28.38 $\pm$ 6.98

Table 5. An example of the meal time and sizes for a human subject.

Day	Time	CHO(g)	Day	Time	CHO(g)
1	7:15	20	4	8:00	36
	11:15	81.7		14:00	53
	18:15	50		18:00	38
2	7:30	38	5	6:00	20
	12:00	80		11:30	85
	19:20	40		17:30	38
	0:30	35	6	7:00	37
3	7:00	35.4		12:40	69
	13:30	73		19:10	33.4
	19:00	34	7	12:30	85
	22:00	37		20:00	24.7

5.3. Experiment III: Identifying carbohydrate disturbance

In this experiment, given that meal intake can be known as a ground truth, we decide to use this information to demonstrate that the parameter U can be identified and has potential in real-world application such as meal identification. As for the performances of OGI model with different estimators on identifying compound uncertainty U , an example of meal identification from a human subject is illustrated in Table 5 and Figure 5 (the meal times and sizes of the subject are summarized in the table and identification results are shown in the figure). The top figure shows the glucose level given by CGM, while the bottom figure illustrates the estimated U and the identification results of carbohydrate disturbances.

As seen, OGI-PFG and OGI-PFM are able to identify most of the carbohydrate disturbances. The performances on meal identification for all the subjects, including the in-silico and human participants, are shown in Table 6 (with the threshold heuristically set to 0.1 mmol/L/min). It shows that the performances from OGI-PFG and OGI-PFM are similar which outperform OGI-EFK and OGI-UKF. The OGI-PFM shows the daily ability with AR 95.46% $\pm$ 0.65%, PR 68.89% $\pm$ 3.57%, FPR 31.11% $\pm$ 3.57%, and RR 71.90% $\pm$ 3.72%.

