17

Glucose Prediction in Type 1 and Type 2 Diabetic Patients Using Data Driven Techniques

Eleni I. Georga[1,2], Vasilios C. Protopappas[2] and Dimitrios I. Fotiadis[1]
[1]*Department of Materials Science and Engineering, University of Ioannina,*
[2]*Department of Mechanical Engineering and Aeronautics, University of Patras, Greece*

1. Introduction

Diabetes mellitus, commonly referred to as diabetes, is a group of metabolic diseases characterized by high blood glucose concentrations resulting from defects in insulin secretion, insulin action or both [American Diabetes Association, 2008[a]]. Diabetes has been classified into two major categories, namely, type 1 and type 2 diabetes. Type 1 diabetes, which accounts for only 5-10% of those with diabetes, is caused by the cell-mediated autoimmune destruction of the insulin producing β-cells in the pancreas leading to absolute insulin deficiency. On the other hand, type 2 diabetes is a more prevalent category (i.e. accounts for ~90-95% of those with diabetes) and is a combination of resistance to insulin action and an inadequate compensatory insulin secretion. The chronic hypergycemia of diabetes is associated with long-term microvascular (diabetic neuropathy, nephropathy and retinopathy) and macrovascular (coronary artery disease, peripheral arterial disease, and stroke) complications.

Diabetes treatment requires the control of clinical and non-clinical variables affecting the blood glucose metabolism [American Diabetes Association, 2008[b]]. It is widely recognized that the tight glycemic control can prevent or reduce the progress of many long-term complications of diabetes. However, a major limiting factor in the glycemic management of type 1 and insulin treated type 2 diabetes is hypoglycemia, which is the condition where the blood glucose is much lower than normal levels. Thus, for most patients with type 1 diabetes, either using multiple insulin injections or insulin pump therapy, self-monitoring of blood glucose should be carried out three or more times a day, whereas, for patients using less frequent insulin injections or non-insulin therapies, the self-monitoring of blood glucose could be useful in achieving their glycemic targets. Recently, continuous glucose monitoring (CGM) systems have been developed which provide many significant benefits in diabetes management, especially for those patients with hypoglycaemia unawareness. Moreover, diabetes control further necessitates the monitoring and analysis of patient's contextual information, such as medication, diet, physical activity and his overall lifestyle. For instance, in type 1 diabetic patients, exercise can cause hypoglycemia in the case where the medication dose or the carbohydrate consumption is not altered.

In addition to the general guidelines that the patient follows during his daily life, several diabetes management systems have been proposed to further assist the patient in the self-management of the disease. One of the most essential components of a diabetes

management system concerns the predictive modelling of the glucose metabolism. It is evident that the prediction of glucose concentrations could facilitate the appropriate patient reaction in crucial situations such as hypoglycemia. Thus, several recent studies have considered advanced data-driven techniques for developing accurate predictive models of glucose metabolism. Data-driven techniques mainly depend on the input – output data from experiments and do not require any knowledge about the physiology of diabetes. These techniques exploit the information hidden in the data (e.g. medication, diet, physical activity, glucose measurements) in order to learn the glucose response to various stimuli. In this direction, the appearance of advanced continuous glucose sensing technologies as well as of activity monitoring devices could significantly enhance the prediction of glucose. CGM technology is used to aid in modelling the real-time trends in glucose data. However, the CGM systems do not measure the blood glucose but the glucose in the subcutaneous (s.c.) tissues. Finally, given the complexity of the glucoregulatory system, the data-driven techniques are considered to be beneficial compared to the contrary approach employing mathematical simulation models.

The scope of this chapter is (a) to present to the reader the current state of the art in predictive models of glucose metabolism in diabetes and (b) to describe a new approach to the problem by employing machine learning techniques using free – living data. The chapter is organised as follows. In Section 2, the related work in the field of modelling the glucose metabolism in diabetic patients is reviewed thoroughly. In Section 3, the proposed glucose prediction method and the derived results are presented. Finally, in Section 4, we discuss the achievements in the field and compare all relative works by identifying their advantages and disadvantages.

2. State of the art in glucose prediction

Several studies have been presented in the literature aiming at the prediction of glucose in diabetic patients. The reported methods can be divided into two major groups. The first one includes mathematical models that simulate the underlying physiology of the glucose – insulin regulatory system. Compartmental models, which are a class of linear dynamic models, have been widely used for studying various aspects of normal physiology and the pathophysiology of diabetes [Carson & Cobelli, 2001; Makroglou et al., 2006]. Recently, new important quantitative knowledge has been gained on glucose metabolism and control by insulin (e.g., the EGP profile during a meal, the hepatic glucose production, the muscle glucose utilization, the kinetics of regular and slowly acting insulin after a s.c. injection), which has allowed the development of new and more accurate simulation models [Dalla Man et al., 2007]. Nevertheless, they are still limited because of the inherent complexity of the glucose – insulin system. On the other hand, the second group of methods provides data-driven models which are able to predict the glucose concentration based only on existing input-output data. Several specific methods are available for formulating such data models, including methods of machine learning and time series analysis. In what follows, we will review the relevant literature on data-driven models of glucose metabolism.

2.1 Machine learning methods

The prediction of the glucose time series as a function of the input variables can be considered as a regression problem with a time component. The fact that the relationship between input variables (i.e. medication, diet, physical activity, stress etc.) and glucose

levels is nonlinear, dynamic, interactive and patient-specific [Tresp et al., 1999], necessitates the application of non-linear regression models such as artificial neural networks, support vector regression and Gaussian processes. Different types of neural networks have been considered in modelling the blood glucose metabolism, such as multilayer perceptron (MLP) [Kok, 2004; Zitar & Al-Labali, 2005; Quchani & Tahami, 2007], radial basis function (RBF) [Baghdadi & Nasrabadi, 2007], wavelet [Zainuddin et al., 2009], time series convolution [Tresp et al., 1999] and recurrent neural networks (RNN) [Tresp et al., 1999; Mougiakakou et al., 2006]. Additionally, Gaussian processes have derived prominent results regarding glucose prediction [Valletta et al., 2009].

The predictive performance of MLP neural network has been compared with that of Elman RNN in [Quchani & Tahami, 2007]. The aim of this study was the prediction of the blood glucose concentration before lunch based on the following features: (a) dosage of short-acting insulin, (b) dosage of long-acting insulin, (c) amount of carbohydrates, (d) stress level (from 1 to 4 discrete levels), (d) exercise level (from 1 - 4), (e) blood glucose concentration before breakfast and (f) period of time between two consecutive measurements of glucose. The data were obtained from 10 type 1 diabetic patients treated by a conventional s.c. insulin therapeutic regimen. The results showed that the Elman RNN outperform the MLP network to a significant extent (mean absolute error 10.4 mg/dl vs. 24.15 mg/dl).

An interesting approach for the prediction of glucose in type 1 diabetes was followed by [Kok, 2004; Baghdadi & Nasrabadi, 2007; Zainuddin et al., 2009] in which the day is split into four intervals (i.e. morning, afternoon, evening, night) and a different model is built for each one based on the fact that the blood glucose concentrations for these intervals are uncorrelated. The predictions of glucose at the end of each interval were made using 19 different features regarding dosage of short-acting insulin, dosage of long-acting insulin, past blood glucose measurements, amount of carbohydrates, exercise level and stress level. The only difference between the above mentioned works concerns the feature selection technique and the neural network that is employed.

