2010

Free-living inferential modeling of blood glucose level using only noninvasive inputs

Derrick K. Rollins, *Iowa State University*
Nidhi Bhandari, *Iowa State University*
Jim Kleinedler, *Iowa State University*
Kaylee Kotz, *Iowa State University*
Amber Strohbehn, *Iowa State University*, et al.

Journal of Process Control 20 (2010) 95–107

Contents lists available at ScienceDirect

Journal of Process Control

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

Free-living inferential modeling of blood glucose level using only noninvasive inputs

Derrick K. Rollins [a,b,*], Nidhi Bhandari [a], Jim Kleinedler [a], Kaylee Kotz [a], Amber Strohbehn [a], Lindsay Boland [a], Megan Murphy [a], Dave Andre [c], Nisarg Vyas [c], Greg Welk [d], Warren E. Franke [d]

[a] Department of Chemical and Biological Engineering, Iowa State University, Ames, IA 50011, United States
[b] Department of Statistics, Iowa State University, Ames, IA 50011, United States
[c] BodyMedia Inc., 4 Smithfield St., Pittsburgh, PA 15222, United States
[d] Department of Kinesiology, Iowa State University, Ames, IA 50011, United States

ARTICLE INFO

Article history:
Received 18 March 2009
Received in revised form 20 August 2009
Accepted 21 September 2009

Keywords:
Diabetes
Blood glucose modeling
Wiener modeling
Block-oriented modeling
Nonlinear regression
Predictive modeling

ABSTRACT

The goal of this work is to present a causation modeling methodology with the ability to accurately infer blood glucose levels using a large set of highly correlated noninvasive input variables over an extended period of time. These models can provide insight to improve glucose monitoring, and glucose regulation through advanced model-based control technologies. The efficacy of this approach is demonstrated using real data from a type 2 diabetic (T2D) subject collected under free-living conditions over a period of 25 consecutive days. The model was identified and tested using eleven variables that included three food variables as well as several activity and stress variables. The model was trained using 20 days of data and validated using 5 days of data. This gave a fitted correlation coefficient of 0.70 and an average absolute error (AAE) (i.e., the average of the absolute values for the measured glucose concentration minus modeled glucose concentration) of 13.3 mg/dL for the validation data. This AAE result was significantly better than the subject's personal glucose meter AAE of 15.3 mg/dL for replicated measurements.
Published by Elsevier Ltd.

1. Introduction

Type 2 diabetes is approaching "epidemic" incidence in the United States, driven in large part by a remarkable rise in obesity rates over the past 15 years [1]. From 1980 through 2003, the number of Americans with diabetes has more than doubled and almost two-thirds of all Americans are now classified as either overweight or obese, a condition that puts them at a high risk for type 2 diabetes. Nearly 18.2 million people in the United States, or about 6.3% of the population, have diabetes [2]. Diabetes mellitus is a group of metabolic diseases characterized by high blood sugar (glucose) levels, which result from defects in insulin secretion, action, or both. Prolonged high glucose levels can cause damage to many areas of the body, increasing the risk of kidney failure, blindness, nerve damage, amputations, heart attack and stroke. Blood glucose (BG) levels are affected by factors such as food intake, stress, physical activity, hormonal changes, medications, illness/infection, and fatigue. All of these, as well as insulin tolerance, make managing diabetes more difficult and increase the chances of either low blood glucose (hypoglycemia) or high blood glucose (hyperglycemia). (This document will use the term "blood glucose" loosely to mean "blood glucose concentration" commonly measured in mg/dL.) The importance of tight glucose control in reducing the complications associated with diabetes is widely recognized. The primary ways that glucose has been managed include diet, exercise, stress management, insulin injections, and different types of drugs [3].

This purpose of this work is to develop a modeling methodology that accurately maps disturbances (i.e., inputs) to changes in BG levels. More specifically, this work seeks to develop *a strong positive correlation* between modeled BG and measured BG using measured input variables that fall into three *causation* classes: food, activity and stress. In addition, this approach includes modeling circadian behavior though a variable we define as the time of day (TOD) or simply, the 24 h clock. Although there are other benefits, the primary motivation for modeling these effects is to advance model-based control. Accurate knowledge of *how much* BG changes or will change from measured inputs and an accurate relationship between BG and *causation inputs* provides the framework for model-based proactive (i.e., feedforward) and predictive control. However, demonstration of improved BG control is not within the scope of this work but is a future objective of our research group. Therefore, since this scope is not to improve BG control and our primary interest is how well modeled BG correlates with measured glucose, performance measures such as *maximum* deviation is not a criterion that we will use to judge the success of this work. This type of criterion will be of importance in a future phase of our research where the control objective will be to reduce the duration and magnitude of BG beyond threshold limits.

* Corresponding author. Address: Department of Chemical and Biological Engineering, Iowa State University, Ames, IA 50011, United States. Tel.: +1 515 294 5516; fax: +1 515 294 2689.
E-mail address: drollins@iastate.edu (D.K. Rollins).

0959-1524/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.jprocont.2009.09.008

Driven by the goal of developing an effective automatic glucose control system, several researchers have been working on developing accurate predictive models [4,5] as well as using advanced control techniques like model-based controllers [6], robust tracking [7], run-to-run control [8], and carbohydrate feedforward control [9]. However, much of this work has been restricted to *in silico* or simulation studies (i.e., not using real subjects). Nonetheless, there have been a few studies which have involved real subjects [10] but the scope of this work has been limited to periods of relatively mild glucose excursions such as night time. Moreover, current compartment-based glucose modeling approaches have at most considered only two inputs – amount of glucose ingestion [11] and exercise [12,13]. In contrast, the aim considered here is to develop a modeling methodology that considers a much larger set of inputs. More specifically, the measurable input set consists of 24 inputs: three nutrient variables, 20 activity and stress-related variables, and TOD. The nutrient variables are carbohydrates, fats and proteins. TOD is measured in minutes starting at 0 (at 12:00 am) and ending at 1339 (at 11:59 pm). The data for the other 20 variables are generated by the SenseWear® Pro_3 Body Monitoring System (BodyMedia Inc., Pittsburgh, PA) worn as an armband on the upper part of the left or right arm (i.e., the triceps area). Thus, two additional challenges are the reduction of this set of variables and the accurate estimation of model parameters from such an extensive set of inputs.

This approach is further challenged by the requirement to determine the independent effects of each input (i.e., causation) over a sufficiently large input space representing normal behavior for the subject. It is not practical to obtain such an input space using designed (controlled) experimental data collection due to the long time duration and uncontrollable inputs such as emotional stress. As a result, in this study modeling data are collected under a free-living protocol. However, freely varying inputs can produce data sets with high input correlations (e.g., carbohydrates and fats typically increase and decrease simultaneously) and thus, impede the ability to obtain input causation. The approach in this paper uniquely addresses this challenge using a modification of the Wiener method developed by [14]. The Wiener network has a unique ability to address this issue because it passes each input through its own linear dynamic block allowing differences in input dynamics to break down cross correlation of inputs. A similar observation has also been reported by [15] where they have used dynamic transformations of correlated inputs (carbohydrate content of meals and insulin) and reported better models. The approach in [15] was not formally framed in the context of block-oriented modeling as it is here.

For experimentally designed data collection, where the inputs are orthogonal, the method in [14] is practical in estimating the parameters using the linear discrete-time equations as proposed. However, when inputs are appreciably correlated for models that are linear in parameters, it does not appear possible to determine parameter estimates that can give the accurate response behavior for independent input changes (i.e., a cause and effect model). This inability is because the information to estimate a parameter is correlated in the same way its associated variable is correlated with other variables. This problem of correlation between the inputs when using models that are linear in parameters has also been highlighted in [15] where they discuss how identification of AR models (and to a lesser degree ARX models) are severely hampered by correlation and can result in models with physically unreasonable gains associated with those inputs.

In contrast, models that have nonlinear response surfaces in the parameter space are not strongly affected by correlated inputs because the information to estimate parameters is not likely to be strongly correlated due to the complexity of the response surface relative to changes in parameter values. The drawback to using a model that is nonlinear in parameters is that the response surface is more restricted which means that this choice should be made carefully from as much phenomenological knowledge as possible. For plasma glucose modeling, current theoretical knowledge is so limited that a theoretical modeling approach does not appear to be practical at this time. Therefore, a semi-empirical approach has been chosen to estimate the linear dynamic blocks of the Wiener network and use nonlinear structures with time derivative behavior that have roots in the laws of conservation (what is commonly called "transfer function models"). Moreover, even though these structures are converted to discrete-time approximations, the dependence on the physically based parameters that come from the continuous-time forms are maintained to preserve the property of a nonlinear response surface in a physically based parameter space.

The proposed method is presented and evaluated in this article by first introducing the methodology in the next section. Then the details of the real subject study and application of the proposed methodology to real glucose data are discussed in Section 3. The model results, a discussion of the models, and insights gained from this work are presented in Section 4. Finally, in Section 5 the most significant results are summarized and plans for future work are discussed.

2. Modeling methodology

As stated above, the purpose of this article is to evaluate the efficacy of the proposed modeling method to infer plasma glucose concentration using noninvasive input variables, such as food components, as well as various kinds of activity and stress-related variables. Since the data are collected under free living conditions, several of these inputs are highly correlated and can have nonlinear, highly interactive, and dynamic effects on glucose level. Dynamic behavior manifests as time lags between the input changes and the output response. Time-lagged or dynamic behavior of food consumption and activity on glucose variation is well accepted by diabetic researchers. Interactive behavior, by definition, requires observing or modeling the effects of two or more variables simultaneously. For two variables, for example, interactive behavior exists when the level of one variable is critical to how changes in the other variable affect changes in the response. An interaction example in this context is the relationship between food and exercise. For example, depending on recent eating history, exercise could cause either a decrease or an increase in blood glucose. It is also important to consider nonlinear behavior because of the wide range of glucose response a diabetic person can experience in a relatively short period of time from modest changes in inputs. Therefore, the method proposed here addresses these conditions in the data and these complexities in the relationship between inputs and the response.

Accordingly, the proposed method is a unique and direct application of block-oriented Wiener modeling that extends the method developed in [14] in a novel way to address the challenges of this application. Note that the Wiener network is in the class of block-oriented modeling, an active area of research in the system identification literature [16–19] and use of a Wiener structure for modeling effects of multiple inputs in biomedical systems is becoming more common [20,21]. The block diagram for the p input, single output Wiener network is shown in Fig. 1. In this application, the inputs are the measured noninvasive variables and the output or response is glucose concentration. As illustrated, each input has its own linear dynamic block. Note that each dynamic block has an unobservable, intermediate output v_i, which represents the independent dynamic response of the input x_i. All the intermediate v_i's then pass through the static nonlinear block to produce the final output, y.