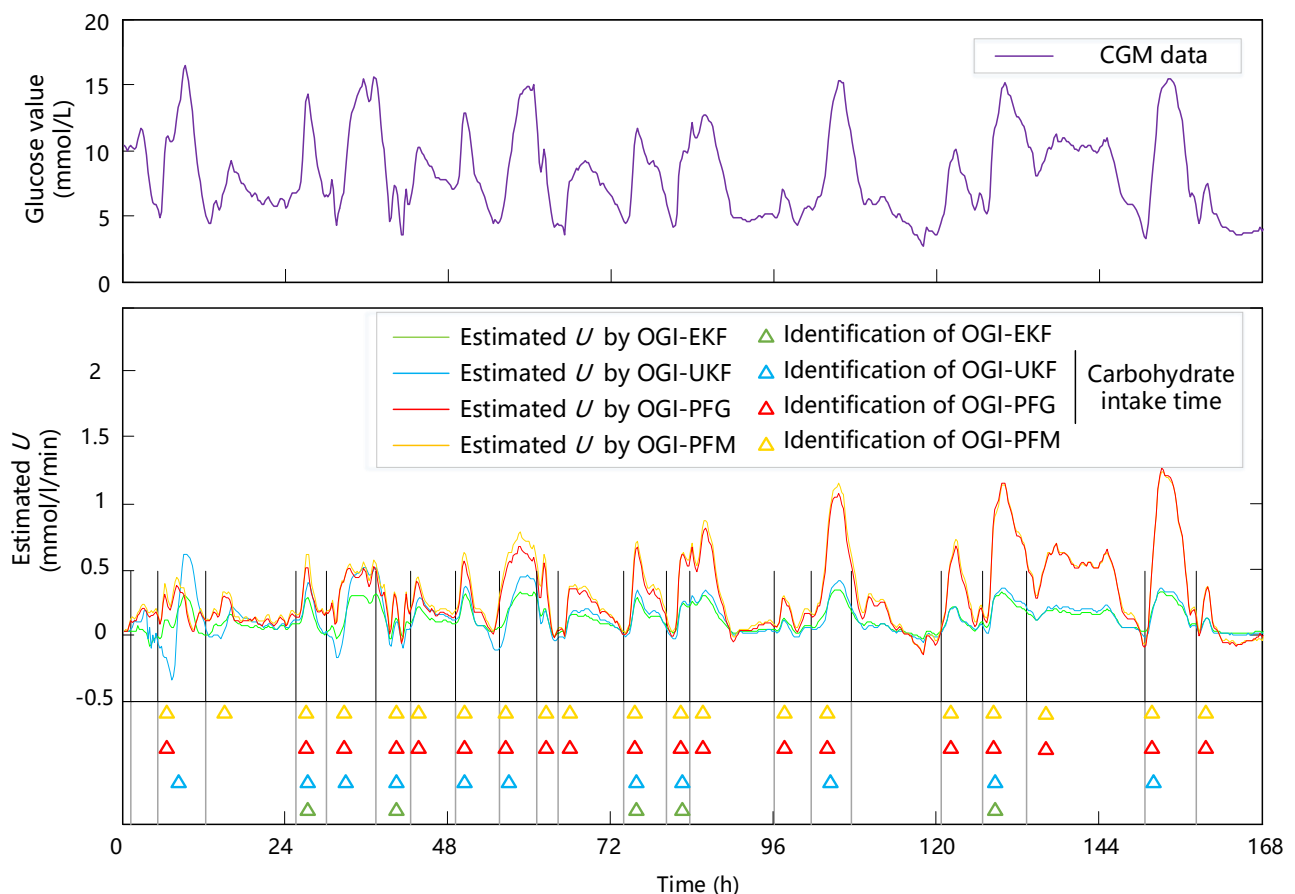


Figure 5. An example of identifying meal disturbance for a human subject. The sampling period is 15min.

Table 6. The daily averages and standard deviation of OGI model with different estimators on meal identification.

Model	AR	PR	RR	FPR
OGI-EKF	92.10% $\pm$ 0.79%	46.35% $\pm$ 4.10%	34.38% $\pm$ 2.77%	53.65% $\pm$ 4.10%
OGI-UKF	93.75% $\pm$ 1.45%	58.74% $\pm$ 8.73%	55.90% $\pm$ 9.37%	41.26% $\pm$ 8.73%
OGI-PFG	95.11% $\pm$ 0.74%	66.62% $\pm$ 4.69%	70.19% $\pm$ 3.10%	33.38% $\pm$ 4.69%
OGI-PFM	95.46% $\pm$ 0.65%	68.89% $\pm$ 3.57%	71.90% $\pm$ 3.72%	31.11% $\pm$ 3.57%

6. Discussion

It is a benefit for a fully automatic AP system to estimate the PGC and PIC based on the collected CGM and IDR data. However, the inter-individual variability limits the accuracy of estimation, because they increase the uncertainty of model parameters from the mathematical perspective. To solve this problem, it is feasible to extend the uncertainty parameters into state variables and estimate them based on Bayesian filtering estimator. But this method can be adopted only if the extended variables are completely observable. This research proposes an observable glucose-insulin (OGI) model in which the model parameters and environment disturbance can be extended to state variables and the observability of the extended system is proved. On this basis, four Bayesian filtering estimators (EKF, UKF, PFG, and PFM) are designed to estimate all the state variables and the PGC and PIC.

It is observed from the experiment of in-silico subjects that the proposed model has high accuracy in PIC and PGC estimation (Figure 2 and Figure 3), and the OGI-PFM outperforms other estimators in most cases. According to the results of different groups, it is found that the OGI model tends to perform better for the older subjects (Table 3). Based on the estimation of physiological states, the OGI model has the potential in identifying parameters and disturbance (Figure 4, Figure 5, Table 4 and Table 6).

In looking at the estimators, the EKF is less effective than the UKF in PIC estimation, which is consistent with the results in [21]. An interesting observation is that PF (PFG, PFM) has better performances than EKF and UKF indicating the promises of PF based estimator, which is not consistent with the conclusion drawn from [39]. This may be because when more model parameters are extended to state variables, the nonlinearity of the model is further enhanced. In this case, the advantages of PF in dealing with nonlinear systems are fully utilized. Comparing PFG vs. PFM, while PFM has the best performance among the four, it is surprising that the outperformance of PFM over PFG is marginal. It is contended that this may be due to the limited dataset where the mixed noises are not apparent. This discovery indeed points out the immediate future effort which is to collect PIC and PGC data from human participants, to conduct the validation on the larger dataset, and to explore the influence of other noise distributions on the estimation.

Although PFG and PFM have better performance in estimating PIC and PGC, PFG and PFM rely on powerful computing power. The simulation is implemented in the computer (Intel(R) Core i9-9900 K 2-Core 3.60 GHz processor). It is worth noting that in real application scenarios, these are not high-performance computation devices but embedded devices with significant constraints in processing power, memory, and power usage. Therefore, it is essential to develop a low-computational and low-power mode in future work, as discussed in [51]. For example, if the computational cost needs to be

reduced due to the limitation of computing power and a little accuracy is allowed to be sacrificed, some parameters with low sensitivity can be fixed.