A number of prediction models specific to type 1 diabetes, including non-linear compartmental models, time series convolution neural networks and RNNs, were compared in [Tresp et al., 1999]. The combination of the RNNs with a linear error model gave the best results deriving a root mean squared error (RMSE) of 51 mg/dl. The inputs that were used for this model are the following: (a) dosage of short-acting insulin, (b) dosage of long-acting insulin, (c) amount of fast carbohydrates, (d) amount of intermediate carbohydrates, (e) amount of slow carbohydrates, (f) duration of regular exercise, (g) duration of intense exercise and (h) past blood glucose level estimates. One remarkable characteristic of this approach is that the effects of food, insulin and exercise on blood glucose were approximated by linear response functions. The efficiency of RNNs has been also demonstrated in [Mougiakakou et al., 2006] where CGM data were used for the prediction of the s.c. glucose concentration in type 1 diabetic patients. Similarly with Tresp et al., compartmental models found in the literature were employed to simulate the kinetics of insulin and the absorption of carbohydrates. They report an average RMSE of 24.08 mg/dl in case where the teacher forcing learning algorithm was applied.

The Gaussian processes have also been used successfully for the prediction of glucose. More specifically, a Gaussian processes prediction model for type 1 diabetic patients was developed in [Valletta et al., 2009] based on continuous glucose measurements, physical activity information as well as information regarding food intake and insulin injections. The prediction model was evaluated on data collected from 18 patients with type 1 diabetes. A

CGM sensor was used to gather the patient's glucose concentration every five minutes. Additionally, physical activity information was collected by a multi-sensor body monitor, the so-called SenseWear armband activity monitor (BodyMedia Inc.). Given the dynamic effects of food, insulin and physical activity on glucose levels, the authors introduced time lag variables for each input that were determined by simulations. Although no quantitative results were provided, it seems that this method can predict glucose in the short-term reasonably well and is able to follow the trends in glucose time series.

2.2 Time series analysis

Time series analysis provides methods that can be used to identify systematic patterns in time series data (such as trends and seasonalities) as well as methods for time series modelling and prediction (i.e. system identification). The autocorrelation analysis of CGM time series [Bremer and Gough, 1999] made clear that glucose dynamics have a detectable structure and, thus, the glucose can be predicted by exploiting its recent history. Since that work, several studies have considered autoregressive (AR) prediction models based on CGM data [Sparacino et al., 2007; Gani et al., 2009; Gani et al., 2010]. In addition, several multivariate time series models have been developed that are enhanced with external information regarding insulin, food and physical activity [Stahl et al., 2009; Rollins et al., 2010]. However, these approaches should take into account the non-stationary behaviour of the glucose time series.

Two simple prediction methods have been applied for the first time to real CGM time series, obtained from 28 type 1 diabetic patients over a period of 48 hours, in [Sparacino et al., 2007]. In particular, the CGM time series was described by either a first order polynomial model or a first order AR model in which the parameters were dynamically identified through weighted linear least squares. In order to remove the high frequency noise from the raw CGM signals, Sparacino et al. applied a low-pass first-order Butterworth filter. For a 30 min prediction length with weight equal to 0.8, the AR model produces a median RMSE of 20.32 mg/dl and detects the positive and negative trends with an average time lag of 3.79 min and 10.06 min, respectively. Overall, the relative performance of the polynomial and the autoregressive models is quite similar during negative trends; in contrast, the autoregressive model performs better during positive trends. In addition, AR models of higher order were found to be unstable and AR models with fixed parameters to yield unacceptable prediction lags with delays equal to the prediction length.

The use of CGM data and AR models for the prediction of glucose has been also suggested by Gani et al. [Gani et al., 2009]. They have proposed an AR model of an order of 30 with fixed coefficients which successfully predicts the s.c. glucose concentration of patients with type 1 diabetes. The s.c. glucose measurements were collected by 9 patients for approximately 5 days. Similarly with Sparacino et al., the CGM data were smoothed by applying the Tikhonov regularization approach. The construction of AR models through regularized least squares resulted in AR coefficients that reflect the temporal dependencies in the glucose signal, and in stable, accurate predictions. In particular, the AR model is able to yield 30 min glucose concentration predictions with an average RMSE of 1.8 mg/dl and a negligible prediction time lag of 0.2 min, and 60 min glucose concentration predictions with an average RMSE of 12.6 mg/dl and average prediction lag of 12.3 min. Gani et al. argue that AR models of low order, such that proposed by Sparacino et al., can produce acceptable predictions, but they introduce significant delays between predicted and measured values, because they are not sufficient to capture the temporal variations of the glucose signal. In

addition, they confirm the instability of the AR models of order higher than one, which is also reported by Sparacino et al., in the case where the AR coefficients are not regularized.
In a subsequent work [Gani et al., 2010], the authors showed that their method results in AR models in which the coefficients do not vary significantly among different individuals, suggesting the feasibility of obtaining individual-independent predictive models. For this purpose, they employed data from three separate studies, involving patients with both type 1 and type 2 diabetes, and utilizing three different CGM devices. The results of this investigation were attributed to the fact that the features of the glucose signals in the frequency domain were found common among patients. Considering that the AR models represent the signal's frequency information and are invariant with respect to the signal's amplitude and phase, the development of similar models was predictable.
Many researchers have incorporated in their time series models the influence of external input variables. Stahl et al. [Stahl et al., 2009] investigated the ability of a variety of linear and non-linear system identification methods (i.e. autoregressive moving average (ARMA) linear regression, autoregressive moving average with exogenous input (ARMAX) linear regression, Wiener model identification, subspace-based identification) to predict the blood glucose concentration for the next two hours with a reasonable accuracy. This target accuracy was defined as a standard deviation of the prediction error less than 18 mg/dl in the 95% of the cases. The identified models were fitted to real data collected during the first six months of a newly diagnosed type 1 diabetic patient who used a traditional blood glucose meter. For modelling purposes, the blood glucose samples were interpolated using a least-squares spline method to obtain a sampling rate of 15 min. In addition, the absorption of injected insulin from the s.c. tissues as well as the digestion and absorption of carbohydrates were described using compartmental models and proposed models found in the literature. The linear models (ARMAX, subspace-based and general transfer function models) were proved insufficient to predict the glucose responses. Therefore, a log-normalized linear model based on subspace-based identification and a GTFM-Wiener model was employed; nevertheless, the prediction error was rather improved.
Recently, a causation modelling methodology with the ability to infer the s.c. glucose concentration using an extensive set of highly correlated non-invasive input variables has been developed [Rollins et al., 2010]. More specifically, the inputs concerned food (i.e. carbohydrates, fats, and proteins), physical activity and stress. This study was initiated by the requirement to determine the independent, dynamic contribution of each input to the overall dynamics of glucose response. For this reason, the predicted glucose was completely determined from measured input data only and previously measured glucose levels did not used in its inference. Accordingly, the s.c glucose concentration was modelled though a block-oriented Wiener network that uses non-linear, in the parameters' space, response surfaces. The prediction method was evaluated using real data of a type 2 diabetic patient collected under free-living conditions over a period of 25 consecutive days. For 5 min predictions, they report an average absolute error of 13.3 mg/dl and a correlation coefficient of 0.7, and they argue that one critical reason for not being able to achieve better results is probably due to lack of information about insulin. However, one important characteristic of this approach is that its predictive accuracy is not limited by the size of the prediction horizon, since it does not depend on past glucose measurements. Moreover, the analysis of the independent dynamic response of each input revealed significant conclusions regarding their effect on dynamic glucose behaviour.

3. The proposed method

Prediction of glucose can be used to provide immediate feedback to the diabetic patients about how the glucose is affected by their lifestyle and treatment. In addition, it offers the means of making real-time suggestions regarding modifications to diet and activity related profile as well as diabetes medications in order to avoid critical events. This study investigates the ability to model the glucose metabolism of type 1 diabetic patients using a multi-parametric set of data recorded under free-living conditions. The proposed method considers the effect of diet, medication, and physical activity on glucose control with the aim to provide accurate glucose predictions.