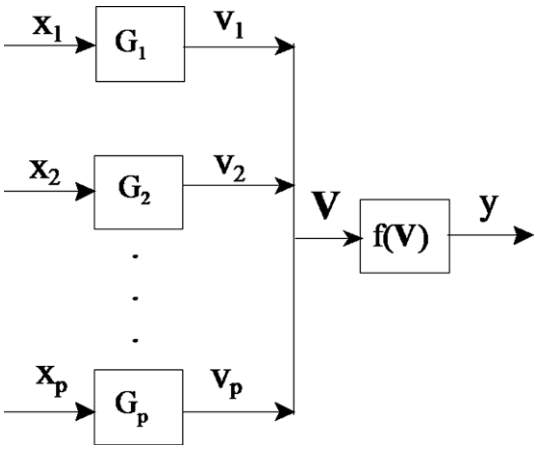

Fig. 1. The Wiener block diagram for a p input, single output system.

A critical characteristic of this approach is the derivation and strict use of functions that are nonlinear in physically interpretable parameters. This property is important to accurately map the inputs (that are highly correlated) to their intermediate dynamic counterparts, i.e., to accurately map x_i to v_i. The Wiener network is defined by its unique structure to map each input to its intermediate dynamic output variable, v_i. After collecting the v_i's and passing them through a nonlinear static block, this network is able to simultaneously address nonlinear static and nonlinear dynamic behavior. Note that, although the x_i's can be highly correlated, the v_i's are not likely to be correlated since different inputs are not likely to have similar dynamic behavior. For example, although carbohydrates and fats will likely be highly correlated since they increase and decrease together, they are likely to have different time constants (due to different time scales for digestion) and hence different dynamic behavior for the v_i's. Therefore, since the v_i's are different, they will be weakly correlated at best. Notwithstanding, given this property, the nonlinear static function can be linear or nonlinear in parameters without adversely affecting causation model development.

Second-order differential equations with first order input dynamics are used for the linear dynamic blocks. That is, all dynamic blocks used here usually have the second-order-plus-lead with dead time (SOPLDT) form given below:

$$\tau_i^2 \frac{d^2 v_i(t)}{dt^2} + 2\tau_i\zeta_i \frac{dv_i(t)}{dt} + v_i(t) = \tau_{ai} \frac{dx_i(t-\theta_i)}{dt} + x_i(t-\theta_i) \tag{1}$$

where $x_i(t)$ is the deviation variable for ith input, i varies from 1 to p, p is the total number of inputs, τ_{ai} is the lead parameter, τ_i is the time constant, ζ_i is the damping coefficient and θ_i is the dead time. The validity of using the second order differential equations given by Eq. (1) can be shown using a set of theoretical component mass balances for plasma glucose and insulin such as given in [22]. The term $2\tau_i\zeta_i$ can be thought of as a *characteristic time-scale parameter* for the ith input in this study. For conservation-based differential equations, $2\tau_i\zeta_i$ can be considered as a measure of an *effective* residence time for input variable i. Note that the dynamic behavior of any input is represented by at most four dynamic parameters giving a maximum number of $4p$ dynamic parameters.

For the static nonlinear function (i.e., the $f(\boldsymbol{V})$ block in Fig. 1), the second order regression form including interaction terms given by Eq. (2) below is being used:

$$\begin{aligned}\hat{\eta}(t) &= f(\boldsymbol{V}) \\ &= a_0 + a_1 v_1(t) + \cdots + a_p v_p(t) + b_1 v_1^2(t) + \cdots + b_p v_p^2(\boldsymbol{t}) \\ &\quad + c_{1,2} v_1(t) v_2(t) + \cdots + c_{p-1,p} v_{p-1}(t) v_p(t)\end{aligned} \tag{2}$$

where $\hat{\eta}$ is the estimated glucose concentration. More specifically, all the input variables x_i's represent deviation from an initial state under *slow* change in this context of glucose modeling. Ideally, this state would be a state of *no* change (i.e., a steady state) but this is not possible for a real subject. Note that the number of static parameters is $1 + 2p + 1 + \cdots + (p-1)$ for a total number equal to

$$N_p = 1 + 2p + \frac{1}{2}(p-1)p \tag{3}$$

For the set of parameters in Eqs. (1) and (2), solutions for $v_i(t)$ can be found under sequential step input changes.

For the proposed method, a discrete-time (DT) Wiener form was used because the oscillatory behavior of the measured glucose level was quite evident and it is desirable to allow the prediction equation to handle varying dynamics from underdamped (i.e., $0 < \zeta_i < 1$) to overdamped (i.e., $\zeta_i > 1$) behavior during parameter optimization. The approximate DT form of Eq. (1) was obtained for a sampling interval of Δt using backward difference derivatives with $v_{i,t} \approx v_i(t)$, $v_{i,t-\Delta t} \approx v_i(t-\Delta t)$ at sampled times $t, t-\Delta t, \ldots$, as

$$v_{i,t} = \delta_{1,i} v_{i,t-\Delta t} + \delta_{2,i} v_{i,t-2\Delta t} + \omega_{1,i} x_{i,t-(\Delta t-\theta)} + \omega_{2,i} x_{i,t-(2\Delta t-\theta)} \tag{4}$$

where $\omega_{2,i} = 1 - \delta_{1,i} - \delta_{2,i} - \omega_{1,i}$ to satisfy the constraint of unity gain and the following equations were obtained:

$$\delta_{1,i} = \frac{2\tau_i^2 + 2\tau_i\zeta_i\Delta t}{\tau_i^2 + 2\tau_i\zeta_i\Delta t + \Delta t^2} \tag{5}$$

$$\delta_{2,i} = \frac{-\tau_i^2}{\tau_i^2 + 2\tau_i\zeta_i\Delta t + \Delta t^2} \tag{6}$$

$$\omega_{1,i} = \frac{(\tau_{ai} + \Delta t)\Delta t}{\tau_i^2 + 2\tau_i\zeta_i\Delta t + \Delta t^2} \tag{7}$$

With $t - \Delta t$ taken as the current time, Eq. (4) estimates the value of $v_i(t)$ at the next sampling instant, t, using inputs only. Thus, strictly speaking, it is not a one-step-ahead predictor because it is not dependent on output values. That is, given values for the inputs, either from the past, present, or anticipated future (e.g., one could enter a meal that they expect to eat in the near future), Eq. (2) is able to estimate glucose response at the next time instant. The dynamic parameters associated with the ith input physically represent time-course quantities. For example, $2\tau_i\zeta_i$ gives a relative measure of the time scale over which input i affects blood glucose. Under the Wiener network, Eqs. (5)–(7) provide the capability to model the independent input effects using free-living data. The method requires that, for each $v_{i,t}$, the parameters in Eq. (4) (i.e., $\delta_{1,i}$, $\delta_{2,i}$, $\omega_{1,i}$ and $\omega_{2,i}$) be determined from the continuous-time dynamic parameters (i.e., τ_{ai}, τ_i, and ζ_i) via Eqs. (5)–(7). After obtaining an Eq. (4) for each i, the modeled glucose value is determined by substituting these results into the DT form of Eq. (2). Consequently, $\hat{\eta}$ is completely determined from measured input data ***only*** and measured glucose level is not used in its inference.

The complete *input-only glucose model* that includes the error term in additive "white" noise that is being defined generally as Model 1 is given as:

$$y_t = \eta_t + \varepsilon_t \tag{8}$$

where y_t is the observed glucose (in deviation variable) at time t, η_t is the true value with its estimate $\hat{\eta}$ given in Eq. (2), ε_t is the error term under the assumptions of independence, normality and constant variance, i.e.,

$$\varepsilon_t \sim N(0, \sigma^2) \quad \forall t \tag{9}$$

Under Eq. (9), in the absence of measured glucose concentration, the following estimate for the true value of glucose concentration under Model 1 is proposed:

$$\hat{y}_t = \hat{\eta}_t \tag{10}$$

Under infrequently measured glucose concentration, such as lancet glucose monitoring, one could use the following feedback correction equation to improve estimation under Model 1 [23]:

$$\hat{y}_t = \hat{\eta}_t + e_{t^*} = \hat{\eta}_t + (y_{t^*} - \hat{\eta}_{t^*}) \tag{11}$$

where t^* is the time of the most recently measured value from the lancet meter.

Under serially correlated noise (N_t) and "continuous glucose monitoring" (CGM) (i.e., frequently measured glucose concentration), following [24], Model 2 is defined as

$$y_t = \eta_t + N_t \tag{12}$$

with

$$N_t = \frac{\varepsilon_t}{1 - \pi_1 B - \pi_2 B^2 - \cdots} \tag{13}$$

where B is the backward shift operator, i.e. $By_t = y_{t-\Delta t}$. Under Model 2, the estimate for glucose concentration at time t is given by

$$\hat{y}_t = \hat{\eta}_t + \hat{\pi}_1 e_{t-\Delta t} + \hat{\pi}_2 e_{t-2\Delta t} + \cdots \tag{14}$$

where $e_{t-i\Delta t}$ is the residual at time $t - i\Delta t$, i.e., $e_{t-i\Delta t} = y_{t-i\Delta t} - \hat{\eta}_{t-i\Delta t}$. See [24] for a discussion on determining the estimates of π_i denoted by $\hat{\pi}_i$.

While all the predictive power based on input changes is contained in $\hat{\eta}_t$, Eq. (14) can be useful for glucose inference in closed-loop control applications with CGM to provide significantly better model accuracy, depending on the quality of the help that comes from the glucose measurements. Eq. (14) is a one-step-ahead predictor as it estimates glucose for the next time step t using current $(t - \Delta t)$ and past $(t - i\Delta t)$ glucose measurements. (A general k-steps-ahead (KSA) predictor, which is more likely to be used, is presented later in the document.) To illustrate the modeling strength of the proposed method under correlated inputs, the Jacobian or so called "derivative matrix" is called upon because the *information* for parameter estimation is observable through this matrix [25]. This matrix is represented for Model 1 in Eq. (15) below, where there is one column for each parameter and one row for each sample time in training

$$\mathbf{V} = \begin{bmatrix} \frac{\partial \eta_t}{\partial \tau_1} & \frac{\partial \eta_t}{\partial \zeta_1} & \cdots & \frac{\partial \eta_t}{\partial c_i} \\ \frac{\partial \eta_{t+1}}{\partial \tau_1} & \frac{\partial \eta_{t+1}}{\partial \zeta_1} & \cdots & \frac{\partial \eta_{t+1}}{\partial c_i} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{15}$$