Table 7. Comparative results with other method in literature on PIC estimation.

Study	Model	Personalized parameter	Estimator	Results (mU/L)
Iman Hajizadeh [21]	Hovorka model	t_I, k_e, U	EKF	RMSE(PIC): 19
			UKF	RMSE(PIC): 13.14
This study	OGI model	t_I, S_i^*, K, τ, U	PFM	RMSE(PIC): 9.49

To the best of our knowledge, research to date has mainly focused on PIC estimation, in which the variability of insulin absorption parameters was focused. Few works of literature discussed the estimation accuracy of PGC. Thus, OGI is compared with the models from the literature on PIC estimations only. Since the objective of the study in [21] aligns with our study, the comparison of our proposed OGI model with the methods in [21] is conducted using our dataset. As seen in Table 7, using the method mentioned in [21], the RMSEs of PIC estimation enhanced by EKF and UKF are 16.64 mU/L and 13.28 mU/L respectively, which is larger than that conducted by OGI-PFM 9.49 mU/L. It can be concluded that OGI-PFM has better performances on our dataset.

In parameter identification, two types of parameters are involved in the model, physiological parameters $[t_I, S_i^*, K, \tau]$ and parameter of compound uncertainty U . Given that (1) in the experiment on in-silico subjects produced by the UVa/Padova simulator, physiological parameters are fixed as in [16, 46]; (2) in human subjects, although the physiological parameters changes all the time, the specific changes of these parameters cannot be continuously monitored, the ability to dynamically identify variable parameters needs to be verified in future work. Nevertheless, since (1) meal intake can induce the U changes; (2) the time and size of meal intake can be set in UVa/Padova Simulator and are also recorded in the experiment of human subject, meal identification is developed to be an application to show the ability to dynamically identify the parameter U . Please note that the focus of this study is to develop a complete observable model. Meal identification is just a preliminary attempt to show the potential of identifying the parameter impacted by external disturbance (carbohydrate intake). It is the author's future work to improve meal identification based on estimated U , such as adding post-processing and choosing an appropriate threshold. Developing a fast and effective way to detect meal intake is worthy of further research.

When the cannula site is no longer absorbing the insulin effectively or even occlusion, the insulin delivery rate displayed in the insulin pump cannot achieve the desired hypoglycemic effect. It is guessed that the parameters which have positive effects on rising glucose level may be overestimated, such as $[t_I, U]$, while the parameters which have negative effects may be underestimated, such as S_i . This may also bring new challenges to meal identification. Exploring the performance of the proposed method in facing such faults is very valuable. Taking fault diagnosis as an example, it may be warning that the cannula site is no longer absorbing the insulin effectively or even occlusion, when the parameter t_I is overestimated (exceeding a certain threshold). In addition, it is also worth exploring how to ensure the robustness of the system in the case of sensor data loss.

7. Conclusion

In this study, an OGI dynamic model is proposed for personalized estimation of PIC and PGC based on CGM and IDR data. The OGI model, which describes the dynamics of glucose and insulin changes, takes full account of the parameter uncertainty caused by inter-individual variability. To deal with the uncertainty, the OGI model is extended to a non-parametric dynamic model, of which all the state variables, including the original ones and those extended from the model parameters, are observable. On this basis, four types of Bayesian filtering estimators are designed to estimate the state variables, including EKF, UKF, PFG, and PFM. For EKF, UKF and PFG, the Gaussian assumption is adopted, but considering the non-negativity of some state variables, PFM explores the non-Gaussian characteristic of process noise and assuming it to be subjected to a distribution mixed of Gaussian and shifted log-normal distribution. Tested by the in-silico experiment, OGI-PFG and OGI-PFM show better abilities to estimate PIC and PGC than OGI-EKF and OGI-UKF. Especially, the OGI-PFM has a particularly promising performance (RMSE of PIC estimation: 9.49 ± 3.81 mU/L, RMSE of PGC estimation: 0.89 ± 0.19 mmol/L). It also indicates that particle filter has advantages in personalizing the parameters and disturbances. It is also found that compared with the influence from filtering types, the distribution of noise has much less influence on the estimation performances. Moreover, the personalized estimations of PIC and PGC for clinic application remains to be verified.