3.1 Materials and methods

3.1.1 Materials

Seven patients with type 1 diabetes participated in this study who were treated with insulin injections (insulin doses and types were different for each patient). The observation period of the study was on average 10 days (range from 5 – 14 days). All patients wore the Guardian Real-Time CGM system (Medtronic Minimed) that monitors the s.c. glucose concentrations every 5 min. The glucose sensor calibration requires at least four blood glucose measurements to be made daily using a standard blood glucose meter. In addition, the glucose sensors have to be replaced every 3 days. The patients were also equipped with the SenseWear body monitoring system (BodyMedia Inc.) which monitors their daily physical activities. The SenseWear armband collects data using five sensors: heat flux, skin temperature, near body temperature, galvanic skin response and a two axis accelerometer. Finally, information regarding the food intake (i.e. type of food, serving sizes and time) and the insulin injections (type, dose and time) was recorded by the patients using a specially designed paper diary. The food composition (i.e. calories, carbohydrates, fat etc.) was postanalyzed by a dietician.

3.1.2 The method

The method for the prediction of the s.c. glucose concentration is presented schematically in Figure 1. It comprises compartmental models of the glucose – insulin regulatory system and a predictive model of glucose. The compartmental models are used to simulate (a) the ingestion and absorption of carbohydrates (the Meal Model), (b) the absorption and the pharmacokinetics / pharmacodynamics of subcutaneously administered insulin (the Insulin Model) as well as (c) the impact of exercise on glucose - insulin metabolism (the Exercise Model). In addition, support vector machines for regression (SVR) are employed to provide individualized glucose predictions. The input variables of the proposed model include the rate of glucose appearance in plasma after a meal, R_a, the plasma insulin concentration, I_p, the s.c. glucose measurements, gl, as well as a set of physical activity related variables. As it can be seen from Figure 1, we assume two different approaches to investigate the physical activity's effects on diabetes. In the first approach, the Metabolic Equivalent of Task (*MET*), the heat flux (*hf*) and the skin temperature (*st*) variables, which are recorded by the SenseWear armband, are used as inputs in the model. The second approach utilizes the alterations in circulating glucose and insulin concentrations (G_{exer}, I_e) during and shortly after exercise as computed by the Exercise Model. The main components of our method are presented in the following subsections.

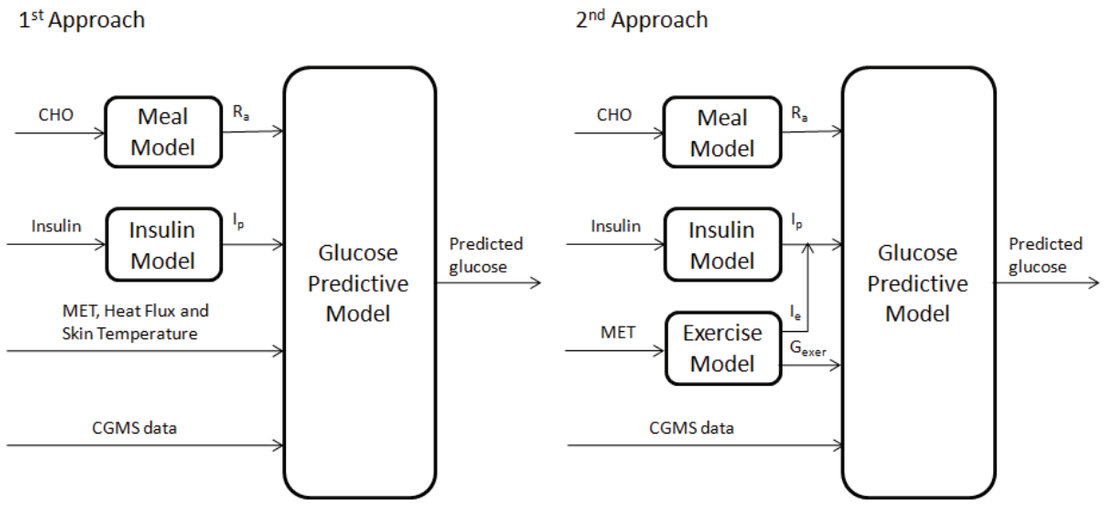

Fig. 1. Schematic representation of the proposed method

3.1.2.1 The insulin model

The absorption of subcutaneously injected insulin is described by the pharmacokinetic model proposed in [Tarin et al., 2005]. This model describes the diffusion of insulin through the s.c. depot, the molecular dissociation of insulin (hexameric/dimeric) and the absorption of insulin into the bloodstream by the following nonlinear partial differential equations:

$$\frac{\partial c_d(t,r)}{\partial t} = P\left(c_h(t,r) - Qc_d(t,r)^3\right) - B_d c_d(t,r) + D\nabla^2 c_d(t,r), \tag{1}$$

$$\frac{\partial c_h(t,r)}{\partial t} = -P\left(c_h(t,r) - Qc_d(t,r)^3\right) + \kappa c_b(t,r)\left(c_{h,\max} - c_h(t,r)\right) + D\nabla^2 c_h(t,r), \tag{2}$$

$$\frac{\partial c_b(t,r)}{\partial t} = -\kappa c_b(t,r)\left(c_{h,\max} - c_h(t,r)\right) + d_b D\nabla^2 c_b(t,r), \tag{3}$$

where c_h, c_d, c_b are the hexameric, dimeric and bound insulin concentrations in the s.c. tissue, respectively, D is the diffusion constant, d_b is a non-dimensional factor that reduces the diffusion effect, P is the dimeric-to-hexameric association rate, Q is the corresponding equilibrium constant, κ is the proportional factor of disengagement of hexameric insulin from the bound state, $c_{h,max}$ is the maximum concentration of hexameric insulin and B_d is the absorption rate constant. The bound state in this model is a virtual state introduced to model the dynamics of long-acting insulin analogues e.g. Glargine. As can be observed from these equations, the diffusion process of insulin in the s.c. tissue is considered to be isotropic i.e. homogeneous and with rotational symmetry with respect to the origin (injection site). Additionally, it is assumed that only the dimeric form of insulin can be absorbed into the plasma with a rate proportional to its concentration. Hence, the exogenous insulin flow (U/min) into the bloodstream is given by:

$$I_{ex}(t) = B_d \int_{V_{sc}} c_d(t,r)dV, \tag{4}$$

where V_{sc} is the complete s.c. volume. This model allows the description of all insulin formulations through the adequate selection of the parameters Q, D, B_d, k, $c_{h,max}$, and d_b. However, the system of partial differential equations has no closed solution and therefore, a time and space discretization is implemented for the numerical calculation of the dimeric insulin concentration.

To estimate the plasma insulin concentration, I_p (uU/ml), a compartmental modelling approach is used [Cobelli et al., 1982]. The model describes the concentration – time evolution of plasma insulin I_p, hepatic insulin I_h and interstitial insulin I_i after a s.c. injection and is given as follows:

$$\dot{I}_p = \frac{I_{ex}(t)}{V_d} - k_1 I_p(t) + k_2 I_h(t) + k_3 I_i(t), \tag{5}$$

where V_d is the plasma insulin distribution volume, and k_1, k_2, k_3 are the rate constants of plasma, hepatic and interstitial insulin elimination, respectively. The input to this physiological model is the exogenous insulin flow, $I_{ex}(t)$, and the output is the plasma insulin concentration I_p. Figure 2(a) shows the exogenous insulin flow profile of Aspart and Glargine insulin injections resulting from the insulin therapy of Patient 4 over a time horizon of two days. It can be seen that the profile varies substantially depending on the injected insulin doses and formulations, i.e. insulin Glargine has a slower onset of action and a longer duration of action than Aspart insulin, whose activity peaks rapidly. The plasma insulin concentration of the combined effect of both insulin types is depicted in Figure 2(b). The long action of Glargine insulin, which resembles the basal insulin secretion of non-diabetic individuals, as well as the effect of Aspart insulin, which is used for controlling the postprandial hyperglycemia, can be observed.