Dropping the subscript t for simplicity, the partial derivative of η is derived with respect to the dynamic parameters τ_i and ζ_i for input variable i. Using Eq. (2) and differentiating give:

$$\frac{\partial \eta}{\partial \tau_i} = a_i \frac{\partial v_i}{\partial \tau_i} + 2b_i v_i \frac{\partial v_i}{\partial \tau_i} + \sum_{k \neq i} c_{i,k} v_k \frac{\partial v_i}{\partial \tau_i} = \left(a_i + 2b_i v_i + \sum_{k \neq i} c_{i,k} v_k \right) \frac{\partial v_i}{\partial \tau_i} \tag{16a}$$

$$\frac{\partial \eta}{\partial \zeta_i} = a_i \frac{\partial v_i}{\partial \zeta_i} + 2b_i v_i \frac{\partial v_i}{\partial \zeta_i} + \sum_{k \neq i} c_{i,k} v_k \frac{\partial v_i}{\partial \zeta_i} = \left(a_i + 2b_i v_i + \sum_{k \neq i} c_{i,k} v_k \right) \frac{\partial v_i}{\partial \zeta_i} \tag{16b}$$

As illustrated by Eqs. (16a) and (16b), the partial derivatives of η with respect to τ_i and ζ_i will be correlated only if the partial derivatives of v_i with respect to τ_i and ζ_i are correlated since these equations are the same except for these terms. Now from Eq. (2) (for $\theta = 0$) these derivatives, for any v_t (dropping subscript i for simplicity), are

$$\begin{aligned} \frac{\partial v_t}{\partial \tau} = & \left(\frac{2\Delta t(\tau^2\zeta + 2\tau\Delta t + \zeta\Delta t^2)}{(\tau^2 + 2\tau\zeta\Delta t + \Delta t^2)^2} \right) v_{t-\Delta t} \\ & + \left(\frac{2\tau^2 + 2\tau\zeta\Delta t}{\tau^2 + 2\tau\zeta\Delta t + \Delta t^2} \right) \frac{\partial v_{t-\Delta t}}{\partial \tau} \\ & - \left(\frac{2\tau\Delta t(\tau\zeta + \Delta t)}{(\tau^2 + 2\tau\zeta\Delta t + \Delta t^2)^2} \right) v_{t-2\Delta t} \\ & - \left(\frac{\tau^2}{\tau^2 + 2\tau\zeta\Delta t + \Delta t^2} \right) \frac{\partial v_{t-2\Delta t}}{\partial \tau} \\ & - \left(\frac{2\Delta t(\tau_a + \Delta t)(\tau + \zeta\Delta t)}{(\tau^2 + 2\tau\zeta\Delta t + \Delta t^2)^2} \right) x_{t-\Delta t} \\ & + \left(\frac{2\tau_a\Delta t(\tau + \zeta\Delta t)}{(\tau^2 + 2\tau\zeta\Delta t + \Delta t^2)^2} \right) x_{t-2\Delta t} \end{aligned} \tag{17a}$$

$$\begin{aligned} \frac{\partial v_t}{\partial \zeta} = & \left(\frac{2\tau\Delta t(\Delta t^2 - \tau^2)}{(\tau^2 + 2\tau\zeta\Delta t + \Delta t^2)^2} \right) v_{t-\Delta t} \\ & + \left(\frac{2\tau^2 + 2\tau\zeta\Delta t}{\tau^2 + 2\tau\zeta\Delta t + \Delta t^2} \right) \frac{\partial v_{t-\Delta t}}{\partial \zeta} \\ & - \left(\frac{2\tau^3\Delta t}{(\tau^2 + 2\tau\zeta\Delta t + \Delta t^2)^2} \right) v_{t-2\Delta t} \\ & - \left(\frac{\tau^2}{\tau^2 + 2\tau\zeta\Delta t + \Delta t^2} \right) \frac{\partial v_{t-2\Delta t}}{\partial \zeta} \\ & - \left(\frac{2\tau\Delta t^2(\tau_a + \Delta t)}{(\tau^2 + 2\tau\zeta\Delta t + \Delta t^2)^2} \right) x_{t-\Delta t} \\ & + \left(\frac{2\tau\tau_a\Delta t^2}{(\tau^2 + 2\tau\zeta\Delta t + \Delta t^2)^2} \right) x_{t-2\Delta t} \end{aligned} \tag{17b}$$

Thus, these results show that Eqs. (17a) and (17b) are quite different, highly complex and highly nonlinear. After substituting these equations into Eqs. (16a) and (16b), it becomes evident that the information for estimating each dynamic parameter is nearly unique (i.e., uncorrelated) since the equation for each parameter is quite different.

A similar analysis is now done for the static gain parameters. The partial derivatives with respect to these parameters (i.e., the a's, b's and c's) are given below as:

$$\begin{aligned} \frac{\partial \eta_t}{\partial a_1} = v_{1,t}, \frac{\partial \eta_t}{\partial a_2} = v_{2,t}, \ldots, \frac{\partial \eta_t}{\partial b_1} = v_{1,t}^2, \frac{\partial \eta_t}{\partial b_2} = v_{2,t}^2 \cdots, \frac{\partial \eta_t}{\partial c_{1,2}} \\ = v_{1,t} v_{2,t}, \ldots \end{aligned} \tag{18}$$

First, all of these derivatives are only a function of the v_i's, which are assumed to be weakly correlated given that they are the outputs from the linear dynamic blocks. Next, as shown by Eq. (18), the parameters associated with the linear and quadratic effects for a particular variable i are only correlated with each other, which is *within* the input variable they represent, i.e., they are *not cross correlated*. However, this is not true for the interaction terms as cross correlation is evident. Even so, one way that this can be addressed during model building is by dropping the interaction terms that adversely affect accuracy as revealed by test data analysis. Hence, from this Jacobian information analysis there is evidence that the model identification approach of the proposed approach is well equipped to address highly correlated inputs.

On the other hand, this ability is not true for model structures that are linear in parameters such as the popular class of **N**onlinear **A**uto **R**egressive **M**odels with e**X**ogenous (NARMAX) variables. The NARMAX form that corresponds to the second order Wiener model is obtained by substituting Eq. (5) into the DT form of Eq. (2) and expanding out the linear terms giving:

$$\eta_t = \alpha_1 y_{t-\Delta t} + \alpha_2 y_{t-2\Delta t} + \alpha_3 y_{t-3\Delta t} + \cdots + \sum_i (\beta_{i,1} x_{i,t-\Delta t} + \beta_{i,2} x_{i,t-2\Delta t} + \cdots) + \sum_i (\chi_{i,1} x_{i,t-\Delta t}^2 + \chi_{i,2} x_{i,t-2\Delta t}^2 + \cdots) + \sum_i \sum_j (\gamma_{i,j,1} x_{i,t-\Delta t} x_{j,t-\Delta t}) + \cdots j \neq i \quad (19)$$

with

$$V = \begin{bmatrix} \frac{\partial \eta_t}{\partial \alpha_1} & \frac{\partial \eta_t}{\partial \alpha_2} & \cdots & \frac{\partial \eta_t}{\partial \beta_{i,1}} & \cdots & \frac{\partial \eta_t}{\partial \beta_{i+1,1}} & \cdots & \frac{\partial \eta_t}{\partial \chi_{i,1}} & \frac{\partial \eta_t}{\partial \chi_{i+1,1}} & \cdots \\ \frac{\partial \eta_{t+1}}{\partial \alpha_1} & \frac{\partial \eta_{t+1}}{\partial \alpha_2} & \cdots & \frac{\partial \eta_{t+1}}{\partial \beta_{i,1}} & \cdots & \frac{\partial \eta_{t+1}}{\partial \beta_{i+1,1}} & \cdots & \frac{\partial \eta_{t+1}}{\partial \chi_{i,1}} & \frac{\partial \eta_{t+1}}{\partial \chi_{i+1,1}} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad (20)$$

where

$$\frac{\partial \eta_t}{\partial \alpha_1} = y_{t-\Delta t}, \frac{\partial \eta_t}{\partial \alpha_2} = y_{t-2\Delta t} \cdots, \frac{\partial \eta_t}{\partial \beta_{i,1}} = x_{i,t-\Delta t}, \frac{\partial \eta_t}{\partial \beta_{i,2}} = x_{i,t-2\Delta t}, \cdots, \frac{\partial \eta_t}{\partial \chi_{i,1}} = x_{i,t-\Delta t}^2, \frac{\partial \eta_t}{\partial \gamma_{i,j,1}} = x_{i,t-\Delta t} x_{j,t-\Delta t}, \ldots \quad (21)$$

Thus, the partial derivatives with respect to the α's are serially correlated as they depend on the lags of y. Likewise, the others are cross correlated as they depend linearly on correlated inputs. This analysis shows the detriment of using models that are linear in parameters when the inputs are correlated, especially when the objective is *cause and effect* modeling. Yet, another drawback is the large number of terms required by NARMAX models to address the equivalent effects of the proposed method. More specifically, for a system with p inputs, the number of NARMAX terms for equivalency to the proposed approach is $n_{\text{terms}} = 4p + 5p^2 + 2p^4$. Therefore, for 11 inputs (the number in the final model in this work), $n_{\text{terms}} = 29{,}931$, which is prohibitively large and thus not practical. A study that directly evaluates the aforementioned limitations of NARMAX is given in [26]. The data were simulated with random noise under the Model given by Eqs. (1) and (2) with two inputs. The equivalent and exact NARMAX structure under this model consisted of 60 terms. For the training data, the inputs were highly correlated and were uncorrelated for the test data set to mimic free living training data and to evaluate cause and effect modeling, respectively. Due to the limitations cited above, NARMAX performed extremely poorly.

3. The study

This section presents the details of the data collection and modeling results using data provided by a type 2 diabetic (T2D) subject – the first author of this article. Hence, it was not necessary to obtain Institutional Review Board (IRB) approval. The subject was male, 50 years old at the start of the study, type 2 diabetic, and in good health with a body mass index (BMI) of 27.9 kg/m^2. The subject was not on diabetic medication or insulin.

To obtain a sufficiently fast sampling rate necessary for discrete-time (DT) dynamic glucose modeling, the MiniMed Continuous Glucose Monitor CGMS® System Gold™ (Medtronic Minimed, Northridge, California) was used to provide glucose measurements. The glucose monitor requires the subcutaneous insertion of a sensor, typically in the torso, and assesses interstitial glucose at a reported rate of one sample every five minutes. The interstitial glucose measurements are used to infer blood glucose levels. The sensors were replaced weekly, which resulted in one to two hours of no measurements for initialization. The CGMS monitor is self-calibrating but is referenced directly to measured blood glucose values obtained four times daily from a blood glucose lancet meter. The subject in this study obtained these values from his personal One Touch Ultra® blood glucose meter (LifeScan, Inc., Milpitas, CA).