Acknowledgments

The authors acknowledge funding support from Beijing Advanced Innovation Center for Big Data-based Precision Medicine. Authors also recognize support by Department of Endocrinology, Xuanwu Hospital Capital Medical University, China, for clinical suggestions, and efforts by Department of Endocrinology, The First Peoples Hospital of Yangquan City, China for recruitment of subjects and collection of physiological data.

Conflict of interest

The authors declare no conflict of interest.

References

1. M. A. Atkinson, G. S. Eisenbarth, A. W. Michels, Type 1 diabetes, *The Lancet*, **383** (2014), 69–82.
2. A. D. Association, Diagnosis and classification of diabetes mellitus, *Diabetes Care*, **33** (2010), S62–S69.
3. C. Cobelli, C. Dalla Man, G. Sparacino, L. Magni, G. De Nicolao, B. P. Kovatchev, Diabetes: Models, signals, and control, *IEEE Rev. Biomed. Eng.*, **2** (2009), 54–96.
4. T. Danne, R. Nimri, T. Battelino, R. M. Bergenstal, K. L. Close, et al., International consensus on use of continuous glucose monitoring, *Diabetes Care*, **40** (2017), 1631–1640.
5. S. V. Edelman, N. B. Argento, J. Pettus, I. B. Hirsch, Clinical implications of real-time and intermittently scanned continuous glucose monitoring, *Diabetes Care*, **41** (2018), 2265–2274.

6. L. Heinemann, G. A. Fleming, J. R. Petrie, R. W. Holl, R. M. Bergenstal, A. L. Peters, Insulin pump risks and benefits: a clinical appraisal of pump safety standards, adverse event reporting, and research needs, *Diabetes Care*, **38** (2015), 716–722.
7. S. Brown, D. Raghinaru, E. Emory, B. Kovatchev, First look at Control-IQ: a new-generation automated insulin delivery system, *Diabetes Care*, **41** (2018), 2634–2636.
8. E. Atlas, R. Nimri, S. Miller, E. A. Grunberg, M. Phillip, MD-logic artificial pancreas system: a pilot study in adults with type 1 diabetes, *Diabetes Care*, **33** (2010), 1072–1076.
9. K. Turksoy, E. Littlejohn, A. Cinar, Multimodule, multivariable artificial pancreas for patients with type 1 diabetes: Regulating glucose concentration under challenging conditions, *IEEE Control Syst. Mag.*, **38** (2018), 105–124.
10. S. O'Neill, Update on technologies, medicines and treatments, *Diabet. Med. J. British Diabetes Assoc.*, **37** (2020), 709–711.
11. G. E. Umpierrez, D. C. Klonoff, Diabetes technology update: Use of insulin pumps and continuous glucose monitoring in the hospital, *Diabetes Care*, **41** (2018), 1579–1589.
12. J. Bondia, S. Romero-Vivo, B. Ricarte, J. L. Diez, Insulin estimation and prediction: A review of the estimation and prediction of subcutaneous insulin pharmacokinetics in closed-loop glucose control, *IEEE Control Syst. Mag.*, **38** (2018), 47–66.
13. J. Pickup, H. Keen, Continuous subcutaneous insulin infusion at 25 years: evidence base for the expanding use of insulin pump therapy in type 1 diabetes, *Diabetes Care*, **25** (2002), 593–598.
14. I. Hajizadeh, M. Rashid, S. Samadi, M. Sevil, A. Cinar, Adaptive personalized multivariable artificial pancreas using plasma insulin estimates, *J. Process Control*, **80** (2019), 26–40.
15. C. Neatpisarnvanit, J. R. Boston, Estimation of plasma insulin from plasma glucose, *IEEE Transact. Biomed. Eng.*, **49** (2002), 1253–1259.
16. C. M. Ramkissoon, P. Herrero, J. Bondia, J. Vehi, Unannounced meals in the artificial pancreas: Detection using continuous glucose monitoring, *Sensors*, **18** (2018), 884–901.
17. L. O. Avila, M. De Paula, C. R. Sanchez-Reinoso, Estimation of plasma insulin concentration under glycemic variability using nonlinear filtering techniques, *Biosystems*, **171** (2018), 1–9.
18. 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, *Computer Methods Biomech. Biomed. Eng.*, **19** (2016), 934–942.
19. K. Turksoy, I. Hajizadeh, S. Samadi, J. Feng, M. Sevil, et al., Real-time insulin bolusing for unannounced meals with artificial pancreas, *Control Eng. Pract.*, **59** (2017), 159–164.
20. A. C. Charalampidis, K. Pontikis, G. D. Mitsis, G. Dimitriadis, V. Lampadiari, et al., 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.
21. I. Hajizadeh, M. Rashid, K. Turksoy, S. Samadi, J. Feng, et al., Plasma insulin estimation in people with type 1 diabetes mellitus, *Indust. Eng. Chem. Res.*, **56** (2017), 9846–9857.
22. J. Ko, D. Fox, GP-BayesFilters: Bayesian filtering using Gaussian process prediction and observation models, *Auton. Robot.*, **27** (2009), 75–90.