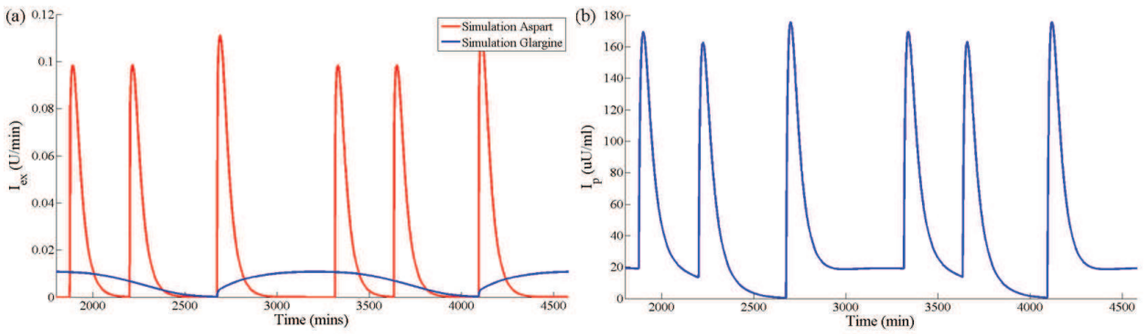

Fig. 2. (a) Exogenous insulin flow of Patient 4 as computed by the insulin compartmental model, (b) Cummulative profile of plasma insulin (Aspart and Glargine) concentration of Patient 4 as computed by the insulin compartmental model

3.1.2.2 The meal model

The model by Lehmann and Deutch [Lehmann & Deutch, 1992] is used to describe the ingestion and absorption of carbohydrates intake. This model describes the rate of appearance of glucose in plasma on the assumption that the rate of gastric emptying is a trapezoidal function and that the intestinal glucose absorption follows first order linear kinetics. The amount of glucose in the gut, q_{gut}, after the ingestion of a meal containing D grams of glucose equivalent carbohydrates is defined as:

$$\dot{q}_{gut}(t) = -k_{abs} q_{gut}(t) + G_{empt}(t, D), \tag{6}$$

where k_{abs} is the rate constant of intestinal absorption and G_{empt} (mg/min) is the gastric emptying function.
The function G_{empt} is described by:

$$G_{empt} = \begin{cases} V_{\max}/T_{asc}, & t < T_{asc} \\ V_{\max}, & T_{asc} < t \le T_{asc} + T_{\max} \\ V_{\max} - (V_{\max}/T_{des})(t - T_{asc} - T_{\max}), & T_{asc} + T_{\max} < t \le T_{asc} + T_{des} \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

where

$$T_{\max} = \frac{2D - V_{\max}(T_{asc} + T_{des})}{2V_{\max}}, \tag{8}$$

corresponds to the duration of the period for which the gastric emptying function is constant and maximum (V_{max}), and T_{asc}, T_{des} are the duration of rising up and dropping periods of G_{empt}, respectively. Then, the rate of appearance of glucose in plasma (mg/min) is given as:

$$R_a(t) = k_{abs} q_{gut}(t). \tag{9}$$

The values for the model parameters have been derived from [Lehmann & Deutch, 1992] and are assumed to be patient-independent.

3.1.2.3 The exercise model

The model used to derive the exercise-induced changes on glucose – insulin metabolism is based on a recent study of Roy & Parker [Roy & Parker, 2007]. In particular, we have developed an algorithm that extracts the most significant exercise events by analyzing the measurements provided by the SenseWear armband. Then, the physiological processes, which occur during an exercise event and at the recovery period, are simulated utilizing the model presented in [Roy & Parker, 2007]. This model describes the effect of exercise on the dynamics of glucose and insulin as follows:

$$\dot{G}_{prod} = a_1 PVO_2^{\max}(t) - a_2 G_{prod}(t), \tag{10}$$

$$\dot{G}_{up} = a_3 PVO_2^{\max}(t) - a_4 G_{up}(t), \tag{11}$$

$$\dot{I}_e = a_5 PVO_2^{\max}(t) - a_6 I_e(t). \tag{12}$$

The terms G_{prod} and G_{up} represent the rates (mg/min) of hepatic glucose production (glycogenolysis) and glucose uptake induced by exercise, respectively, while the I_e (uU/(ml.min)) denotes the rate of insulin removal from the circulatory system during and after exercise. In addition, although the corresponding equation is not given here, the rate of glycogenolysis during prolonged exercise decreases by a factor of G_{gly} due to the depletion

of glycogen stores in the liver. The dynamics of glycogenolysis are described in detail in [Roy & Parker, 2007]. The intensity of the exercise (intense walking) as recorded by the activity device over time for Patient 1 is shown in Figure 3(a) along with the computed metabolic response of the patient. More specifically, Figure 3(b) illustrates the glucose uptake rate (G_{up}) and the hepatic glucose production rate (G_{prod} - G_{gly}) during and after exercise, where it can be observed that the effects of exercise progressively attenuate during the recovery period. The rate of insulin removal from plasma (I_e), as shown in Figure 3(c), exhibits also similar behaviour.

As shown in the equations (10-12), the exercise intensity is quantified by the percentage of the maximum oxygen consumption ($PVO_2^{\max}$). Since the SenseWear armband does not report the oxygen uptake (VO_2) during exercise, the term $PVO_2^{\max}$ was calculated by:

$$PVO_2^{\max} = \frac{VO_2}{VO_2^{\max}} = \frac{3.5MET}{VO_2^{\max}}, \tag{13}$$

where $VO_2^{\max}$ is the maximal oxygen uptake and depends on patient's age, gender and physical status. For each patient, the $VO_2^{\max}$ value was derived from reference tables.

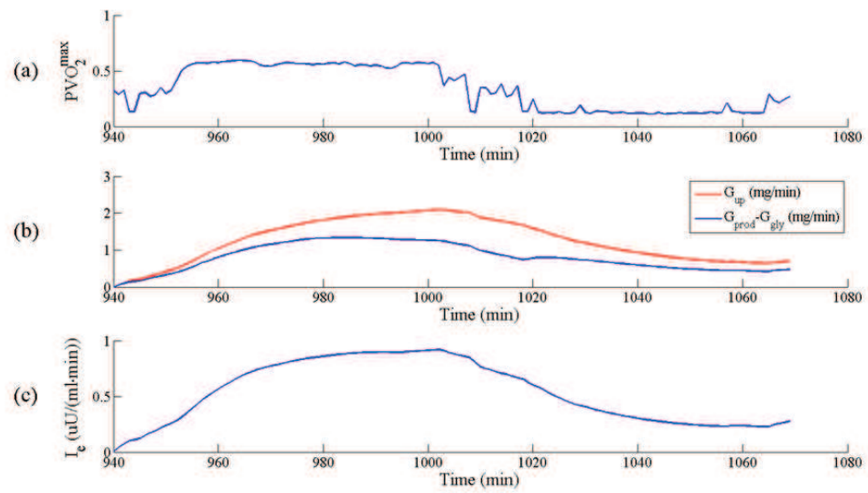

Fig. 3. The effects of an exercise event on metabolism of Patient 1. (a) Data from sensors, (b) Glucose uptake and production rate as computed by the exercise compartmental model, (c) Insulin removal rate as computed by the exercise compartmental model

The implications of exercise in the glucose – insulin regulatory system are incorporated into the proposed method by introducing an additional input variable, the G_{exer}, which describes the blood glucose variation (mg/min) during exercise and at the recovery period:

$$G_{exer} = \left(G_{prod} - G_{gly}\right) - G_{up}. \tag{14}$$

Accordingly, the insulin dynamics are modified by adding the term I_e in the equation (5), resulting in:

$$\dot{I}_p = \frac{I_{ex}(t)}{V_d} - k_1 I_p(t) + k_2 I_h(t) + k_3 I_i(t) - I_e(t). \tag{15}$$

Regarding the parameters for the exercise model, they are obtained from [Roy & Parker, 2007].