The activity measurements were obtained using the SenseWear® Pro$_3$ Body Monitoring System (BodyMedia Inc., Pittsburgh, PA). The SenseWear® Body Monitoring System generated the values for the twenty activity variables. The SenseWear® Armband, shown in Fig. 2, utilizes pattern detection algorithms [27,28] that employ physiologic signals from a unique combination of sensors. The raw physiological data include movement, heat flux, skin temperature, near body temperature, and galvanic skin response (GSR). The unit collects data using five sensors: heat flux, skin temperature, near body temperature, GSR, and a two-axis accelerometer. The heat flux sensor measures the amount of heat being dissipated from the body by measuring the heat loss along a thermally conductive path between the skin and a vent on the side of the armband. Skin temperature and near-armband temperature are also measured by sensitive thermistors. The armband also measures GSR, which is the conductivity of the wearer's skin that varies due to physical and emotional stimuli. The two-axis accelerometer tracks the movement of the upper arm and provides information about body position [28]. The SenseWear® Armband samples at a rate of one sample per minute; however, measurements at five minute intervals were used here to match the sampling rate of the glucose monitor.

Because the study is free-living, no constraints were placed on diet or lifestyle. The subject recorded the food ingested, the approximate serving sizes, the time of eating, and meal durations in a food log. In addition, the subject also obtained at least four daily measurements of blood glucose using his One-Touch glucose meter and entered them into the CGMS for calibration. During a 24 h period, the arm band device was typically removed twice; once for an hour during a time of low activity, which was most often in the evening around bed time, and once in the morning for about 30 min during showering.

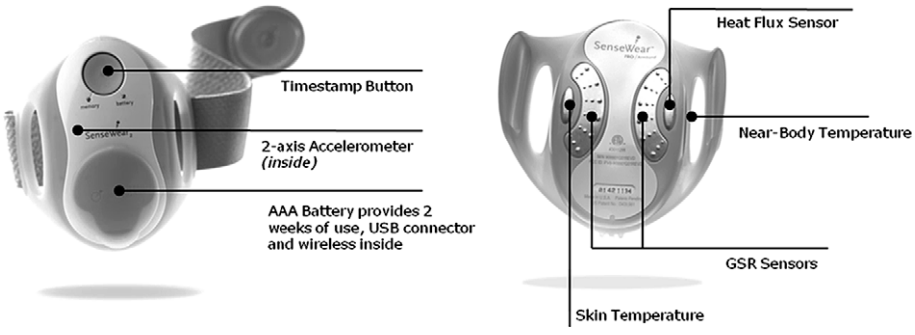

Fig. 2. The SenseWear® Armband.

The full set of input variables considered in this study is given in Table 1. As shown in this table, the total number of input variables initially considered in this study was 24. Using nutrition information from Calorie King [29], the grams of carbohydrates, fats, and proteins ingested per meal or snack were determined. Except for four variables (three nutrient variables and TOD), all other variables listed in Table 1 were reported by the SenseWear armband. Development of the final model required finding the best model that, under an input reduction procedure, minimized the number of variables and parameters while maximizing test set performance. A number of forward and backward strategies were applied to develop an optimal set of parameters for the model. A forward strategy is one that defines a minimum set of variables and adds variables sequentially based on improvements in test set performance. A backward strategy starts with the full set of variables and eliminates variables sequentially based on improvements in test set performance. The final model was determined from a forward strategy. The first input variable that was selected had the highest correlation coefficient with glucose. A second variable was selected only if it had a low correlation coefficient with the first variable and a high correlation coefficient with glucose. The final model had eleven variables: seven activity variables and four other variables as highlighted in Table 1.

All the measured variables were downloaded daily and merged into one data file on a weekly basis. The new data were then combined with the previous data files containing all the other data. For the armband, data missing only for a short period of time (less than 30 min) were interpolated with the average of the values on each side of the interval of the missing data. Data missing for longer periods were set equal to the initial steady state values. This was done to avoid setting them to very high or low values for a long period and inducing large errors. By setting them equal to initial steady state values, their deviation values become zero and the effect of those missing values is removed from the model for that period.

Table 1
The initial (24) and final (11) sets of input variables.

	Initial inputs	Final inputs
Food	**Carbohydrates**	Carbohydrates
	Fats	Fats
	Proteins	Proteins
The SenseWear® Armband	**Transverse accel – peaks**	Transverse accel – peaks
	Longitudinal accel - peaks	Longitudinal accel – average
	Longitudinal accel – average	Transverse accel – MAD
	Transverse accel - average	
	Transverse accel – MAD	
	Longitudinal accel – MAD	
	Heat flux – average	Heat flux – average
	Skin temp – average	Near-body temp – average
	Near-body temp – average	
	GSR – average	GSR – average
	Step Counter	Energy expenditure
	Energy expenditure	
	METS	
	Physical Activity	
	Lying down	
	Sleep	
	Sedentary	
	Moderate	
	Vigorous	
	Very Vigorous	
Circadian	**Time of Day (TOD)**	Time of Day (TOD)

This study uses three types of data sets. The first one is called the *training set*. The data in the training set are used to obtain the value of the optimization criterion and hence, determines directly the estimates for the model parameters. The second one is called the *validation set*. This set is used to guard against over fitting and when a *validation set* is used, it is referred to as *supervised training*. Under *supervised training*, while none of the data of the *validation set* are used in the objective function, the model performance on this data set is used to control the direction and progress towards optimality. Conversely, when the model is trained without influence from fit of another data set, the training is said to be *unsupervised*, and this set is called the *test set*.

The objective of this modeling problem is to maximize the true but unknown correlation coefficient between measured and fitted glucose concentration that is defined as $\rho_{y,\hat{y}}$ and estimated by r_{fit}. More specifically, under this objective a model is declared *useful*, if, and only if,

$$\rho_{y,\hat{y}} > 0 \tag{22}$$

The meaning of this criterion is that glucose predictions from the model decrease and increase with measured glucose concentrations beyond some degree of mere chance, i.e., there is true positive correlation. Notwithstanding, *the closer this value is to the upper limit of one (1), the more useful the model.* Therefore, to achieve this objective, one seeks to identify a model under Eq. (2) with a sufficiently large value of r_{fit}. The approach used involves splitting the data into two sets and using the training set to build the model and the testing (or validation) set to evaluate the model against data that were not *directly* used by the optimization process to estimate the model parameters. However, due to the highly complex mapping of the parameters into the response space of r_{fit}, the following indirect criterion was used:

$$\begin{aligned} &\text{Maximize } r_{\text{fit}} \text{ by Minimizing} \quad SSE_{\Theta} = \sum_{t=\Delta t}^{n\Delta t} (y_1 - \hat{y}_t)^2 \\ &\text{Subject to :} \qquad \zeta_i > 0,\ \tau_i > 0,\ \theta_i \geqslant 0 \quad \forall i \end{aligned} \tag{23}$$

As indicated, Eq. (23) was used under the assumption that minimizing *SSE* is equivalent to maximizing r_{fit}. While there is no formal proof for this assumption, experimental evidence supports a strong tendency for this relationship. In Section 4, one such graphical illustration is provided which supports the use of Eq. (23) as an effective indirect criterion for maximizing r_{fit}.

In Eq. (23), Θ is a vector representing the dynamic and static estimated parameters (i.e., τ_{ai}, τ_i, ζ_i, θ_i, for all i, and all the steady state parameters in Eq. (2)). These constraints and the physical meaning of the parameters offered advantages in setting initial guesses over using Eq. (4) directly. Optimization was carried out using the Solver tool in MS Excel®. While there is no guarantee of reaching the global minimum in this context of nonlinear regression, the strategy consisted of using several starting values to increase confidence. Depending on the size of the training data set, the optimization step could take several hours.

A model that is nonlinear in parameters, such as the proposed structure, does not have the condition that the sum of the residuals equal 0 as in the case of linear regression. However, under Eq. (23), the sum of the residuals in training should be small and thus, allowing for a secondary (or lesser) criterion on the closeness of y_t and $\hat{y}_t$, i.e., for accuracy. The measure of accuracy used here is defined as the average absolute error (AAE) and is given by Eq. (24):

$$AAE = \frac{\sum_{t=t_{\text{initial}}}^{t_{\text{final}}} |y_t - \hat{y}_t|}{M} \tag{24}$$

where M = the number of glucose measurements between t_{initial} and t_{final}. Hence, accuracy is judged to increase as AAE decreases.

Thus, in addition to sufficiently large r_{fit} values for both the training and testing/validation data sets, an acceptable model must also have a relatively small value of *AAE* in training. This secondary criterion is not imposed in testing/validation because Eq. (23) forces small residuals for training data only. Furthermore, as shown in Section 4, if a model is capable of a high r_{fit} as demonstrated in training then high accuracy can be obtained with effective calibration and adaptive procedures.

The models were developed under two different nonlinear static structures that are called the "*Reduced Model*" (RM) and the "*Full Model*" (FM). The RM consisted only of linear terms (i.e., the *a's* in Eq. (2)) in the model with the elimination of all second order terms in Eq. (2); the FM consisted of all the terms in Eq. (2). Recall from the previous section that the RM guards against a decline in accuracy due to cross correlation and extrapolation. But, its drawback is reduced accuracy when the second order static effects are significant. As will be shown momentarily, a RM trained on a 2006 data set fit a 2008 test data set quite well. The 2006 data set has 25 days data and 2008 data set has 16 days. In addition to using the full 16 days of data in 2008 as a test set to evaluate model longevity, this data set is used to evaluate model adaptation and calibration. Model adaptation is when recently collected (i.e., "on-line") data are used to estimate (i.e., "update") values of model parameters. By "calibration," it is meant that on-line data, and more specifically, lancet data in this context, are used to adjust the values determined by the model. Model adaptation is evaluated under frequent (e.g., CGMS sampling) and infrequent (e.g., lancet sampling) on-line data collection. Under frequent data collection it was found that only a small number of days were needed to update model parameters under Eq. (23) using unsupervised training. Thus, in this case, a validation set was not used and all the remaining 2008 data formed the test set. However, when using only four measurements per day for the first seven days for model adaptation, it was found necessary to perform supervised training on the remaining nine days due to the small number of samples.

All the calibration data are assumed to come from infrequent lancet data and that it is the only measured glucose data available. In this evaluation, four values per day were used. These came from the CGMS measurements that consisted of two high values and two low values spread out over the time that normal lancet measurements would be taken (e.g., not during sleep). Thus, it is assumed that the lancet meter agrees perfectly with the CGMS for the purpose of evaluation. It is reasonable to do the evaluation this way since in this study all cases are evaluated based on the agreement of the model to the CGMS data. However, in practice, a lancet meter will be used and calibration will be relative to this device. The calibration scheme consisted of two steps. The first step made an adjustment to correct for systematic bias under the assumption that previously determined deviations from measured values are random. Thus, the first correction uses the average of these previous deviations and adjusts the model as follows:

$$\hat{y}_t = \hat{\eta}_t + \bar{e}_{t^*} = \hat{\eta}_t + \sum_{i=1}^{n} (y_{t_i^*} - \hat{\eta}_{t_i^*}) n^{-1} \tag{25}$$

where t_i^* is the time of the ith glucose measurement and n is the number of measurements used. The second correction is a local correction so that the modeled glucose concentration is "adjusted" to agree with the most recent measurement and as time passes it adjusts back to the value given by Eq. (25) as described through Eq. (26)

$$\begin{aligned} &Y_t = \hat{\eta}_t + (y_{t^*} - \hat{\eta}_{t^*})^{\frac{1-t^*}{\lambda \cdot \Delta t}} \\ &\text{Subject to}: \quad t \geqslant t^*, \quad 0 < \lambda < 1 \end{aligned} \tag{26}$$

where λ is an adjustable constant. This scheme is evaluated under both validation and testing using the 2006 RM and the 2008 data sets. The final evaluation of this study examines the potential of the proposed method to improve k-steps-ahead (KSA) prediction for various values of k. The KSA prediction model form of the proposed method allows easy determination of the contributions of the inputs and outputs separately. Thus this study will reveal the importance of including inputs for predictive modeling as the prediction horizon increases.