23. Y. Ruan, M. E. Wilinska, H. Thabit, R. Hovorka, Modeling day-to-day variability of glucose–insulin regulation over 12-week home use of closed-loop insulin delivery, *IEEE Transact. Biomed. Eng.*, **64** (2016), 1412–1419.
24. W. Wang, S. Wang, X. Wang, D. Liu, Y. Geng, T. Wu, A glucose–insulin mixture model and application to short-term hypoglycemia prediction in the night time, *IEEE Transact. Biomed. Eng.*, **68** (2021), 834–845.
25. W. Liu, A mathematical model for the robust blood glucose tracking, *Math. Biosci. Eng.*, **16** (2019), 759–781.
26. K. Fessel, J. B. Gaither, J. K. Bower, T. Gaillard, K. Osei, G. A. Rempała, Mathematical analysis of a model for glucose regulation, *Math. Biosci. Eng.*, **13** (2016), 83–99.
27. H. F. Arguedas, M. A. Capistran, Bayesian analysis of glucose dynamics during the oral glucose tolerance test, *Math. Biosci. Eng.*, **18** (2021), 4628–4647.
28. X. Shi, Q. Zheng, J. Yao, J. Li, X. Zhou, Analysis of a stochastic IVGTT glucose–insulin model with time delay, *Math. Biosci. Eng.*, **17** (2020), 2310–2329.
29. N. Chopin, A sequential particle filter method for static models, *Biometrika*, **89** (2002), 539–552.
30. J. Wei, G. Dong, Z. Chen, Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression, *IEEE Transact. Indust. Elect.*, **65** (2017), 5634–5643.
31. M. S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Transact. Signal Process.*, **50** (2002), 174–188.
32. R. N. Bergman, L. S. Phillips, C. Cobelli, Physiologic evaluation of factors controlling glucose tolerance in man: measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose, *J. Clin. Invest.*, **68** (1981), 1456–1467.
33. A. De Gaetano, O. Arino, Mathematical modelling of the intravenous glucose tolerance test, *J. Math. Biol.*, **40** (2000), 136–168.
34. E. Horová, J. Mazoch, J. Hiłgertová, J. Kvasnička, J. Škrha, et al., Acute hyperglycemia does not impair microvascular reactivity and endothelial function during hyperinsulinemic isoglycemic and hyperglycemic clamp in type 1 diabetic patients, *Exper. Diabetes Res.*, **2012**, 59–66.
35. N. Magdelaine, L. Chaillous, I. Guilhem, J.-Y. Poirier, M. Krempf, et al., A long-term model of the glucose–insulin dynamics of type 1 diabetes, *IEEE Transact. Biomed. Eng.*, **62** (2015), 1546–1552.
36. R. Hermann, A. Krener, Nonlinear controllability and observability, *IEEE Transact. Autom. Control*, **22** (1977), 728–740.
37. Q. Ma, Structural conditions on observability of nonlinear systems, *Int. J. Inform. Technol. Computer Sci.*, **3** (2011), 16–22.
38. C. Yardim, Z. H. Michalopoulou, P. Gerstoft, An overview of sequential Bayesian filtering in ocean acoustics, *IEEE J. Ocean. Eng.*, **36** (2011), 71–89.
39. C. Eberle, C. Ament, The Unscented Kalman Filter estimates the plasma insulin from glucose measurement, *Biosystems*, **103** (2011), 67–72.