3.1.2.4 The glucose predictive model

In this study, a support vector machine for regression [Smola & Scholkopf, 2003; Bishop, 2006] is employed to predict the s.c. glucose concentrations. Let us consider that the training data set D comprises N input vectors $x^1, \ldots, x^N$ ($x^i \in R^d$) with corresponding target glucose values $t^1, \ldots, t^N$. In ε-SVR our goal is to find a linear model of the form:

$$y(x) = w^T \phi(x) + b, \tag{16}$$

which must satisfy the following conditions:

$$t^n \le y(x^n) + \varepsilon + \xi_n, \tag{17}$$

$$t^n \le y(x^n) - \varepsilon + \hat{\xi}_n. \tag{18}$$

The function $\phi(x)$ denotes a fixed feature-space transformation and w and b are the weights and bias parameters, respectively. The error function for ε-SVR is defined as:

$$C\sum_{n=1}^{N}(\xi_n + \hat{\xi}_n) + \frac{1}{2}\|w\|^2, \tag{19}$$

which must be minimized subject to the constraints $\xi_n, \hat{\xi}_n \ge 0$, as well as (17) and (18). This can be achieved by introducing the Lagrange multipliers $a_n, \hat{a}_n \ge 0$ and $\mu_n, \hat{\mu}_n \ge 0$ and by minimizing the Lagrangian:

$$\begin{aligned} L = C\sum_{n=1}^{N}(\xi_n + \hat{\xi}_n) + \frac{1}{2}\|w\|^2 - \sum_{n=1}^{N}(\mu_n\xi_n + \hat{\mu}_n\hat{\xi}_n) \\ - \sum_{n=1}^{N}\alpha_n(\varepsilon + \xi_n + y_n + t_n) - \sum_{n=1}^{N}\hat{\alpha}_n(\varepsilon + \hat{\xi}_n + y_n + t_n). \end{aligned} \tag{20}$$

Solving the optimization problem, it is found that the predictions for the new inputs can be made using:

$$y(x) = \sum_{n=1}^{N}(a_n - \hat{a}_n)k(x, x^n) + b, \tag{21}$$

where $k(x, x') = \phi(x)^T \phi(x')$ is the kernel function. From the corresponding Karush-Kuhn-Tucker (KKT) conditions, which state that at the solution the product of the dual variables and the constraints must vanish, results that $a_n\hat{a}_n = 0$. Therefore, the SVR provides a sparse solution, since the only terms that have to be evaluated in the predictive model are those which involve the support vectors, i.e. the data in the training set for which exactly one of the Lagrange multipliers is greater than zero.

To be more specific, given the input x, the prediction of the s.c. glucose concentration, y, at the time $t+l$, assuming that t is the current time, is given by:

$$y_{t+l}(x) = y_{t+l}(x_1,\ldots,x_d). \quad (22)$$

where $x_i = x_i(t),\ldots,x_i(t-n_i\Delta t)$, with i =1,…, d, denotes the inputs in the model, $n_i\Delta t$ is the time lag for the input x_i, Δt is the sampling time and l is the prediction length.

3.2 Model training and evaluation

The proposed method is evaluated using the dataset obtained from seven type 1 diabetic patients. The SVR is trained individually for each patient and a V-fold cross validation algorithm is used to avoid over-fitting. More specifically, V-fold cross validation splits the dataset D in k equal parts, where k is defined as the total number of days for which the patient is monitored. Thus, the value of V coincides with the value of k and each fold contains the data of the i^{th} day, with i =1,…, k. The SVR is built using a linear kernel function and the parameter ε in the ε-insensitive loss function is set equal to 0.001. The regularization parameter C is optimized using a grid search method. Similarly, V-fold cross validation is used by the search method to calculate the optimal values for that parameter.

Time lags of 30 min are considered for the I_p, Ra and gl input variables, while the time lag for the exercise-related inputs (i.e. MET, st, hf and G_{exer}) is assumed to be 3 hours. The sampling time, Δt, was 5 min for all the above cases. Predictions are performed for four different values of prediction length l, i.e. 15, 30, 60 and 120 min.

The predictive accuracy of the proposed method is assessed by calculating the RMSE, and the correlation coefficient, r, for each patient's test set. Furthermore, the Clarke's Error Grid Analysis (EGA) [Kovatchev et al., 2004; Clarke, 2005] is used to assess the clinical significance of the errors between the predicted and the measured s.c. glucose concentrations. The Clarke's EGA method uses a Cartesian diagram, in which the predicted values are displayed on the y-axis, whereas the values from glucose sensor are displayed on the x-axis. This diagram is subdivided into 5 zones: A, B, C, D and E. The points that fall within zones A and B represent sufficiently accurate or acceptable glucose results, points in zone C may result in unnecessary corrections, points in zone D could lead to incorrect treatments, and points in zone E represent erroneous treatment.

3.3 Results

The RMSE (mg/dl) and r values obtained from the first approach are reported in Table 1. It can be observed that the short-term glucose predictions (i.e. for 15 and 30 min) present low error and high degree of correlation with the real glucose profiles. More specifically, the average value of RMSE for 15 min and 30 min predictions is equal to 9.60 mg/dl and 16.23 mg/dl, respectively. In both cases, the predicted glucose concentrations exhibit a strong correlation with the measured values (i.e. 0.95 and 0.88). However, as prediction length increases (i.e. for 60 and 120 min), the performance of the proposed method significantly decreases. Concerning the 60 min predictions, the derived results are still adequate compared to the previous values, whereas, the accuracy of the 120 min predictions is considerably lower. In addition, the predictions for some patients (i.e. Patient 2, 3, 5, 6, 7) are found systematically more accurate, in terms of RMSE, than for Patient 1 and Patient 4 which most probably resulted from better model training due to the longer follow-up period; nevertheless, slight differences are observed in the associated r values. For the second approach, the derived results, as it is shown in Table 2, are almost equal to those for the first approach.