4. Results and discussion

Table 2 gives the first set of results to be discussed. This table consists of training and validation results for both periods of data. Thus, all the cases shown here involved *supervised training*. For the duration of training given, the models in Table 2 are the best fits in validation and serve as a benchmark for comparison of other models. As shown, three cases of modeling the 2006 data are given and one case of modeling the 2008 data is given. The first three models in Table 2 are on the 2006 data. The first one was developed from 20 days of data for training while using the other 5 sequential days for validation. (Each case will be referred to by the number of training days to the number of validation or testing days, e.g., a 20/5 model would have 20 days in training and 5 days in validation or testing.) As shown, for 2006 20/5 FM, training r_{fit}, and AAE are 0.78, and 12.4 mg/dL, respectively. To give a relative context to AAE, AAE was determined from 26 pairs of replicated One Touch® Ultra blood glucose meter readings collected during this study. This value was found to be 15.3 mg/dL and significantly greater than both the 20/5 training AAE and the validation AAE of 13.3 mg/dL. For the 20/5 model, the r_{fit} values in training and validation are 0.78 and 0.70, respectively, strongly supporting the ability of the model to track glucose response using the eleven inputs. A plot of fitted and observed glucose concentration over time for this case is shown in Fig. 3 for the validation period. As these plots show, the model correlated quite well with variations in observed glucose level except for some of the most extreme hypoglycemic response and some of the higher frequency behavior.

A 20/5 evaluation was also produced for the 2006 data set with TOD removed. For this case, unsupervised training gave an r_{fit} and AAE of 0.75 and 12.9 mg/dL, respectively; a little worse than the model with TOD. The testing performance was also just slightly worse with r_{fit} and AAE of 0.67 and 15.0 mg/dL, respectively. Thus, for this data set, TOD appears to have a small but significant contribution to model accuracy. Fig. 4 shows the training progression of SSE and r_{fit} for each iteration in the minimization of SSE. As shown, as SSE decreases, r_{fit} tends to increase with each iteration

Table 2
Training and validation results for different models using data from the same period.

Data period	Model type	Model name	Training			Validation		
			# Days	r_{fit}	AAE (mg/dL)	# Days	r_{fit}	AAE (mg/dL)
2006	Full	20/5 FM	20	0.78	12.4	5	0.70	13.3
2006	Full	7/18 FM	7	0.80	11.6	18	0.64	14.5
2006	Reduced	20/5 RM	20	0.62	14.8	5	0.61	14.6
2008	Full	7/9 FM	7	0.75	13.4	9	0.57	11.8

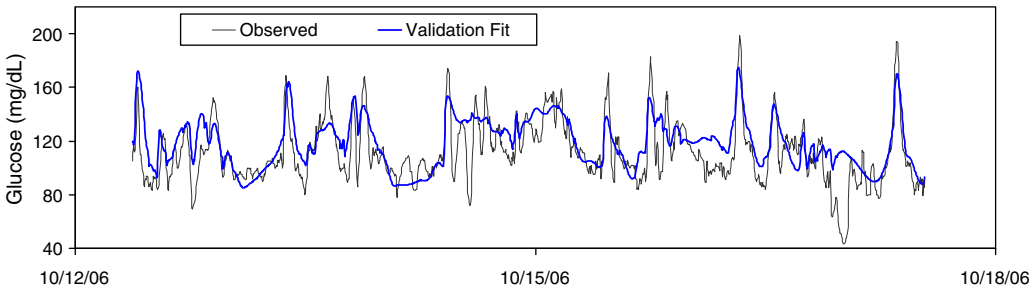

Fig. 3. The fit of Model 1 to the real glucose data in validation using the 2006 20/5 Full Model.

until leveling off in support of using the objective criterion given by Eq. (23) as a way to maximize r_{fit}.

Another informal way of analyzing the fit is by plotting measured glucose concentration by paired fitted values. This plot is statistically represented by the value of r_{fit}. More specifically, as r_{fit} increases towards 1, the data lie more along a line. The Clarke Error Grid analysis [30] is a popular procedure that divides this plot into regions or zones to aid in the diagnosis of specific types of poor performance, especially at the extremes. This plot is given in Fig. 5 for the 2006 20/5 FM validation results. As shown, there are five zones from A to E. Zone A is the region where the fitted values are within ±20% of the measured value and the model is considered sufficiently accurate. Zone B is when fitted values are more than ±20% different from measured values and serves as a caution or warning zone. Zones C, D, E are "red flag" areas because of the potential for critical misdiagnosis of hypo- and hyperglycemic glucose levels. The 2006 validation data in Fig. 5 show most of the values in Zone A (90.1%), a small number in Zone B (7.4%) and a much smaller number in Zones C–E (2.5%); thus, provides further support for the proposed approach.

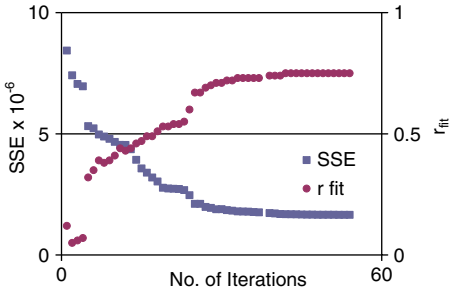

Fig. 4. The progress of SSE and r_{fit} during the minimization of SSE as the number of iterations increase for the 2006 20/5 FM and with TOD omitted as an input.

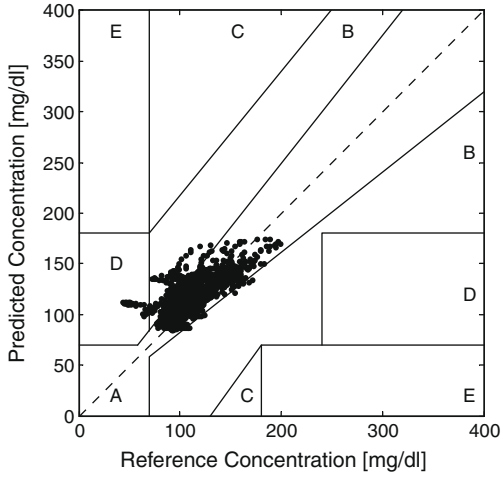

Fig. 5. Plot of Clarke Error Grid for validation data using 2006 20/5 FM.

For the 2006 data set, the correlation between carbohydrates and proteins, between carbohydrates and fats, and between fats and proteins, were 0.82, 0.87, and 0.88, respectively. These correlations were broken down as the inputs passed through the Wiener network as evidenced by the weak correlations of their corresponding dynamic variables. More specifically, the correlation between the outputs from the dynamic blocks, i.e. v_i's, for carbohydrates and proteins, carbohydrates and fats, and fats and proteins, were −0.03, 0.11, and −0.02, respectively.

The next case in Table 2, the 2006 7/18 FM case, was done to determine if an accurate model could be determined using a much smaller set of training data and to evaluate validation performance over a much longer period of time. As the results show, to no surprise, training is better for the smaller training set (i.e., 7/18) and although the validation results are slightly worse, they are still quite good. For the third case in Table 2, the 20/5 RM case, the training r_{fit} of 0.62 is much lower than the full model's value of 0.78 although AAE is still quite good at 14.8 mg/dL. The drop in performance is an indication of the necessity of modeling second order static behavior for this subject. For validation, the 2006 20/5 RM results (r_{fit} = 0.61 and AAE = 14.6 mg/dL) are closer to the 2006 7/18 FM results (r_{fit} = 0.64 and AAE = 14.5 mg/dL). Thus, although RM performance is worse, this loss might be acceptable if RM shows long-term stability and accuracy. This evaluation will be done when the results in Table 3 are reviewed.

The last case in Table 2 is the 7/9 FM for 2008. In comparison to the 2006 7/18 FM, the performance is similar with better training for 2006. In validation, the model for 2008 is better for AAE (11.8 mg/dL versus 14.5 mg/dL) but slightly worse for r_{fit} (0.57 versus 0.64). The smaller AAE with a smaller r_{fit} for 2008 is due to less spread in glucose variation over the data collection time period. Nonetheless, both cases performed well and support efficacy of the proposed approach. The 2008 FM case will serve as the benchmark for the cases that are presented in Table 3.

Table 3 contains modeling results of applying the 2006 20/5 RM to the 2008 data. The first two cases used 2008 data sampled at five minute intervals, or 12 samples per hour, to estimate the linear static coefficients while maintaining all the dynamic parameters at their 2006 20/5 RM values. In these cases the estimation protocol was to apply the least squares criterion until convergence was achieved. This was done without interruption and so, was done under *unsupervised training*. Thus, no 2008 data outside of the training periods had any influence on values of the parameters. More specifically, the 2008 data outside of the training periods were test data. This case evaluates the model adaptation potential of the proposed approach under continuous glucose sampling (CGS) rates (as in type 1 monitoring and automatic control) to adjust a minimum set of parameters (i.e., the static linear parameters only). When

Table 3
Modeling 2008 data using the 2006 20/5 reduced model.

Training				Evaluation				
Sampling rate	# Days	r_{fit}	AAE (mg/dL)	Type	Calibration	# Days	r_{fit}	AAE (mg/dL)
12/h	4	0.51	15.3	Testing	No	12	0.45	15.9
	7	0.47	17.0		No	9	0.48	12.1
All training was on 2006 Data				Testing	No	16	0.42	16.1
					Yes	16	0.60	14.1
4/day	7	0.69[a]	24.9[a]	Validation	No	9	0.48	14.8
		(0.47)	(19.8)		Yes	9	0.60	12.9

[a] The top value in this cell is for the data sampled at a rate of 4 samples per day and the bottom value (in parentheses) is for the data sampled at 12 samples per hour.

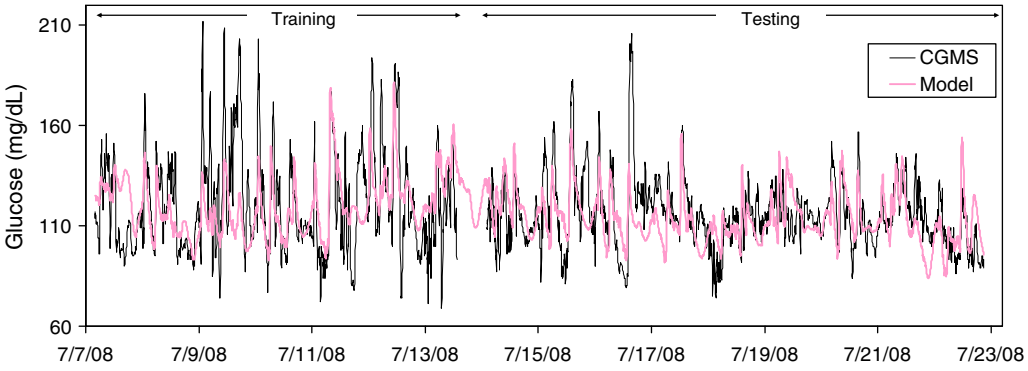

Fig. 6. Plot of 2006 RM with static linear adaptation using one week of CGMS data collected in 2008. The first 7 days are training (i.e., data before the gap) and the remaining 9 days are testing (i.e., data after the gap).

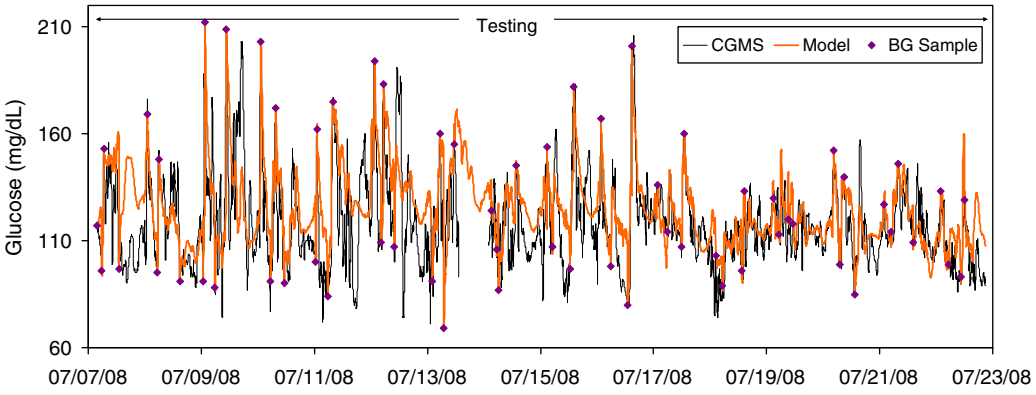

Fig. 7. Plot of 2006 20/5 RM on 2008 data with calibration using 4 samples per day under testing.

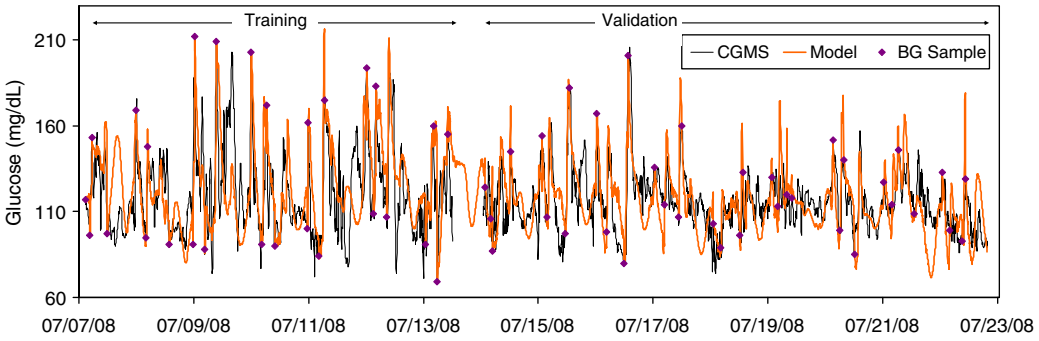

Fig. 8. Plot of 2006 20/5 RM on 2008 data with supervised adaptive training on static linear parameters and calibration using 4 samples/day of 2008 data. All dynamic parameters are at the 2006 20/5 RM values. Training is before the gap and validation is after the gap.

attempting to train the model using 1 or 2 days of data, severe over-fit resulted on the test data. Three days of training did not severely over-fit the training data at convergence but did however result in an *AAE* of 21.8 mg/dL (with a large systematic bias) and a r_{fit} of 0.43. As Table 3 shows, four days of training data greatly improved test results giving an *AAE* of 15.9 mg/dL and a r_{fit} of 0.45. The only other case considered was for seven days of training, the same number days as for the base case in Table 2. The results for this case were nearly as good as the 2008 7/18 FM base case giving an *AAE* of 12.1 mg/dL and r_{fit} of 0.48. Fig. 6 plots CGMS glucose concentration and the modeled glucose concentration for the second case in Table 3. The gap around July 14th represents missing CGMS data and the data before this gap is the training data for this case and the data after the gap is the testing data. As one can see, the model is quite stable and directionally follows GGMS glucose behavior quite well. Thus, when using the RM, the proposed approach appears to have the potential for long-term model stability and the capability of model adaptation using only a few days of CGMS data.

The next two cases in Table 3 evaluate the long-term effectiveness of the proposed approach when the 2006 20/5 RM is applied directly without any training. The difference in these two *unsupervised training* cases is in the use of infrequently sampled data for calibration using Eq. (26). The number of samples is 4/day which is the amount of data requested by the CGMS protocol for on-line calibration from the personal glucose meter. Without calibration the results were: *AAE* = 16.1 mg/dL and r_{fit} = 0.42. With calibration the results were: *AAE* = 14.1 mg/dL and r_{fit} = 0.60. The results for this case are plotted in Fig. 7. The infrequent blood glucose (BG) samples are shown by diamonds on this plot. As seen in this plot,

Table 4
The estimated dynamic parameters for the 20/5 FM for 2006 data set that are defined in Eq. (1).

Input variable	i	τ_i (min)	ζ_i	τ_{ai} (min)	θ_i (min)	$2\tau i\zeta_i$ (min)
Carbohydrates	1	30.2	1.4	−3.2	15	83.0
Fats	2	238.4	0.9	20.7	15	408.1
Proteins	3	1688.0	0.3	−73.8	15	918.9
Transverse accel – peaks	4	25.1	0.1	−13.1	0	3.4
Heat flux – average	5	91.3	0.1	5.4	0	14.4
Longitudinal accel - average	6	98.9	9.3	−37.3	0	1843.3
Near-body temp – average	7	483.1	1.0	26.9	0	976.1
Transverse accel – MAD	8	59.3	0.9	132.2	0	103.5
GSR – average	9	902.8	0.05	−157.4	0	85.1
Energy expenditure	10	264.7	2.5	−55.5	0	1316.8
Time of day	11	98.2	0.01	−0.5	0	2.0

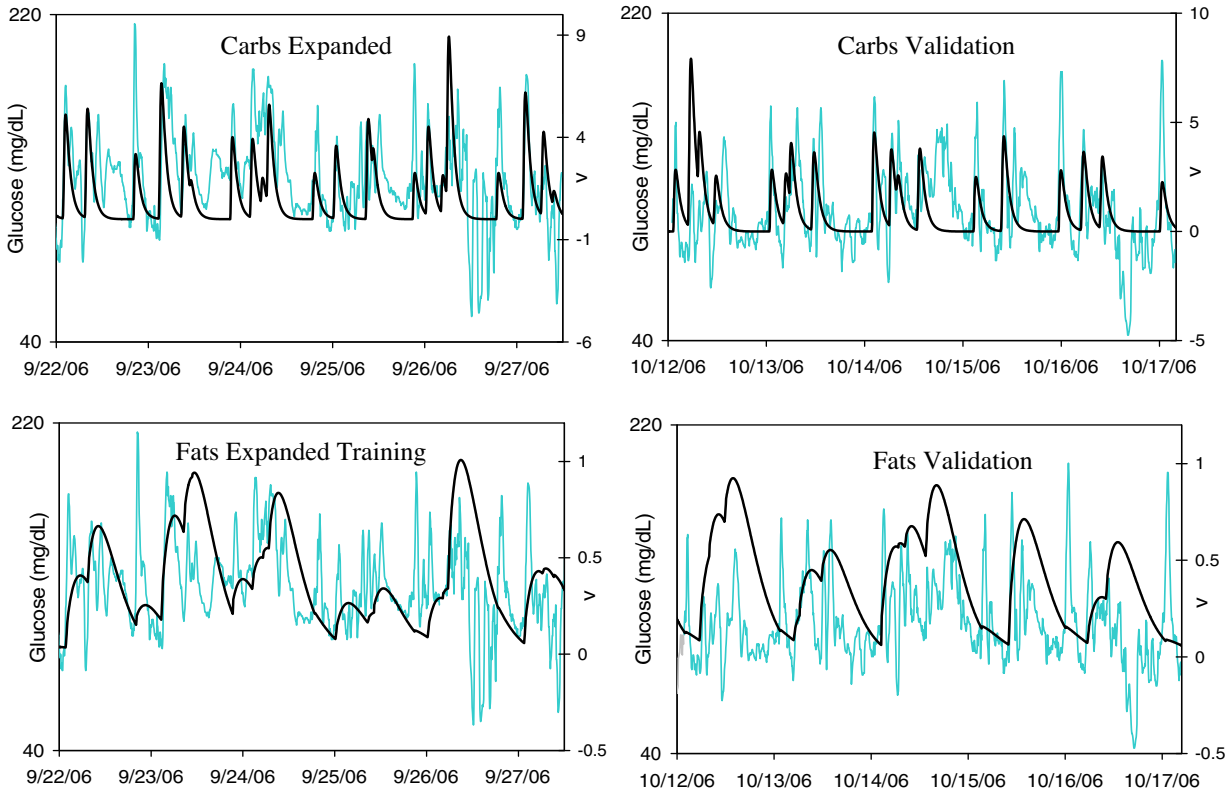

Fig. 9a. Overlay plots for carbohydrates (top) and fats (bottom) to illustrate graphically how the dynamic glucose behavior matches with the dynamic input effects. The scales on the left are for observed glucose (mg/dL) and the scales on the right are for the v_i's. The light color lines is the measured glucose and the black lines are the dynamic input effects as given by v_i's. As shown, carbohydrates have a shorter residence time than the fats with little to no overlap of peaks. The meals are clearly distinguished by non-overlapping peaks. In contrast, the longer residence time of fats indicates significant merging of effects from meal to meal. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

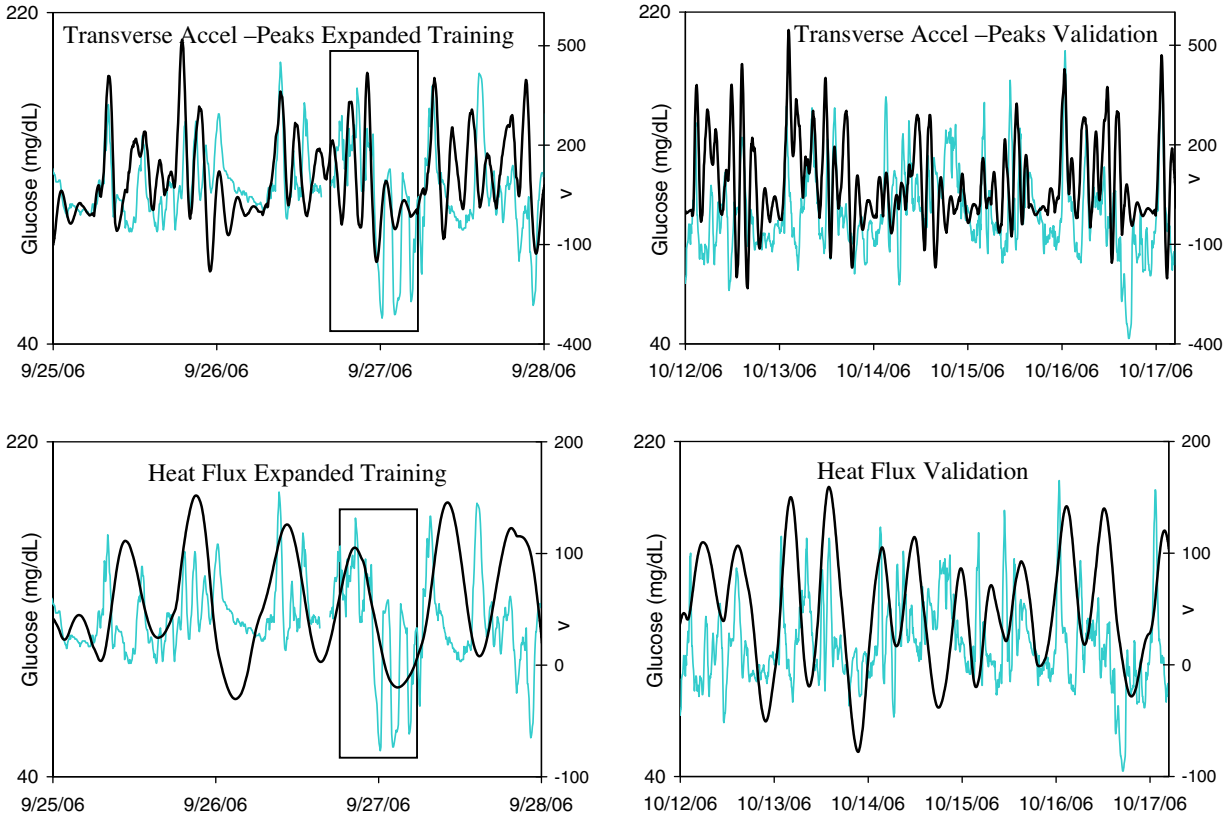

Fig. 9b. Overlay plots for transverse accel-peaks (top) and heat flux (bottom). The scales on the left are for observed glucose (mg/dl) and the scales on the right are for the v_i's. The light color lines are glucose and the black lines are the v_i's. These plots show how activity matches well with high and low components of glucose response frequencies. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

the model with calibration correlates quite well with the measured CGMS glucose concentrations. These results strongly support efficacy of the proposed approach, especially given that they were produced by a model that was almost two years old and used only 4 glucose measurements per day for calibration.

The last two cases in Table 3 evaluate adaptive estimation of the linear static parameters using only 4 samples/day for seven days (i.e., 28 samples total) starting from the 2006 20/5 RM. Since the amount of data is not sufficient for *unsupervised training*, validation sets were used to *supervise training*. Observe that the second of these two cases also used four measurements per day for calibration (using Eq. (26)) and for validation data. These two cases were done to evaluate the monitoring potential of the proposed approach using a subject's personal glucose meter data for *model adaptation and on-line calibration* since the 4 daily measurements were the only 2008 data used here. Without calibration, the nine days of validation gave an $AAE = 14.8$ mg/dL and $r_{\text{fit}} = 0.48$. With calibration, the performance improved to an $AAE = 12.9$ mg/dL and $r_{\text{fit}} = 0.60$. These results are plotted in Fig. 8. Thus, it appears the approach also has potential in noninvasive glucose monitoring using personal meter data for model adaption and calibration.

Next, the individual dynamic characteristics of each input are examined to gather insight of their effect on dynamic glucose behavior. This can be accomplished by examining the values of

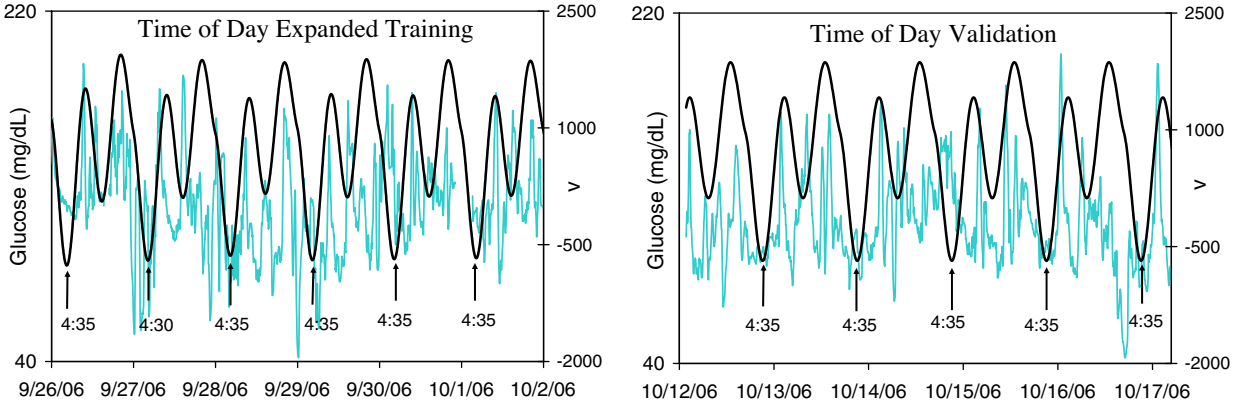

Fig. 9c. Overlay plots for Time of Day. The scales on the left are for observed glucose (mg/dL) and the scales on the right are for the v_i's. The light color lines are glucose and the black lines are the v_i's. These plots illustrate support for significant circadian effect on glucose behavior for this subject. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

the dynamic parameters and from plotting the v_i's over time. For this discussion the 2006 20/5 FM results are used. For this case the dynamic parameters are given in Table 4 and selected plots of the v's are given in Figs. 9a, 9b and 9c. These plots also contain glucose response and thus, represent individual dynamic behavior laid over its corresponding glucose behavior. As such, these plots provide insight on the dynamic contributions of the different inputs to the overall dynamics of glucose response.

First, the dynamic effects of the nutrient components are analyzed. Fig. 9a contains plots of v_i over time for the carbohydrates and fats variables (protein is not shown for space considerations) and over laid by the measured glucose values. Note that these plots have two vertical axes, one on the left for BG and one on the right for v_i. As shown in Fig. 9a, the periodic behavior of these variables is quite different and covers a very wide range. Carbohydrates (CHO) have the smallest *effective* residence time among the food variables as seen from the last column of Table 4. The overall *effective* residence times of fats and proteins are more than 4 and 11 times, respectively, greater than that for carbohydrates. These values provide a measure of the relative rates at which these three nutrients impact the blood glucose for this subject. From these plots it can be seen that the effects of carbohydrates are much faster and relatively short term. In addition, the effects of fats from different meals generally seem to overlap and their v's reach minimum levels once in a 24-h period.

Fig. 9b gives the time series plot for v_4 (Transverse accel. – peaks) and v_5 (heat flux). As shown in Fig. 9b, v_4 matches high frequency behavior of glucose quite well. It is found that v_8 (from Transverse accel. – MAD) matched the highest frequency behavior the best, but the plot has not been included for space considerations). Nevertheless v_4, even with a slightly lower frequency behavior than v_8, still matches the boxed region (representing the low glucose values changing quickly with high amplitudes) around 9/27 very well. The validity of this very low level, highly oscillatory behavior of glucose seemed questionable at first but given the excellent match with v_4, and with the food components exhibiting different dynamic behaviors, it appears to be valid.

In addition, one can also examine the plots for v_5 (heat flux) which appears to match the low level of these data well and other patterns in the data on a period of roughly a half of a day; but its periodicity does vary, especially for the validation data. The very low levels appear between 4:00–5:00 am, which were during sleep. Given that the heat flux is the amount of heat loss through the skin, it is expected that this loss would be the least during times corresponding to the longest periods of low activity. However, during sleep, other factors, such as changes in covering, room temperature, dreaming, etc., can also affect heat flux. Thus, it appears that the combination of v_4, in terms of periodicity (frequency and amplitude), and v_5 in terms of low level, validates the behavior of the boxed region and, therefore, appears to be quite useful in explaining this behavior.

The final analysis concerns the only variable that was included from observing patterns in the data – the time of day (TOD). This variable is the 24 h clock. Its dynamic output, v_{11}, is plotted in Fig. 9c. In a 24 h cycle, the pattern of v_{11} was very periodic with the minimum occurring around 4:30 am each day. The circadian rhythms of glucose and insulin in humans are well reported in literature [31,32] and the findings from this analysis also indicate that there appears to be an "internal clock" for this subject contributing to glucose behavior, especially during periods of low effects from food such as during sleep. To establish this more conclusively and widely, more subjects will need to be evaluated under this approach given its ability to obtain v_{11}. (Also for space considerations, v_6, v_7, and v_9 are not discussed or shown graphically.)

Next, one might consider the value of this approach for KSA prediction which is predictive ability k sampling periods (i.e., k times Δt) ahead from the most recent glucose measurement. Note that since the proposed method does not depend on glucose measurements but only on inputs (i.e. the models are output error models), it does not suffer from the limitations of KSA prediction; namely, the need for output measurements and their ability to impact future prediction. For this discussion, an additional model is introduced that depends *only* on glucose measurements that will be called *Model 3* and is described as:

$$\hat{y}_{t+k} = \hat{\varphi}_0 y_t + \hat{\varphi}_1 y_{t-1} + \hat{\varphi}_2 y_{t-2} + \hat{\varphi}_3 y_{t-3} + \cdots \quad (27)$$

Model 3 exploits serial correlation of current and past outputs to find optimal values for the $\hat{\varphi}'s$. Model 2 is determined using the 2006 20/5 FM and Model 3 using measured data for $k = 1$, 6 and 12. Model 2 had two additional parameters (i.e., $\hat{\theta}_1$ and $\hat{\theta}_2$) and Model 3 had four parameters (i.e., $\hat{\varphi}_0$ to $\hat{\varphi}_3$). Since for both models the parameterization was low compared to the amount of data, the impact of these models can be effectively examined through the behavior of r_{fit} for the training data. These results are given in Fig. 10 where r_{fit} is plotted against k for both models.

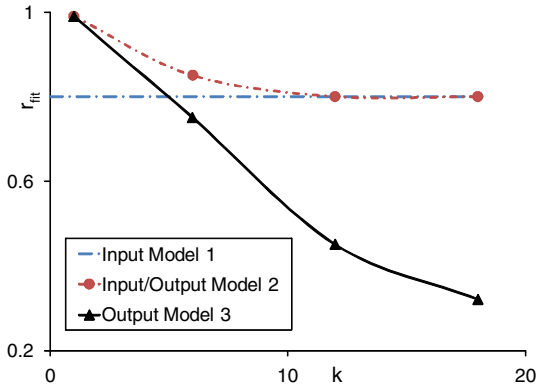

Fig. 10. r_{fit} versus k for the 2006 20/5 FM under Models 2 and 3.

As shown, r_{fit} increases as k increases for both models. Model 2 approaches Model 1 performance and Model 3 drops rapidly. Essentially, this indicates that while the outputs can aid in prediction when the number of steps ahead is small, it will not aid much when the number of steps is greater than 12 (60 min in the future). However, the results of Model 2 indicate that predictive accuracy is not limited by the size of k but by the accuracy of the input model, $\hat{\eta}_{t+k\Delta t}$. Therefore, not only can the proposed method play a key role in feedforward control but also provide critical modeling ability for model predictive control since insulin infusion can affect glucose performance much longer than one hour or for $k > 12$.

5. Concluding remarks

This article presents preliminary evidence demonstrating that it is possible to accurately model blood glucose concentration for individuals under free-living conditions for an extended period of time using *an extensive* set of food, activity and stress inputs. In actuality, these variables can be considered as disturbances in that they act to change glucose levels away from normal levels. By modeling the effects of these disturbances this work represents a critical advancement toward the goal to tighten glucose levels for insulin-dependent as well as non-insulin dependent diabetics. This accomplishment can aid considerably in understanding *how* food, activity, and stress affect glucose behavior on a broad level and, as shown here, on an individual level. In addition, it may also provide insight for researchers attempting to understand the processes and mechanisms on a more fundamental level involving invasive variables.

The success of this modeling work has been realized due to recent advancements in glucose and activity measurement technology. Continued success will depend on progress in sensor technology to measure critical variables accurately and rapidly. Preliminary work involving known insulin amounts due to insulin infusion for a type 1 diabetic has shown tremendous modeling improvements producing a much higher r_{fit} in training and validation of 0.90 and 0.96, respectively. This data set and the results can be accessed on the website: http://www.public.iastate.edu/~drollins/Research%20Data.htm. (The plan is to submit type 1 results for publication after modeling more subjects.) It seems that one critical reason for not being able to achieve values of r_{fit} greater than 0.65 for the test data is probably due to lack of information about insulin. Some of our future modeling efforts will aim to infer insulin information from other variables if possible, when the subjects are not using any external insulin.

The "closing of the loop" for automatic glucose control through an artificial pancreas usually requires the ability to predict glucose at least one to two hours ahead in the future. But as discussed in the last section, the ability to predict glucose that far ahead based on past glucose measurements is hampered by its low information content. However, as demonstrated, this critical need can be fulfilled by an input based model like the one proposed in this work provided that some accurate information about anticipated behavior is made available to the model. This can be provided in ways such as using patterns of past behavior or information entered by subjects about factors under their control such as meals and exercise.

Our future work will involve modeling more individuals to create the knowledge base to fully exploit the unique information that this approach provides. Ambitiously, our long-term goals include: 1. the development of a truly noninvasive glucose monitoring system from easily measured input sensor data; 2. the development of automatic model-based control methods – in particular, feedforward and model predictive control using multiple input modeling; and 3. providing medical workers with the knowledge via software development that will increase their understanding of the factors affecting glucose and tools for improved glucose regulation.

Acknowledgements

We gratefully acknowledge the assistance of Ai-Ling Teh and Lucas Beverlin.

References

[1] L.C. Liburd, F. Vinicor, Rethinking diabetes prevention and control in racial and ethnic communities, Journal of Public Health Management and Practice (Suppl.) (2003) S74–S79.
[2] American Diabetes Association, Economic costs of diabetes in the US in 2002, Diabetes Care 26(3) (2003) 917–932.
[3] P. Home, A. Chacra, J. Chan, A. Emslie-Smith, L. Sorenson, P.V. Crombrugge, Considerations on blood glucose management in Type 2 diabetes mellitus, Diabetes/Metabolism Research and Reviews 18 (2002) 273–285.
[4] R. Gillis, C.C. Palerm, H. Zisser, L. Jovanovic, D.E. Seborg, F.J. Doyle, Glucose estimation and prediction through meal responses using ambulatory subject data for advisory mode model predictive control, Journal of Diabetes Science and Technology 1 (6) (2007) 825–833.
[5] R. Hovorka, V. Canonico, L.J. Chassin, U. Haueter, M. Massi-Benedetti, M.O. Federici, T.R. Pieber, H.C. Schaller, L. Schaupp, T. Vering, M.E. Wilinska, Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes, Physiological Measurement 25 (4) (2004) 905–920.
[6] P. Dua, F.J. Doyle, E.F. Pistikopoulos, Model-based blood glucose control for type 1 diabetes via parametric programming, IEEE Transactions on Biomedical Engineering 53 (8) (2006) 1478–1491.
[7] E. Ruiz-Velazquez, R. Femat, D.U. Campos-Delgado, Blood glucose control for type 1 diabetes mellitus: a robust tracking H_∞ problem, Control Engineering Practice 12 (2004) 1179–1195.
[8] C.C. Palerm, H. Zisser, L. Jovanovic, F.J. Doyle, A run-to-run control strategy to adjust basal infusion rates in type 1 diabetes, Journal of Process Control (2007) 33, doi:10.1016/j.jprocont.2007.07.010.
[9] G. Marchetti, M. Barolo, L. Jovanovic, H. Zisser, D.E. Seborg, A feedforward–feedback glucose control strategy for type 1 diabetes mellitus, Journal of Process Control 18 (2008) 149–162.
[10] S.A. Weinzimer, G.M. Steil, K.L. Swan, J. Dziura, N. Kurtz, W.V. Tamborlane, Fully automated closed-loop insulin delivery versus semi-automated hybrid control in pediatric patients with type 1 diabetes using an artificial pancreas, Diabetes Care 31 (2008) 934–939, doi:10.2337/dc07-1967.
[11] Y. Zheng, M. Zhao, Modified minimal model using a single-step fitting process for the intravenous glucose tolerance test in type 2 diabetes and healthy humans, Computer Methods and Programs in Biomedicine 79 (2005) 73–79.
[12] M. Hernandez-Ordonez, D.U. Campos-Delgado, An extension to the compartmental model of type 1 diabetic patients to reproduce exercise periods with glycogen depletion and replenishment, Journal of Biomechanics 41 (2008) 744–752.
[13] P.J. Lenart, R.S. Parker, Modeling exercise effects in type I diabetic patients, in: Proceedings of the 15th Triennial World Congress, Barcelona, Spain, 2002.
[14] D.K. Rollins, N. Bhandari, Constrained MIMO dynamic discrete-time modeling exploiting optimal experimental design, Journal of Process Control 14 (6) (2004) 671–683.
[15] D.A. Finan, H. Zisser, L. Jovanovic, W.C. Bevier, D.E. Seborg, Practical issues in the identification of empirical models from simulated type 1 diabetes data, Diabetes Technology and Therapeutics 9 (5) (2007) 438–450.
[16] F.J. Doyle, R.K. Pearson, B.A. Ogunnaike, Identification and Control Using Volterra Models, Springer-Verlag London Limited, Great Britain, 2002.
[17] Y. Zhu, Estimation of an N–L–N Hammerstein–Wiener model, Automatica 38 (2002) 1607–1614.
[18] J.C. Gomez, E. Baeyens, Identification of block-oriented nonlinear systems using orthonormal bases, Journal of Process Control 14 (2004) 685–697.
[19] H.C. Park, S.W. Sung, J. Lee, Modeling of Hammerstein–Wiener processes with special input test signals, Industrial & Engineering Chemistry Research 45 (2006) 1029–1038.
[20] D.T. Westwick, R.E. Kearney, Identification of multiple input Wiener systems, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 12, No. 4, 1990, pp. 1895–1896.
[21] D.T. Westwick, E.A. Pohlmeyer, S.A. Solla, L.E. Miller, E.J. Perreault, Identification of multiple-input systems with highly coupled inputs: application to EMG prediction from multiple intracortical electrodes, Neural Computation 18 (2006) 329–355.
[22] B. Topp, K. Promislow-, G. Devries, R.M. Miuraa, D.T. Finegood, A model of ***b***-cell mass, insulin, and glucose kinetics: pathways to diabetes, Journal of Theoretical Biology (2000) 605–619, doi:10.1006/jtbi.2000.2150.
[23] D. Zhai, D.K. Rollins, N. Bhandari, H. Wu, Continuous-time Hammerstein and Wiener modeling under second-order static nonlinearity for periodic process signals, Computers & Chemical Engineering 31 (2006) 1–12.
[24] D.K. Rollins, N. Bhandari, S. Chin, T.M. Junge, K.M. Roosa, Optimal deterministic transfer function modeling in the presence of serially correlated noise, Chemical Engineering Research and Design 84 (A1) (2006) 9–21.
[25] D. Bates, D. Watts, Nonlinear Regression Analysis and its Application, Wiley, New York, USA, 1988.
[26] D.K. Rollins, N. Bhandari, K.R. Kotz, Critical modeling issues for successful feedforward control of blood glucose in insulin dependent diabetics, Proceedings of American Control Conference Seattle (2008) 832–837.
[27] D. Andre, A. Teller, Health care anywhere today, Studies in Health Technology and Informatics 118 (2005) 89–110.
[28] G.J. Welk, S.N. Blair, K. Woof, S. Jones, R.W. Thompson, A comparative evaluation of three accelerometry-based physical activity monitors, Medicine and Science in Sports and Exercise 32 (9 Suppl) (2000) S489–S497.
[29] Calorie King – Calorie Counter and Nutritional Information, <http://www.calorieking.com/foods/>.
[30] W.L. Clarke, D. Cox, L.A. Gonder-Frederick, W. Carter, S.L. Pohl, Evaluating clinical accuracy of systems for self-monitoring of blood glucose, Diabetes Care 10 (1987) 622–628.
[31] E. Van Cauter, K.S. Polonsky, A.J. Scheen, Roles of circadian rhythmicity and sleep in human glucose regulation, Endocrine Reviews 18 (1997) 716–738.
[32] A.P. Hansen, K. Johansen, Diurnal patterns of blood glucose, serum free fatty acids, insulin, glucagon and growth hormone in normals and juvenile diabetics, Diabetologia 6 (1970) 27–33.