40. T. Li, M. Bolic, P. M. Djuric, Resampling methods for particle filtering: classification, implementation, and strategies, *IEEE Signal Proc. Mag.*, **32** (2015), 70–86.
41. B. M. Hill, The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic, *J. Am. Stat. Assoc.*, **58** (1963), 72–84.
42. C. D. Man, F. Micheletto, D. Lv, M. Breton, B. Kovatchev, C. Cobelli, The UVA/PADOVA type 1 diabetes simulator: New features, *J. Diabetes Sci. Technol.*, **8** (2014), 26–34.
43. R. Visentin, C. Dalla Man, C. Cobelli, One-day Bayesian cloning of type 1 diabetes subjects: toward a single-day UVA/Padova type 1 diabetes simulator, *IEEE Transact. Biomed. Eng.*, **63** (2016), 2416–2424.
44. B. P. Kovatchev, M. Breton, C. D. Man, C. Cobelli, In silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes, *J. Diabetes Sci. Technol.*, **3** (2009), 44–55.
45. C. Cobelli, C. D. Man, M. G. Pedersen, A. Bertoldo, Advancing our understanding of the glucose system via modeling: A perspective, *IEEE Transact. Biomed. Eng.*, **61** (2014), 1577–1592.
46. S. Samadi, K. Turksoy, I. Hajizadeh, J. Feng, M. Sevil, A. Cinar, Meal detection and carbohydrate estimation using continuous glucose sensor data, *IEEE J. Biomed. Health Inform.*, **21** (2017), 619–627.
47. K. Turksoy, S. Samadi, J. Feng, E. Littlejohn, L. Quinn, A. Cinar, Meal detection in patients with type 1 diabetes: a new module for the multivariable adaptive artificial pancreas control system, *IEEE J. Biomed. Health Inform.*, **20** (2015), 47–54.
48. M. Vettoretti, A. Facchinetti, G. Sparacino, C. Cobelli, Type-1 diabetes patient decision simulator for in silico testing safety and effectiveness of insulin treatments, *IEEE Transact. Biomed. Eng.*, **65** (2017), 1281–1290.
49. M. Breton, B. Kovatchev, Analysis, modeling, and simulation of the accuracy of continuous glucose sensors, *J. Diabetes Sci. Technol.*, **2** (2008), 853–862.
50. D. J. Albers, M. Levine, B. Gluckman, H. Ginsberg, G. Hripcsak, L. Mamykina, Personalized glucose forecasting for type 2 diabetes using data assimilation, *PLoS Comput. Biol.*, **13** (2017), 1–38.
51. F. Rahmanian, M. Dehghani, P. Karimaghaee, M. Mohammadi, R. Abolpour, Hardware-in-the-loop control of glucose in diabetic patients based on nonlinear time-varying blood glucose model, *Biomed. Signal Process. Control*, **66** (2021), 1–12.

Appendix

Algorithm of EKF

The pseudocode of the EKF algorithm is presented in Algorithm 1.

Algorithm 1 Algorithm of EKF

- 1: Input: $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{X}_0$, z
 - 2: Initialization: $\mathbf{P}_1 = \mathbf{P}_0$, $\mathbf{X}_1 = \mathbf{X}_0$, where $\mathbf{P}$ is covariance matrix, $\mathbf{P}_0$ is set as unit matrix.
 - 3: **for** $k = 2 : t$ **do**
 - 4: $\mathbf{X}'_k = F(\hat{\mathbf{X}}_{k-1}, u_k)$, $z'_k = H(\mathbf{X}'_k)$
 - 5: Covariance matrix prediction:
 $\mathbf{P}'_k = \mathbf{\Phi}_k \mathbf{P}_{k-1} \mathbf{\Phi}_k^T + \mathbf{Q}$, where $\mathbf{\Phi}_k = \frac{\partial F(\mathbf{X}'_k)}{\partial \mathbf{X}'_k}$
 - 6: Kalman gain:
 $\mathbf{K}_k = \mathbf{P}'_k \boldsymbol{\varphi}_k^T (\boldsymbol{\varphi}_k \mathbf{P}'_k \boldsymbol{\varphi}_k^T + \mathbf{R})^{-1}$
 where $\boldsymbol{\varphi}_k = \frac{\partial H(\mathbf{X})}{\partial \mathbf{X}}$
 - 7: Covariance matrix update:
 $\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \boldsymbol{\varphi}_k) \mathbf{P}'_k$
 - 8: State variability update:
 $\hat{\mathbf{X}}_k = \mathbf{X}'_k + \mathbf{K}_k (z_k - z'_k)$
 - 9: **end for**
-