No. of Patient	Prediction Length							
	15 min		30 min		60 min		120 min	
	RMSE (mg/dl)	r	RMSE (mg/dl)	r	RMSE (mg/dl)	r	RMSE (mg/dl)	r
Patient 1	12.57	0.96	21.36	0.90	33.06	0.75	62.29	0.28
Patient 2	9.69	0.95	16.32	0.87	24.52	0.68	31.11	0.37
Patient 3	9.33	0.95	15.85	0.87	25.77	0.67	33.91	0.46
Patient 4	11.92	0.92	19.24	0.81	29.06	0.62	39.25	0.40
Patient 5	6.45	0.91	11.04	0.91	17.89	0.70	26.22	0.48
Patient 6	10.85	0.95	18.38	0.86	24.82	0.72	34.64	0.49
Patient 7	6.42	0.98	11.45	0.93	18.84	0.82	23.48	0.70
Average (SD)	9.60 (2.45)	0.95 (0.02)	16.23 (3.87)	0.88 (0.04)	24.85 (5.33)	0.71 (0.06)	35.84 (12.81)	0.45 (0.13)

Table 1. Prediction results obtained from the first approach (exercise described only by sensor data)

No. of Patient	Prediction Length							
	15 min		30 min		60 min		120 min	
	RMSE (mg/dl)	r	RMSE (mg/dl)	r	RMSE (mg/dl)	r	RMSE (mg/dl)	r
Patient 1	12.07	0.96	19.93	0.91	30.99	0.80	55.43	0.46
Patient 2	9.58	0.96	15.91	0.88	24.06	0.69	31.24	0.42
Patient 3	9.28	0.95	15.59	0.87	25.35	0.69	33.82	0.44
Patient 4	11.73	0.92	18.97	0.82	28.70	0.63	37.99	0.35
Patient 5	6.50	0.90	11.09	0.91	16.78	0.65	23.80	0.47
Patient 6	11.18	0.95	18.95	0.86	26.58	0.70	37.18	0.46
Patient 7	6.20	0.98	11.69	0.94	21.19	0.80	33.60	0.51
Average (SD)	9.51 (2.39)	0.95 (0.03)	16.02 (3.55)	0.88 (0.04)	24.81 (4.74)	0.71 (0.07)	36.15 (9.70)	0.44 (0.05)

Table 2. Prediction results obtained from the second approach (exercise described by compartmental modelling)

Clarke's EGA clearly shows that the vast majority of the predicted-measured glucose points lay in zones A and B, which indicate clinically acceptable results. On the other hand, a small amount of points belong to the other zones (i.e. C, D and E), which indicate potentially dangerous overestimation or underestimation of the actual values. Figure 4 represents the Clarke's EGA plots for Patient 5, in the case where the first approach is followed. In this figure, it is shown that as the prediction length increases, the plots become more spread, as expected. Tables 3 and 4 report the average results obtained from the two different approaches, respectively. We observe that nearly all the points lay in zones A and B, even if a higher prediction length is considered. Occasional points belong to the C zone, whereas only a few points belong to the D zone. Finally, no points belong to the erroneous E zone. In addition, no differences are evident between the results obtained from the two approaches.

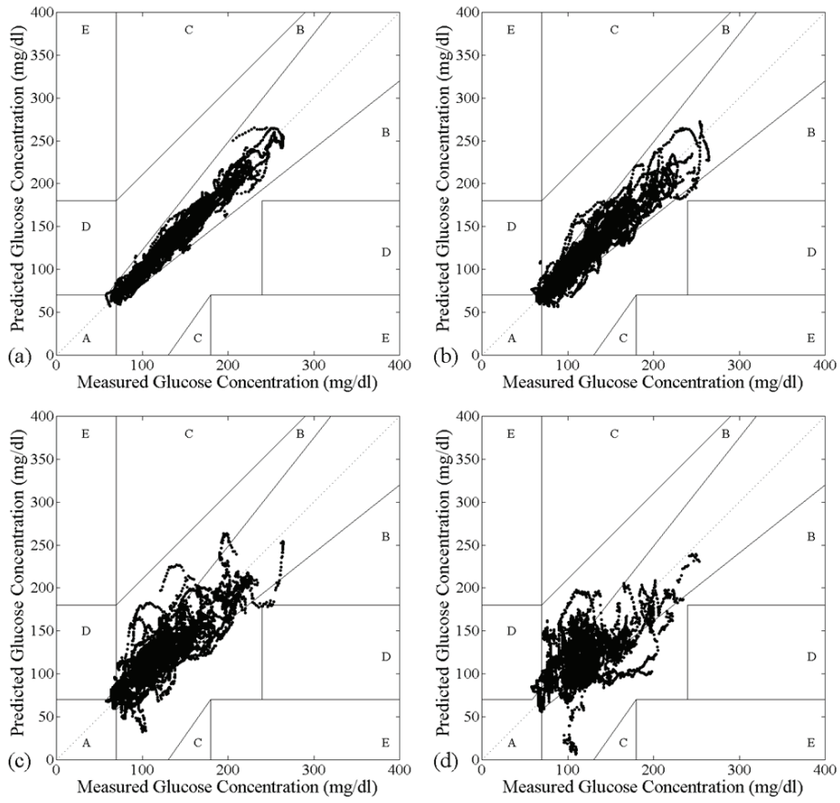

Fig. 4. Clarke's EGA diagrams for Patient 5 based on the first approach. (a) – (d) correspond to different prediction lengths (i.e. 15, 30, 60, 120 min)

Zone	Prediction Length			
	15 min	30 min	60 min	120 min
Zone A	98.86 %	92.54 %	80.02 %	62.91 %
Zone B	1.08 %	6.97 %	18.49 %	33.78 %
Zone C	0.00 %	0.02 %	0.07 %	0.37 %
Zone D	0.06 %	0.47 %	1.42 %	2.94 %
Zone E	0.00 %	0.00 %	0.00 %	0.00 %

Table 3. Average percentages of points falling into the different zones of the Clarke's EGA for the first approach (exercise described only by sensor data)

Zone	Prediction Length			
	15 min	30 min	60 min	120 min
Zone A	98.58 %	92.54 %	79.96 %	60.10 %
Zone B	1.35 %	6.89 %	18.33 %	36.84 %
Zone C	0.00 %	0.02 %	0.09 %	0.20 %
Zone D	0.07 %	0.55 %	1.62 %	2.86 %
Zone E	0.00 %	0.00 %	0.00 %	0.00 %

Table 4. Average percentages of points falling into the different zones of the Clarke's EGA for the second approach (exercise described by compartmental modelling)

4. Discussion

Glucose metabolism is a non-linear, dynamic system, the behaviour of which has been extensively modelled by data-driven methods. Table 5 provides a summary of most of the studies on glucose prediction in diabetes reported in the literature, along with a short description of methods used, input variables and patient types.

During the last decade, the application of machine learning methods in predictive modelling of glucose concentration in patients with diabetes has gained much attention. Simple feed forward neural networks [Kok, 2004; Zitar & Al-Labali, 2005; Quchani & Tahami, 2007; Baghdadi & Nasrabadi, 2007] as well as more sophisticated types such as recurrent [Tresp et al., 1999; Mougiakakou et al., 2006] and wavelet neural networks [Zainuddin et al., 2009] have been utilised up to now for the prediction of the glucose concentration in diabetic patients. The results obtained in these works show that reasonably accurate glucose predictions can be made; however, a direct comparison between them is not feasible since they refer to different prediction horizons. Furthermore, the performance of these methods highly depends on the input which is used. Firstly, the fact that the predictions are mainly based on glucose measurements recorded 3-4 times per day inevitably affects the output of the prediction. Nevertheless, the development of glucose sensors introduced the utilization of CGM data for the prediction of glucose which was a breakthrough in the field [Mougiakakou et al., 2006]. In addition, in most of these studies the physical activity is qualitatively described, except for the work of Valletta et al. [Valletta et al., 2009] which employs Gaussian processes to model the glucose variations in response to real activity data recorded continuously throughout the day. Given that activity plays an important role in glucose regulation, this consideration constitutes a substantial limitation.