Algorithm 2 Algorithm of UKF

-
- 1: Input: $\mathbf{Q}, \mathbf{R}, \mathbf{X}_0, z$
 - 2: Initialization: $\mathbf{P}_1 = \mathbf{P}_0, \mathbf{X}_1 = \mathbf{X}_0, n = 10, \lambda = -7, \alpha = 1, \beta = 2$
 - 3: where $\mathbf{P}$ is covariance matrix, $\mathbf{P}_0$ is set as unit matrix, λ is a scaling parameter, α determines the spread of the sigma points, β is used to incorporate prior knowledge of the distribution.
 - 4: **for** $k = 2 : t$ **do**
 - 5: $2n+1$ sigma points and associated scalar weights:
$$\begin{aligned} \mathbf{X}_{k-1}^{(0)} &= \hat{\mathbf{X}}_{k-1} \\ \mathbf{X}_{k-1}^{(i)} &= \hat{\mathbf{X}}_{k-1} + \left(\sqrt{(n+\lambda)\mathbf{P}_{k-1}} \right)_i, \quad i = 1, \dots, n \\ \mathbf{X}_{k-1}^{(i)} &= \hat{\mathbf{X}}_{k-1} - \left(\sqrt{(n+\lambda)\mathbf{P}_{k-1}} \right)_i, \quad i = n+1, \dots, 2n \end{aligned}$$

$$\begin{aligned} w_m^{(0)} &= \frac{\lambda}{n+\lambda} \\ w_c^{(0)} &= \frac{\lambda}{n+\lambda} + (1 - \alpha^2 + \beta) \\ w_m^{(i)} &= w_c^{(i)} = \frac{\lambda}{2(n+\lambda)}, \quad i = 1, \dots, 2n \end{aligned}$$

where $(\sqrt{\mathbf{A}})_i$ is the i th column of the matrix square root, i.e., lower triangle Cholesky factorization.
 - 6: Sigma points prediction:
$$\mathbf{X}'_k = F(\mathbf{X}_{k-1}^{(i)}, u_k), z'_k = H(\mathbf{X}'_k)$$
 - 7: State Prediction:
$$\begin{aligned} \mathbf{X}'_k &= \sum_{i=0}^n w^{(i)} \mathbf{X}'_k \\ \mathbf{P}'_k &= \sum_{i=0}^{2n} w^{(i)} [\mathbf{X}'_k - \mathbf{X}'_k] [\mathbf{X}'_k - \mathbf{X}'_k]^T + \mathbf{Q} \end{aligned}$$
 - 8: UT transformation:
$$\begin{aligned} \mathbf{X}''_k^{(0)} &= \mathbf{X}'_k \\ \mathbf{X}''_k^{(i)} &= \mathbf{X}'_k + \left(\sqrt{(n+\lambda)\mathbf{P}'_k} \right)_i, \quad i = 1, \dots, n \\ \mathbf{X}''_k^{(i)} &= \mathbf{X}'_k - \left(\sqrt{(n+\lambda)\mathbf{P}'_k} \right)_i, \quad i = n+1, \dots, 2n \end{aligned}$$
 - 9: Output prediction:
$$\begin{aligned} z_k^{(i)} &= H(\mathbf{X}''_k^{(i)}) \\ \hat{z}_k &= \sum_{i=0}^{2n} w^{(i)} z_k^{(i)} \end{aligned}$$
 - 10: Covariance matrix prediction:
$$\begin{aligned} \mathbf{S}_k &= \sum_{i=0}^{2n} w^{(i)} (z_k^{(i)} - \hat{z}_k)(z_k^{(i)} - \hat{z}_k)^T + \mathbf{R} \\ \boldsymbol{\psi}_k &= \sum_{i=0}^{2n} w^{(i)} (\mathbf{X}'_k - \mathbf{X}'_k)(z_k^{(i)} - \hat{z}_k)^T \end{aligned}$$
 - 11: Kalman gain: $\mathbf{K}_k = \boldsymbol{\psi}_k \mathbf{S}_k^{-1}$
 - 12: Covariance matrix update: $\mathbf{P}_k = \mathbf{P}'_k - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^T$
 - 13: State update: $\hat{\mathbf{X}}_k = \mathbf{X}'_k + \mathbf{K}_k (z_k - z'_k)$
 - 14: **end for**

Algorithm of PFG

The pseudocode of the PFG algorithm is presented in Algorithm 3.

Algorithm 3 Algorithm of PFG

-
- 1: Input: $\mathbf{Q}, R, \mathbf{X}_0, z$
 - 2: Initialization: $\mathbf{X}_1 = \mathbf{X}_0$
 N initial particles production for each variable $X \in \mathbf{X}$:
 $X_1^i \sim N(X_0, Q_X), i = 1 : N$
 - 3: **for** $k = 2 : t$ **do**
 - 4: State prediction:
 $X_k'^i \sim N(F(X_{k-1}^i, u_k), Q)$
 $z_k'^i = H(X_k^i)$
 - 5: Weight calculation:
 $\tilde{W}_k^i = \frac{W_k^i}{\sum_{i=1}^N W_k^i}$
 - 6: Resampling $\hat{X}_k^i$ according to the $\tilde{W}_k^{(N)}$, and weight reset: $W_k^i = \frac{1}{N}$
 - 7: $\hat{X}_k \approx \frac{1}{N} \sum_{i=1}^N \delta(\hat{X}_k^i)$
 - 8: **end for**
-

Algorithm of PFM

The pseudocode of the PFM algorithm is presented in Algorithm 4.

RMSEs of PGC and PIC estimation with various particle numbers

We have presented the RMSEs of PGC and PIC estimation with various particle numbers in the Table 8 and added explanation on it.

Table 8. RMSEs of PGC and PIC estimation with various particle numbers.

The Number of particles	200	500	800	1000	2000	5000
Time cost (s)	27	81	268	402	886	2623
RMSE of PGC (mmol/L)	1.47	1.23	0.98	0.89	0.88	0.86
RMSE of PIC (mU/L)	9.95	9.77	9.61	9.49	9.47	9.45

As expected, we can see that as the number of particles increases, the accuracy gradually increases. Nevertheless, when the number of particles exceeds 1000, the increase of particles has less and less influence on the improvement of estimation accuracy, and the corresponding calculation cost is heavier and heavier. Considering the limited computing power of the artificial pancreas, we balanced the relationship between the accuracy and the amount of calculation. For this reason, we set the number of particles to 1000.

Algorithm 4 Algorithm of PFM

- 1: Input: Q, R, X_0, z
- 2: Initialization: $X_1 = X_0$
Hyperparameter computation:

$$\lambda_X^1 = X_0$$

$$\mu_X^1 = \ln(X_0) - \frac{1}{2} \ln\left(1 + \frac{Q_X}{X_0^2}\right)$$

$$\sigma_X^1 = \sqrt{\ln\left(1 + \frac{Q_X}{X_0^2}\right)}$$
 N initial particles production for each non-negativity variability: $X_1^i \sim GLN(\lambda_X^1, \mu_X^1, \sigma_X^1)$, and for variability U : $U_1^i \sim N(0, Q_U)$.
 Note that the X represents each non-negativity variability, i.e., all parameters except U .
- 3: **for** $k = 2 : t$ **do**
- 4: Hyperparameter computation:

$$\lambda_X^k = f(X_{k-1})$$

$$\mu_X^k = \ln(f(X_{k-1})) - \frac{1}{2} \ln\left(1 + \frac{Q_X}{f(X_{k-1})^2}\right)$$

$$\sigma_X^k = \sqrt{\ln\left(1 + \frac{Q_X}{f(X_{k-1})^2}\right)}$$
- 5: State prediction:

$$X_k^i \sim GLN(\lambda_X^k, \mu_X^k, \sigma_X^k)$$

$$U_k^i \sim N(U_{k-1}^i, Q_U)$$

$$z_k^i = H(X_k^i)$$
- 6: Weight calculation:

$$W_k^i = \frac{1}{\sqrt{R}} \exp\left(-\frac{(z_k - z_k^i)^2}{2R}\right)$$
- 7: Weight normalization:

$$\tilde{W}_k^i = \frac{W_k^i}{\sum_{i=1}^N W_k^i}$$
- 8: Resampling $\hat{X}_k^i$ according to the $\tilde{W}_k^{(N)}$, and weight reset: $W_k^i = \frac{1}{N}$
- 9: $\hat{X}_k \approx \frac{1}{N} \sum_{i=1}^N \delta(\hat{X}_k^i)$
- 10: **end for**



AIMS Press

© 2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)