The prediction of glucose in diabetic patients has also been addressed through time series analysis. The fact that glucose can be predicted by exploiting the recent history of CGM data was initially suggested by Bremer and Gough [Bremer and Gough, 1999]. This was further demonstrated by the findings of three subsequent studies [Sparacino et al., 2007; Gani et al., 2009; Gani et al., 2010] showing that AR models can provide stable, accurate predictions. One advantage of AR models consists in the interpretability of the AR coefficients, which describe the temporal dependencies in the glucose signal. In addition to this, the estimation of the model parameters involves a convex optimization problem with a unique minimum. Apart from AR models, linear and non-linear time series models with external input variables have also been developed [Stahl et al., 2009; Rollins et al., 2010]. It is noteworthy that in the study of Rollins et al. [Rollins et al., 2010] employing a block-orient Wiener network, real data from an activity device are utilised to quantify the effect of physical activity. Equally important, the authors made an attempt to examine the individual dynamic characteristics of each input regarding food, insulin, and activity in order to interpret their effects on glucose behaviour. However, much of modern theory of time series is concerned with stationary time series and, therefore, it is needed to establish some conditions, e.g. CGM data must be a first and second order stationary process.

The problem of glucose prediction in diabetic patients from a multi-parametric set of free-living data (i.e. food, insulin, physical activity and continuous glucose measurements) has been addressed in the context of Gaussian processes [Valletta et al., 2009] and Wiener networks [Rollins et al., 2010]. The same problem is treated here with the aid of support vector machines for regression. Accounting for the physiological processes related to diabetes (i.e insulin absorption, gut absorption), we employed appropriate compartmental

models found in the literature. In addition, we assumed two different approaches to investigate the activity's effects on diabetes. For the first time, to the best of our knowledge, the changes on glucose - insulin levels induced by exercise are incorporated into a glucose predictive model, and, moreover, an exercise model is fed with real sensor data to indicate the exercise intensity.

Study	Diabetes Type (No of Patients)	Input Variables	Method
Zitar & Al-Labali, 2005	Type 2 (70)	BG, Insulin, Meal Announcement (1 or 0), Exercise Announcement (1, 0)	MLP Neural Network
Quchani & Tahami, 2007	Type 1 (10)	BG, Insulin, CHO, Exercise Levels, Stress Levels	Elman RNN
Kok, 2004	Type 1 (1)	BG, Insulin, CHO, Exercise Levels, Stress Levels	MLP Neural Network
Baghdadi & Nasrabadi, 2007	Type 1 (1)	BG, Insulin, CHO, Exercise Levels, Stress Levels	RBF Neural Network
Zainuddin et al., 2009	Type 1 (1)	BG, Insulin, CHO, Exercise Levels, Stress Levels	Wavelet Neural Network
Tresp et al., 1999	Type 1 (1)	BG, Insulin, CHO, Exercise Duration	RNN
Mougiakakou et al., 2006	Type 1(4)	CGM Data, Insulin, CHO	RNN
Valletta et al., 2009	Type 1 (18)	CGM Data, Insulin, CHO, Exercise Data	Gaussian Processes
Sparacino et al., 2007	Type 1 (28)	CGM Data	AR model
Gani et al., 2009	Type 1 (9)	CGM Data	AR model
Stahl et al., 2009	Type 1(1)	BG, Insulin, CHO	ARMA, ARMAX, Wiener and Subspace-Based System Identification
Rollins et al., 2010	Type 2 (1)	CGM Data, CHO, Fats, Proteins, Exercise Data	Block-Oriented Wiener Network
This work	Type 1(7)	CGM Data, Insulin, CHO, Exercise Data	SVR

Table 5. Summary of works on glucose prediction in diabetic patients using data-driven techniques (BG: blood glucose, CHO: carbohydrates)

The application of compartmental models describing the absorption of subcutaneously administered insulin and the absorption of glucose from the gut following a meal is also reported in several studies [Mougiakakou et al., 2006; Valleta et al., 2009; Stahl et al., 2009]

dealing with the problem of the prediction of glucose. Compartmental analysis of these processes is necessitated because the data collected by the patients (i.e. food, insulin) are non-uniformly sampled; on the contrary most of the predictive methods require uniformly sampled data. Nevertheless, response functions can also be used for this purpose as in [Tresp et al., 1999; Rollins et al., 2010]. On the other hand, the reason why we used exercise compartmental models was to examine if accurate predictions could be achieved from simulation outputs (made from real exercise data) which model the metabolic response not only during exercise but also during the recovery period. The overall approach has the advantages of SVR. First we must consider that the optimization problem is transformed into a dual convex quadratic programming leading to a global minimum. Moreover, compared with the existing kernel regression modelling approaches (i.e. RBF), it gives significant algorithmic and representation advantages by producing sparser models. Finally, it has to be mentioned that SVR is effective even on large and high dimensional datasets, which is the case in glucose prediction problems.

The results obtained in the present study make clear that the glucose concentration in patients with type 1 diabetes can be predicted with a sufficient numerical accuracy in the short-term. The increase in the length of prediction leads to more significant deviations of the obtained predictions from the reference glucose concentrations as also reported in previous studies [Sparacino et al., 2007; Stahl et al., 2009; Gani et al., 2009]. Small differences were observed in the predictive accuracy among the patients of our study, which indicates that the proposed scheme could be applied to most of type 1 patients (given that lifestyle data are recorded in a similar way). It becomes apparent from the Clarke's EGA that the performance of the proposed prediction method is also significant from a clinical point of view since practically all of our predictions do not fall in the zones which would lead to incorrect or erroneous treatment (i.e. C - E). A direct comparison of the present study could be performed with that of Valletta et al. [Valletta et al., 2009]; however, the authors provide no quantitative results. Compared to [Rollins et al., 2010], we found more accurate predictions, but we have to consider that the model proposed by Rollins et al. does not exploit information about insulin, since it concerns type 2 diabetic patients. Although the studies employing AR models [Stahl et al., 2009; Gani et al., 2009] produced better results, they largely depend on the assumption that the CGM data are described by a stationary process.

Tables 1 and 2 show that the two approaches which were used to describe the physical activity yielded almost equal results. Since in the second approach the predictions are based only on segments containing significant (discrete) exercise events, this would imply that sufficient predictions could still be achieved without necessitating the activity monitor to be worn continuously throughout the day, but only during exercise, which practically enhances the possibilities of a predictive system to be acceptable by the patients. Moreover, this second modelling approach can accept as input descriptive exercise event announcements manually notified by the patient; however, the predictive accuracy in that case should be tested. In addition, the ability to analyse and predict the effects of exercise on glucose metabolism can be exploited for providing to the patient advices on hypothetical scenarios for forthcoming exercise events. The above advantages offered by exercise compartmental modelling are extremely useful for diabetes advisory systems.

Considering the intra- and inter-individual variability in the metabolic response to food, insulin and exercise, it could be very important to estimate the parameters involved in the compartmental models from each patient's data. However, this process would require the conduction of tracer experiments, and thus it was not included in this study. Another

simplification in our study was that the influence of the fats, proteins and other food nutrients on the dynamics of the digestive and absorptive processes, as well as the effect of the glycemic index, was not considered. Also, there are factors affecting the insulin absorption and insulin kinetics (e.g. site of injection, ambient and body temperature), which have not been investigated. The introduction of these variables in our study would lead to more realistic modelling of the glucose metabolism; therefore, it will be taken into account in the future. We also intend to improve the performance of the SVR by determining the most appropriate kernel function and by estimating the parameter ε in the insensitive loss function, but our final objective is to form a generalized predictive model that can be applied to group of patients.

5. References

American Diabetes Association (2008)[a]. Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care*, Vol. 31, No. 1, (January 2008) 55-60, 1935-5548

American Diabetes Association (2008)[b]. Standards of Medical Care in Diabetes. *Diabetes Care*, Vol. 31, No. 1, (January 2008) 12-54, 1935-5548

Baghdadi, G. & Nasrabadi, A. M. (2007). Controlling blood glucose levels in diabetics by neural network predictor, *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3216-3219, 978-1-4244-0787-3, France, August 2007, IEEE, Lyon

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*, Springer, 0-387-31073-8, New York

Bremer, T. & Gough D. A. (1999). Is Blood Glucose Predictable From Previous Values? A Solicitation for Data. *Diabetes*, Vol. 48, (March 1999) 445-451, 1939-327X

Caumo A.; Simeoni M. & Cobelli C. (2001). Glucose Modelling, In: *Modelling Methodology for Physiology and Medicine*, Carson, E. & Cobelli C., (Ed. 2001), 337-372, Academic Press, 0-12-160245-1, San Diego

Clarke, W. L. (2005). The original clarke error grid analysis (EGA). *Diabetes Technology and Therapeutics*, Vol. 7, No. 5, (October 2005) 776-779, 1520-9156

Cobelli, C.; Federspil, G.; Pacini, G.; Salvan A. & Scandellari C. (1982). An integrated mathematical model of the dynamics of blood glucose and its hormonal control. *Mathematical Biosciences*, Vol. 58, No. 1, (February 1982) 27-60, 0025-5564

Dalla Man, C.; Rizza, R. A. & Cobelli, C. (2007). Meal Simulation Model of the Glucose-Insulin System. *IEEE Transactions on Biomedical Engineering*, Vol. 54, No. 10, (October 2007) 1740-1748, 0018-9294

Gani, A.; Gribok, A. V.; Rajaraman, S.; Ward, W. K. & Reifman, J. (2009). Predicting subcutaneous glucose concentration in humans: Data-driven glucose modelling. *IEEE Transactions on Biomedical Engineering*, Vol. 56, No. 2, (February 2009) 246-254, 0018-9294

Gani, A.; Gribok, A. V.; Lu, Y.; Ward, W. K.; Vigersky, R. A. & Reifman, J. (2010). Universal Glucose Models for Predicting Subcutaneous Glucose Concentration in Humans. *IEEE Transactions on Information Technology in Biomedicine*, Vol. 14, No. 1, (January 2010) 157-165, 1089-7771

Kok P. (2004). Predicting blood glucose levels of diabetics using artificial neural networks, *Research Assignment for Master of Science*, Delft University of Technology

Kovatchev, B.P.; Gonder-Frederick, L.A.; Cox, D.J. & Clarke, W.L. (2004). Evaluating the accuracy of continuous glucose monitoring sensors: Continuous glucose-error grid analysis illustrated by TheraSense Freestyle Navigator data. *Diabetes Care*, Vol. 27, No. 8, (August 2004) 1922-1928, 1935-5548

Lehmann E. D. & Deutsch, T. (1992). A physiological model of glucose-insulin interaction in Type 1 diabetes mellitus. *Journal of Biomedical Engineering*, Vol. 14, No. 3, (May 1992) 235-242, 0141-5425

Makroglou, A.; Li, J. & Kuang, Y. (2006). Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: An overview. *Applied Numerical Mathematics*, Vol. 56, No. 3-4, (March 2006) 559-573, 0168-9274

Mougiakakou, S. G.; Prountzou, A.; Iliopoulou, D.; Nikita, K. S.; Vazeou, A. & Bartsocas, C. S. (2006). Neural network based glucose - insulin metabolism models for children with type 1 diabetes, *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3545-3548, 1557-170X, USA, August 2006, IEEE, New York

Quchani, S. A. & Tahami, E. (2007). Comparison of MLP and Elman Neural Network for Blood Glucose Level Prediction in Type 1 Diabetics, *Proceedings of the 3rd Kuala Lumpur International Conference on Biomedical Engineering*, Vol. 15, pp. 54-58, 978-3-540-68016-1, Malaysia, December 2006, Springer Berlin Heidelberg, Kuala Lumpur

Rollins, D.; Bhandari, N.; Kleinedler, J.; Kotz, K.; Strohbehn, A.; Boland, L.; Murphy, M.; Andre, D.; Vyas, N.; Welk, G. & Franke, W. E. (2010). Free-living inferential modelling of blood glucose level using only noninvasive inputs. *Journal of Process Control*, Vol. 20, No. 1, (January 2010) 95-107, 0959-1524

Roy, A. & Parker, R. S. (2007). Dynamic Modelling of Exercise Effects on Plasma Glucose and Insulin Levels. *Journal of Diabetes Science and Technology*, Vol. 1, No. 3, (May 2007) 338—347, 1932-2968

Smola, A. J. & Scholkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, Vol. 14, No. 3, (August 2004), 199-222, 0960-3174

Sparacino, G.; Zanderigo, F.; Corazza, S.; Maran, A.; Facchineti, A. & Cobelli, C. (2007). Glucose Concentration can be Predicted Ahead in Time from Continuous Glucose Monitoring Sensor Time-Series. *IEEE Transactions on Biomedical Engineering*, Vol. 54, No. 5, (May 2007) 931-937, 0018-9294

Stahl, F. & Johansson, R. (2009). Diabetes mellitus modelling and short-term prediction based on blood glucose measurements. *Mathematical Biosciences*, Vol. 217, No. 2, (February 2009), 101-117, 0025-5564

Tarin, C.; Teufel, E.; Pico, J.; Bondia, J. & Pfleiderer, H. J. (2005). A Comprehensive Pharmacokinetic Model of Insulin Glargine and Other Insulin Formulations. *IEEE Transactions on Biomedical Engineering*, Vol. 52, No. 12, (December 2005) 1994—2005, 0018-9294

Tresp, V.; Briegel T., & Moody, J. (1999). Neural-network models for the blood glucose metabolism of a diabetic. *IEEE Transactions on Neural Networks*, Vol. 10, No. 5, (September 1999) 1204-1213, 1045-9227

Valleta, J. J.; Chipperfield, A. J. & Byrne, C. D. (2009). Gaussian process modelling of blood glucose response to free-living physical activity data in people with type 1 diabetes, *Proceedings of the 1st Annual International Conference of the IEEE Engineering in*

Medicine and Biology Society, pp. 4913-4916, 978-1-4244-3296-7, USA, September 2009, IEEE, Minneapolis

Zainuddin, Z.; Pauline O. & and Ardil C. (2009). A Neural Network Approach in Predicting the Blood Glucose Level for Diabetic Patients. *International Journal of Computational Intelligence*. Vol. 5, No. 1, (Winter 2009) 72-79, 2070-3821

Zitar, R. A. & Al-Jabali, A. (2005). Towards neural network model for insulin/glucose in diabetics-II. *Informatica*, Vol. 29, No. 2, (June 2005) 227-232, 0350-5596

Knowledge-Oriented Applications in Data Mining

Edited by Prof. Kimito Funatsu

ISBN 978-953-307-154-1
Hard cover, 442 pages
Publisher InTech
Published online 21, January, 2011
Published in print edition January, 2011

The progress of data mining technology and large public popularity establish a need for a comprehensive text on the subject. The series of books entitled by 'Data Mining' address the need by presenting in-depth description of novel mining algorithms and many useful applications. In addition to understanding each section deeply, the two books present useful hints and strategies to solving problems in the following chapters. The contributing authors have highlighted many future research directions that will foster multi-disciplinary collaborations and hence will lead to significant development in the field of data mining.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Eleni I. Georga, Vasilios C. Protopappas and Dimitrios I. Fotiadis (2011). Glucose Prediction in Type 1 and Type 2 Diabetic Patients Using Data Driven Techniques, Knowledge-Oriented Applications in Data Mining, Prof. Kimito Funatsu (Ed.), ISBN: 978-953-307-154-1, InTech, Available from: http://www.intechopen.com/books/knowledge-oriented-applications-in-data-mining/glucose-prediction-in-type-1-and-type-2-diabetic-patients-using-data-driven-techniques

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen