

## Development of a Model-Based Noninvasive Glucose Monitoring Device for Non-Insulin Dependent People

Derrick K. Rollins<sup>1,2\*</sup>, Lucas Beverlin<sup>2</sup>, Yong Mei<sup>1</sup>, Kaylee Kotz<sup>1</sup>, David Andre<sup>3</sup>, Nisarg Vyas<sup>3</sup>  
Greg Welk<sup>4</sup> and Warren D. Franke<sup>4</sup>

<sup>1</sup>Department of Chemical and Biological Engineering, Iowa State University, Ames, Iowa 50011

<sup>2</sup>Department of Statistics, Iowa State University, Ames, Iowa 50011

<sup>3</sup>BodyMedia Inc., 4 Smithfield St., Pittsburgh, Pennsylvania 15222

<sup>4</sup>Department of Kinesiology, Iowa State University, Ames, Iowa 50011

### ABSTRACT

Continuous-time glucose monitoring (CGM) effectively improves glucose control, as oppose to infrequent glucose measurements (i.e. using Lancet Meters), by providing frequent blood glucose concentration (BGC) to better associate this variation with changes in behavior. Currently, the most widely used CGM devices rely on a sensor that is inserted invasively under the skin. Because of the invasive nature and also the replacement cost of sensors, the primary users of current CGM devices are insulin dependent people (type 1 and some type 2 diabetics). Most non-insulin dependent diabetics use only lancet glucose measurements. The ultimate goal of this research is the development of CGM technology that overcomes these limitations (i.e. invasive sensors and their cost) in an effort to increase CGM applications among non-insulin dependent people. To meet this objective, this preliminary work has developed a methodology to mathematically infer BGC from measurements of non-invasive input variables which can be thought of as a "virtual" or "soft" sensor approach. In this work virtual sensors are developed and evaluated on 20 subjects using four BGC measurements per day and eight input variables representing meals, activity, stress, and clock time. Up to four weeks of data are collected for each subject. One evaluation consists of 3 days of training and up to 25 days of testing data. The second one consists of one week of training, one week of validation, and 2 weeks of testing data. The third one consists two weeks of training, one week of validation and one week of testing data. Model acceptability is determined on an individual basis based on the fitted correlation to CGM testing data. For 3 day, 1 week, and 2 weeks training studies, 35%, 55% and 65% of the subjects, respectively, met the Acceptability Criteria that we established based on the concept of usefulness.

**Corresponding author:** Derrick K. Rollins, Email: [drollins@iastate.edu](mailto:drollins@iastate.edu)

**Keywords:** Virtual Sensor, Wiener Modeling,

**Running title:** Block-Oriented Modeling, Type 2 Diabetes.

**Received :** Jul 30, 2013;

**Accepted:** Dec 06, 2013;

**Published :** May 10, 2014;

## Introduction

Recent research suggests that real-time, frequent, glucose monitoring can improve blood glucose control over infrequent monitoring provided through the use of lancet glucose meters for both insulin dependent [1]-[6], and non insulin dependent diabetics [7]. Frequent glucose measurement capability is referred to as continuous-glucose monitoring (CGM); although not really continuous, current devices can deliver on-line glucose measurements as often as every one to five minutes [8]. Nonetheless, this is a substantial improvement over lancet monitoring that only produces a few values (e.g. four values) per day, at best. CGM therefore improves the user's ability to achieve better glucose control by providing frequent, real-time, glucose concentration levels that enables correlation with activity and food consumption. For example, a user is able to see with a high frequency display rate the extent to which the size of a meal affects glucose changes. Currently, the most widely used and effective CGM devices rely on a sensor that is inserted invasively under the skin. Sensors cost from \$35 to \$60 and last 3 days to a week. Thus, two significant drawbacks of these devices are comfort and cost [9]. Given these drawbacks, these devices are not widely used except by insulin dependent diabetics that rely heavily on a fast sampling rate for better control. For this reason, these devices are less likely to be used by non-insulin dependent people, including non-diabetic, pre-diabetics and diet-controlled type 2 diabetics.

Hence, the motivation of this work is the development of a useful, non-invasive, subject-specific (personalized), continuous monitoring system in an effort to increase CGM among non-insulin dependent people.

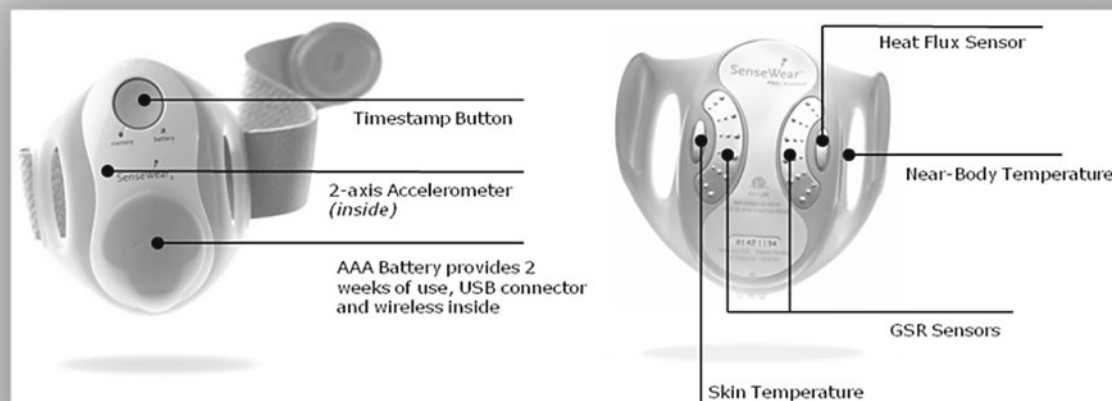
To achieve this goal we seek to develop a low maintenance, high frequency monitoring system with an accuracy that is high enough to be useful for non-insulin dependent people. Moreover, this preliminary work proposes an inferential (i.e., virtual) sensor approach for predicting blood glucose concentration (BGC) from noninvasive inputs. This virtual sensor updates at the

same rate as conventional physical sensor CGM devices. The model is developed from lancet BGC measurements that are obtained at a rate of four measurements per day. Since each sensor is calibrated from user data, the model developed for each person is said to be "subject-specific." While inferential modeling of BGC has been done by a number of researchers [19]-[24], [29], [31] particularly in type 1 diabetic applications using frequent glucose measurements, this is the first approach that we are aware of that seeks to develop an inferential model for non-insulin dependent subjects using infrequent lancet measurements from the subject's personal lancet glucose meter. Our approach to achieve this goal is to use a novel modeling method to infer glucose concentration using non-invasive input measurements for each subject from variables representing food, activity, clock time[10]-[12], and stress[13],[14].

## Methods

The main physical component of this system is a BodyMedia<sup>®</sup> armband of the type shown in Fig. 1. This device is a multi-sensor monitoring device that provides accurate estimates of physical activity data using accelerometers, heat related sensors and galvanic skin response (GSR) [15]. GSR is the conductivity of the wearer's skin that varies due to physical and emotional stimuli. For more details see [27], [28]. Given that the armband currently uses complex algorithms (e.g., for pattern recognition) it should also be able to incorporate our proposed BGC prediction algorithm. However, this research is beyond the scope of this article which is focused on the development of the modeling methodology.

The most critical challenge in this highly complex, non-linear, multiple-input, highly underdetermined modeling problem is the estimation of a large set of dynamic and static parameters from a very small set of BGC data, with a sampling frequency of only 4 values per day. To achieve accuracy under these conditions is a significant advancement over the work of Rollins et al. and a unique accomplishment. Other challenges include adequately guarding against over-fitting, the lack of initial steady state data, low quality



**Fig. 1.** The SenseWear® Armband of BodyMedia, Inc.

meal information that uses a designation of small, medium and large, and frequent and arbitrary removal of the armband monitor. Through novel modifications of the Rollins et al. [16] approach, this work demonstrates an ability to overcome these challenges, and thus, has promising potential to develop an effective inferential continuous-time BGC sensor for the target population of non-insulin dependent people. The details of the proposed modeling approach are now described.

### The Modeling Approach

The basic objective of this work is the development of a subject-specific “soft sensor” or “virtual” sensor methodology that provides “useful” information to help individuals monitor and control their glucose more effectively than with lancet glucose meters. The most critical and challenging objective in this highly underdetermined problem is that the model must be developed from a BGC sampling rate of only four samples per day. These samples will come from the lancet meter of the subject and the idea is to transform these measurements to a CGM display frequency during the period of the day that the subject is not sleeping. This virtual sensor approach is an inferential model that is developed from measured variables that are termed inputs. This virtual sensor idea has seen wide applications in process monitoring and control applications in recent years [17], [18] due to advancements in computer hardware, software, and

measurement technology. Note that to distinguish the type of sensor, i.e., “virtual” versus “physical,” we will use the terms “virtual-sensor” and “physical-sensor.” In addition, it should be noted that our use of “monitoring” include both the use of a virtual-sensor or physical-sensor although virtual sensors do not measure the process variable being monitored directly. This major challenge in this work is the frequency of BGC data for model building (in this research, 4 times per day) is much less than the virtual measurement rate of 5 minutes. This limitation means that the information available for model identification, i.e., parameter estimation, is quite limited and could thus, severely impact accuracy.

The information for the development of a virtual sensor comes from two sources -- the response data set and the input data set. Since the information content of lancet BGC is quite limited, the proposed approach strongly relies on the input data set for information on glucose behavior. More specifically, this data set consists of meal size with three levels, six (6) variables from the BodyMedia armband, and the time of day (TOD) in minutes on the 24 hour clock. The inputs that we selected for this study from the armband are those selected by Rollins et al [16]. We eliminated near body temperature as we determined it was not contributing significantly to glucose behavior for any of the subjects.

*(Continued on page 22)*

The inputs are shown in Table 1.

**Table 1.** Input variables: Meal Size (1), Armband (2-7), and TOD (8).

Input	Name
1.	Meal Size Index
2.	Transverse accel – peaks
3.	Heat flux – average
4.	Longitudinal accel – average
5.	Transverse accel – MAD
6.	GSR – average
7.	Energy expenditure
8.	Time of day (TOD)

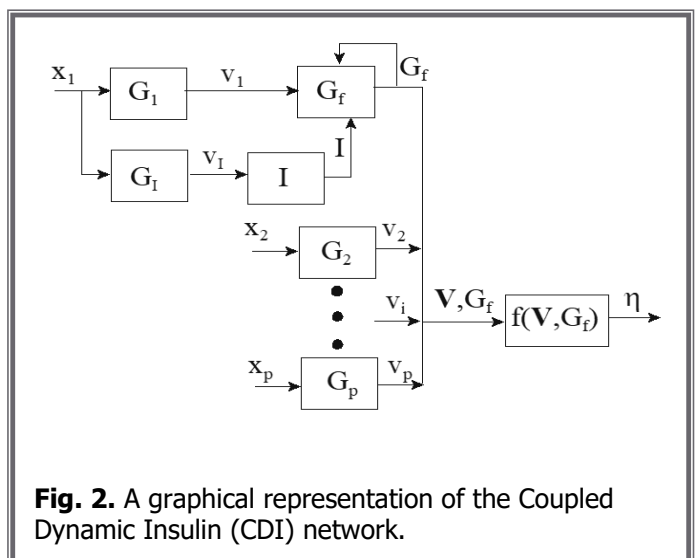
The ability to map the available input/output information to accurate sensor measurements depends on the model structure, the model building procedure, and the inferential algorithm that we are calling the “Inferential Engine.” The model structure consists of the mathematical functions and the network that tie these functions together. The model building (i.e., identification) procedure is the process of using input/output information to estimate the values of unknown parameters in the mathematical functions. The Inferential Engine is the equation used to obtain the virtual sensor measurements at the desired sampling frequency. This equation represents input selection, parameter estimates, and the use of lancet glucose measurements to enhance reliability. The purpose of this section is to describe these three components of the proposed technique in detail.

**Modeling Structure**

The modeling structure of this application must permit accurate parameter estimation under a small number of sampling times ( $n$ ), effectively handling several inputs with different dynamic behavior, and mild extrapolation. The proposed modeling network is what we call the Coupled Dynamic Insulin (CDI) network (its structure is given in Fig. 2). As shown, the first input, meal size ( $x_1$ ), enters both a linear dynamic food block ( $G_1$ ) and a linear

dynamic unmeasured insulin block ( $G_I$ ). The output from the unmeasured dynamic insulin block ( $v_1$ ) enters a pseudo blood insulin block which is coupled with the food block ( $G_f$ ) which produces the dynamic food input ( $v_1$ ) to the pseudo BGC block. Then, the unmeasured output from the coupled food block is the dynamic glucose ( $G_f$ ) input due to food consumption. Each of the other inputs (e.g. inputs 2-8) enters a separate linear dynamic block and the outputs from these blocks are collected into non-observable variables ( $v_i$ ) and together with  $G_f$  are passed through a static block which can be any type of function. The CDI model simulates the process where food digestion is responsible for the rise of blood glucose after each meal, while the secretion of insulin is responsible for the fall of blood glucose level a period of time later after the meal. The CDI network is defined by the attributes of allowing separate dynamic behavior for each input and the use of variables for unmeasured insulin generation ( $v_i$ ) and unmeasured blood insulin concentration ( $I$ ). To our knowledge, this is first application of unmeasured pseudo insulin in modeling blood glucose concentration. This idea is a key reason for the success of our modeling approach in this application of infrequent BGC measurements.

The dynamic functions for  $G_i, i = 1, \dots, p, I$  (the  $I$  is for insulin), follow the modeling work of Rollins et al. [16] and are second order differential equations of the form:



**Fig. 2.** A graphical representation of the Coupled Dynamic Insulin (CDI) network.

(Continued on page 23)

$$\tau_i^2 \frac{d^2 v_i(t)}{dt^2} + 2\tau_i \xi_i \frac{dv_i(t)}{dt} + v_i(t) = \tau_{ai} \frac{dx_i(t)}{dt} + x_i(t) \quad (1)$$

where  $x_i(t)$  is the  $i^{\text{th}}$  input,  $i$  varies from 1 to  $p$ ,  $p$  is the total number of inputs,  $\tau_{ai}$  is the lead parameter,  $\tau_i$  is the time constant, and  $\xi_i$  is the damping coefficient, with  $x_1$  = meal size input variable,  $x_i$ ,  $i = 2, \dots, p-1$ , are armband input variables, and  $v_p$  is the TOD input variable.

Using backward difference finite derivative approximations, Eq. (1) gives (Rollins et al., [16])

$$\Rightarrow v_{i,t} = \frac{2\tau_i^2 + 2\tau_i \xi_i \Delta t}{\tau_i^2 + 2\tau_i \xi_i \Delta t + \Delta t^2} v_{i,t-\Delta t} + \frac{-\Delta t^2}{\tau_i^2 + 2\tau_i \xi_i \Delta t + \Delta t^2} v_{i,t-2\Delta t} + \frac{\tau_{ai} \Delta t + \Delta t^2}{\tau_i^2 + 2\tau_i \xi_i \Delta t + \Delta t^2} x_{i,t-\Delta t} + \frac{-\tau_{ai} \Delta t}{\tau_i^2 + 2\tau_i \xi_i \Delta t + \Delta t^2} x_{i,t-2\Delta t} \quad (2)$$

with

$$v_{i,t} = \delta_{1,i} v_{i,t-\Delta t} + \delta_{2,i} v_{i,t-2\Delta t} + \omega_{1,i} x_{i,t-\Delta t} + \omega_{2,i} x_{i,t-2\Delta t} \quad (3)$$

such that  $w_{2,i} = 1 - d_{1,i} - d_{2,i} - w_{1,i}$ . This constraint is used to impose a unity gain restriction for the linear dynamic blocks.  $\Delta t$  is the sampling time for the inputs. In the Laplace domain, the linear dynamic functions are

$$G_i(s) = \frac{V_i(s)}{X_i(s)} = \frac{\tau_{ai}s + 1}{\tau_i^2 s^2 + 2\tau_i \xi_i s + 1} \quad (4)$$

Note that the number of dynamic parameters associated with each input is three. This small number is a strength that we exploit to obtain parameter estimates under limited sampling, as discussed below. The CDI model for food alone is represented by the following coupled Eqs. (5) and (6):

$$\frac{dG_f(t)}{dt} = \alpha_1 v_1(t) - \alpha_2 G_f(t) I(t) \quad (5)$$

$$\frac{dI(t)}{dt} = \alpha_3 v_I(t) - \alpha_4 I(t) \quad (6)$$

where  $v_1$  and  $v_2$  are outputs from dynamic blocks  $G_I$  and  $G_T$  respectively, and  $\alpha_1$  to  $\alpha_4$  are the "coupled" model parameters.

We also use backward difference finite derivative approximation on Eqs. (5) and (6) to give

$$G_{f,t} = \frac{\alpha_1 v_{1,t} \Delta t + G_{f,t-\Delta t}}{1 + \alpha_2 I_t \Delta t} \quad (7)$$

$$I_t = \frac{\alpha_3 v_{I,t} \Delta t + I_{t-\Delta t}}{1 + \alpha_4 \Delta t} \quad (8)$$

Note there are four additional parameters ( $\alpha_1, \alpha_2, \alpha_3$  and  $\alpha_4$ ) that need to be identified.

The function  $f(\mathbf{V})$  is called "the static function" and is a function of all of inputs. This function can theoretically be of any form. For effectiveness under mild extrapolation and minimum parameter estimation (as discussed below) we have chosen a first order linear regression model of the form:

$$y_t = \eta_t + \varepsilon_t = a_0 + G_{f,t} + a_2 v_{2,t} + \dots + a_p v_{p,t} + \varepsilon_t \quad (9)$$

Where  $\varepsilon_t$  is the error term and assumed to be independently normally distributed with mean 0 and variance  $\sigma^2$  for all  $t$ , and  $a_j$ 's are static parameters.

As stated in Rollins et al. [16], the modeling objective is simply to maximize the true but unknown correlation coefficient between measured and fitted BGC. This quantity is represented by  $\rho_{y,\hat{y}}$  and estimated by  $r_{fit}$ . Thus, under this criterion, as a minimum, a model is considered **useful**, if, and only if,

$$\rho_{y,\hat{y}} > 0 \quad (10)$$

(Continued on page 24)

Since the degree of usefulness increases with  $\rho_{y,\hat{y}}$ , the goal is to obtain the largest (as close to the upper limit of 1) value as possible. Due to the highly complex mapping of the parameters into the response space of  $r_{fit}$ , the following indirect criterion is used in obtaining the parameter estimates as described in Rollins et al. [16].

$$(11)$$

Note that only training data are used to compute SSE under Eq. (11).

### Model Identification Procedure

We use the CDI network with Eqs. (1)-(9) and developed a procedure that can accurately estimate the  $3(p+1)$  dynamic parameters, 4 coupled parameters ( $\alpha_1, \alpha_2, \alpha_3$  and  $\alpha_4$ ) and the  $p$  static parameters even when the number of sampling times ( $n$ ) is much less than  $4p + 7$ , the total number of parameters. This procedure requires each input to have a separate set of dynamic parameters as uniquely met by the Wiener network but not by other common networks (e.g. such as the Auto Regressive Moving Average with eXogenous (ARMAX) variables network) [16].

Let  $G_{i,t} = 0$  and  $v_{i,t} = 0$  for all  $i$  in Eq. (9) except for one value of  $i = j$ , i.e.,  $v_i = v_j \neq 0$ , for one value of  $i$ ,  $i = 2, \dots, p$ . Thus, with only one input variable  $v_i = v_j$ , Eq. (9) becomes a *simple linear regression model* (SLRM). To distinguish this SLM from Eq. (9), the fitted form is written as

$$\hat{y}_{i,t} = \hat{\gamma}_{0i} + \hat{\gamma}_i \hat{w}_{i,t} \quad (12)$$

Where  $\hat{y}_{i,t}$  = the fitted BGC for the one input  $i$  at  $t$ ;  $\hat{\gamma}_{0i}$  and  $\hat{\gamma}_i$  are the estimated intercept and slope parameters,

re- Maximize  $r_{fit}$  by Minimizing  $SSE = \sum_{t=\Delta t}^{n\Delta t} (y_t - \hat{y}_t)^2$

Subject to:  $\zeta_i > 0, \tau_i > 0, \theta_i \geq 0 \quad \forall i$

spectively, for the SLRM for input  $i$ ; and  $\hat{w}_{i,t}$  is the SLRM estimate of  $v_{i,t}$ .

Note that for fitting the SLRM, only five (5) parameters (the temporary static parameters  $\gamma_0$  and  $\gamma_i$ , and the permanent dynamic parameters  $\tau_i, \zeta_i$  and  $\tau_{ai}$ ) are estimated each time which, as necessary, is less than  $n = 12$  for three days of data collection, for example.

In Appendix A, a proof is given to show that for the SLRM,  $r_{fit} = r_{y_t, v_{i,t}}$ . More specifically, for the SLRM,  $r_{fit}$  is determined by only  $\hat{w}_{i,t}$  and not by the static model coefficients,  $\hat{\gamma}_{0i}$  and  $\hat{\gamma}_i$ . Thus, for the SLRM, since  $v_i$  only depends on the dynamic parameters for input  $i$ , one can find the set of dynamic parameters that results in the *best*  $r_{fit}$  for each input  $i$  separately (i.e.,  $\tau_i, \zeta_i$  and  $\tau_{ai}$ ). We exploit this result by decomposing the modeling problem into separate sub-problems that will be identified in 3 steps: 1. the dynamic parameters for each input  $i, i = 2, \dots, p$ , under Eq. (12) (five parameters are estimated for each  $i$ ); 2. the insulin and food dynamic parameters under Eqs. (7),(8), (13) (eleven parameters are estimated; one temporary intercept parameter, four coupled parameters for initial values to be used in Step 3, and six permanent dynamic parameters); and 3. the permanent static and coupled parameters with all the inputs included under Eq. (9) (at most  $p +$  four coupled parameters are estimated).

In Step 1, our current procedure is to manually adjust the dynamic parameters one input at a time to find the "best" set of values for each input. Our definition of "best" will be given momentarily. In Step 2, the following reduced form of Eq. (9) is applied:

$$y_t = \lambda_0 + G_{f,t} + \varepsilon_t \quad (13)$$

where  $\lambda_0$  is a temporary parameter only used in Step 2. This step is the most challenging. With a given set of initial values, either some or all the parameters are estimated simultaneously using an effective nonlinear regression algorithm. This process is the most iterative and time consuming as some parameters are manually

set and fixed and the rest are estimated using the optimization algorithm. This process is iteratively repeated until no more improvement can be made in  $r_{fit}$ . The only input involved in Step 2 is food, i.e., meal size. If an adequate  $r_{fit}$  ( $r_{fit}$  for training should be positive in agreement with Eq (10)) is not found in this step, the modeling procedure is terminated, and we conclude that the procedure failed to find an adequate model for this subject from the given data sets. Step 3 is completed in one estimation trial when the dynamic and couple modeling parameters from Steps 1 and 2 are used since Eq. (9) becomes a first order linear regression model. However, if one desires to estimate the couple modeling parameters also, Eq. (9) becomes a nonlinear regression model and the estimation process can be more challenging. Note that, for example, with  $n = 12$ , or three days of data collection and  $p = 8$ , at most twelve parameters are estimated in Step 3, which is still not exceeding  $n$ . At the end of all three estimation steps, with  $p = 8$ ,  $4p + 7$  or 39 total parameters have been uniquely estimated in a highly nonlinear modeling problem from at least  $n = 12$  or three days of data collection, for example. This ability is a critical novelty and a powerful benefit of this approach.

The "best" set of modeling parameters is determined for two given scenarios. The first one only uses a Training set of data. In this scenario, the goal is to maximize  $r_{fit}$  of the training set. Consequently, the procedure is to reach convergence at the global minimum for the least squares objective criterion. This estimation procedure is "unsupervised" training (note this is a different definition from that of T. Hastie's book [25] and A.J. Izenman's book [30]). The second scenario is when there are both Training and Validation sets of data. In this scenario, the goal is to determine the largest  $r_{fit}$  for the Validation data set with a "close" value of  $r_{fit}$  for the Training data set. Here, convergence for the Training set may not be reached and the Validation set determines when the iterative process terminates. Since the Validation results determine when the optimization process terminates, this is a type of "supervised" training. This procedure is used to guard against over fitting, (i.e., fitting BGC behavior in the Training set that is not due to true variation in BGC).

The success of both types of training is evaluated through the use of an additional set of data called the "test set" which had no influence on the model identification process (i.e., the parameter estimates). The first scenario is used when  $n$  is small, say 12, the number after 3 days, whereas the second scenario is used otherwise, e.g., when  $n$  is 24, the number after 6 days. Parameter estimation was done using the Excel<sup>®</sup> Solver Routine.

Successful model identification relies on effective selection of initial conditions and starting values for model parameters and the dynamic inputs (i.e., the  $v_i$ 's). The following procedure is given under a protocol where the armband is worn nearly 24 hours a day and removed only for showering. The initial steady state is chosen during a period of slow change, commonly early in the morning. The set of initial values in our procedure are  $\tau_i = 1.1$ ,  $\zeta_i = 0.9$ ,  $\alpha_2 = 20$ ,  $\alpha_4 = 0.1$ , and all other initial parameter values are equal to zero. The initial values for the  $v_i$ 's have to be determined iteratively. When the dynamic parameters are set to values so are the  $v_i$ 's as shown by Eq. (2). Our procedure is to set the initial values of the  $v_i$ 's to their average values over the training data. These values are to remain fixed during estimation and changed after estimation of dynamic parameters. The estimation process for a set of dynamic parameters is completed when the "best"  $r_{fit}$  is obtained with initial values of  $v_i$ 's close to their average values for the training data.

For the missing data due to removal of armband[32]: if data missing lasts for a short period of time (e.g. no more than one hour of missing data), the missing data were interpolated with the average value of the two sides of the missing data interval. If data missing lasts longer than one hour, we set the missing data to its initial value.

#### Development of the Inferential Engine

After obtaining a full set of parameter estimates, the proposed model development procedure has two more refinements. The first one is elimination of any armband

(Continued on page 26)

inputs that adversely affect the value of  $r_{fit}$ . This is done by setting each, and only one,  $a_i$  (for  $i = 2, \dots, 7$ ) to zero at a time, and observing  $r_{fit}$ . If  $r_{fit}$  increases for the Training set in Scenario 1 or for the Validation set for Scenario 2, this input is removed. After this process is completed for each input, all the remaining static parameters are estimated under Step 3 for a final time.

The final refinement involves the use of lancet glucose to help to reduce model bias. Since these measurements are infrequent and are not measured at a constant rate, it is not possible to build a correction model based on the correlation of residuals. The correction equation that we use comes from Rollins et al. [16] where only the most recent measurement, at  $t = t^*$ , is used. This equation, which represents the proposed virtual sensor, is given as:

$$\hat{y}_t = \hat{\eta}_t + (y_{t^*} - \hat{\eta}_{t^*}) \lambda^{\frac{t-t^*}{\Delta t}} \quad (14)$$

subject to:  $t > t^*$  and  $0 < \lambda < 1$ , where  $l$  is an adjustable constant,  $y_{t^*}$  is the lancet BGC measurement at  $t = t^*$ ,  $\hat{\eta}_t$  = the estimated BGC at time  $t$  under the Eq. (9) model,  $\hat{\eta}_{t^*}$  = the estimated BGC at time  $t = t^*$  under the Eq. (9) model, and  $\hat{y}_t$  = the virtual (i.e., soft) sensor value for the proposed method at time  $t$ . Note that  $(y_{t^*} - \hat{\eta}_{t^*})$  represents that amount of correction and this correction diminishes as time increases based on the value of  $\lambda$ , which is close to 1. Thus, by the time the next lancet measurement is taken, usually  $\hat{y}_t \approx \hat{\eta}_t$ . This means that at  $t = t^*$ ,  $\hat{y}_t \approx \hat{\eta}_{t^*}$ ; at  $t = t^* + \Delta t$ ,  $\hat{y}_t \approx y_{t^*}$ ; and for  $t = t^* + k\Delta t$ , with  $k \gg 1$  and before the next lancet measurement,  $\hat{y}_t \approx \hat{\eta}_t$ . That is, at the time of the lancet measurement, the proposed virtual monitor would display a value close to  $\hat{\eta}_t$ , the next value would be close to the lancet measurement, and as time proceeded, the lancet value would have less corrective influence as the predictor would rely more on the model to infer BGC. When two sets are used to estimate model parameters,  $\lambda$  can be set to give the most accurate

values in the validation set. When only a training set is used to estimate the model parameters, a default value can be used based on results from modeling several subjects.

### Clinical Study for 22 Subjects

For the proposed method, the development of a virtual-sensor requires 4 lancet measurements per day spread as evenly as possible over the time the subject is awake in about a 14 hour period. We did not have access to data meeting this requirement. However, from a previous study, we had physical-sensor CGM data sets which were collected with Institutional Review Board (IRB) approval, and the data sets were used to develop and evaluate the methodology. Thus, these data sets played two roles. First, for each subject, they played the role of a surrogate person, i.e., the true BGC for the purpose of evaluation. Secondly, they played the role of the lancet sampled data, i.e., the data used to build the virtual-sensors.

Using 22 test subjects (see Table 2) with 4 weeks of data collection (in most cases and slightly under 4 weeks in other cases except for Subject 1 and 8 which had only about 3 weeks of data due to loss data), we have obtained results to support the modeling viability. As just stated, these data sets were collected for another study (see Beverlin et al. [26], [32]). Modifications had to be made to these data sets for use in this study. First, food quantities, which were in grams of carbohydrates, fats and proteins, had to be converted to a food index representing meal sizes with 0 for no meal, 1 (two time stamps) for a small meal, 2 (three time stamps) for a medium meal and 3 (four time stamps) for a large meal. In practice the time stamps will be entered by the user pressing the time stamp button on the armband at the start of a meal. The conversion we used was based on the grams of carbohydrates only with less than 20 grams being a small meal, more than 100 grams being a large meal and all other amounts considered a medium size meal. Secondly, infrequent BGC measurements were not obtained from a lancet meter but converted from a

(Continued on page 27)



**Table 2.** Characteristic information on the 22 subjects used in this study.

Subject	Gender	Age	BMI	BGC Average (mg/dl)	BGC Standard Deviaion (mg/dl)
1	F	44	25.8	106.7	26.5
2	F	40	36.7	107.3	19.5
3	F	55	26.8	110.5	19.3
4	F	51	22.5	112.9	14.9
5	M	55	24.0	115.0	22.0
6	M	57	27.3	124.9	17.0
7	F	42	31.9	154.6	33.1
8	F	55	26.6	132.3	24.5
9	F	55	36.0	131.8	27.0
10	M	47	30.8	135.7	32.6
11	F	55	42.5	143.6	61.1
12	F	49	23.1	103.3	14.8
13	F	40	29.0	122.1	23.2
14	M	56	23.4	105.9	19.6
15	M	35	39.5	137.6	28.4
16	F	53	22.1	107.1	22.2
17	M	60	29.0	180.5	61.5
18	M	56	23.5	113.9	23.1
19	M	43	25.5	122.8	23.4
20	M	52	27.8	117.7	18.3
21	F	54	36.3	232.2	60.4
22	M	47	28.0	120.5	18.6

continuous glucose monitoring system (CGMS) at a sampling rate of only four values per day at particular and fixed times (i.e. only 4 values per day out of the continuous readings from CGMS were used) to mimic infrequent lancet sampling. CGMS values were taken only at 8 am, noon, 4 pm and 8 pm; if data were unavailable, then the nearest value was taken with no more than 4 values used per day. The monitoring period was taken to be from 8 am to 10 pm daily which means that this was the only period that virtual BGC were reported. Thus, the period from 10 pm to 8 am was taken to be a non-monitoring period in order to mimic that monitoring is not required during the sleeping period.

proteins. The amounts were calculated from self reporting logs of the type and quantities of food eaten. Hence, the errors of these quantities are likely quite high at times and it is likely that a significant number of meals were not recorded or logged at the proper times. When we converted the quantities to an index value for meal size for this study (i.e. "1" represents small meal size, "2" for medium size, and "3" for large size), we applied the same conversion equation to all of the subjects. Thus, the quality of food information that we developed our models from in this study is quite poor. Therefore, since these results are obtained under poor food information they indicate the robustness of the technique to low quality food information.

Note that, the original data sets contain meal information in terms of grams of carbohydrates, fats and

(Continued on page 28)

Before evaluating model acceptability, subject 21 and 22 were rejected due to poor food information. As a result, subject 21 and 22 were removed from this research from this point on.

### Measures of Performance

Model acceptability will be determined on an individual subject basis given that the models are subject-specific and each individual will only be concerned about model accuracy as it pertains to model developed for them. Thus, this study is evaluated based on the number of subject-models that meet a particular Acceptability Criteria. But we state and justify this criteria momentarily, after we present the statistics they it uses.

The first one is called the averaged error (AE) and is simply the average value of the residuals:

$$AE = \frac{\sum_{i=1}^m y_i - \hat{y}_i}{m} \quad (15)$$

where  $m$  is the number of terms being averaged.

The second one is called the averaged absolute error and is similar to Eq. (15) except that the absolute difference is used for the term in the summation as follows:

$$AAE = \frac{\sum_{i=1}^m |y_i - \hat{y}_i|}{m} \quad (16)$$

A scaled AAE value to adjust for spread is used called the relative AAE (RAAE). This measure of performance is determined by dividing Eq. (16) by the standard deviation of the values used to calculate AAE as follows:

$$RAAE = \frac{AAE}{\sqrt{\frac{\sum_{i=1}^m (y_i - \bar{y}_i)^2}{m-1}}} = \frac{AAE}{stdev(y_i)} \quad (17)$$

RAAE is a relative AAE statistic that accounts for large spread in the glucose variation of subjects. For replicated lancet measurements, the study in Rollins et al. [16] determined RAAE to be about 0.60. Thus, we will assume that a value around 0.60 is comparable to the performance of a glucose lancet meter. However, lancet accuracy or even repeatability can vary widely from individual to individual due to the accuracy of the device, inherent variability in the measurement protocol, and human error. Nonetheless, since this is the only result that likely exists in this type of study (i.e., four weeks of data collection under the protocol of this study) we will use it in our criteria. It is also noted that, given that the models in this study are developed from three discrete levels of food size and not from the three types of consumed quantities, and from a much lower frequency of glucose data when comparing to typical CGMS values (i.e., four values per day versus 12 values per hour), we expected RAAE to be higher and allow for slightly higher values in the Acceptability Criteria.

The last statistic or performance measure is  $r_{fit}$ . Based on the results in Rollins et al. [16] for a type 2 diabetic and in Beverlin et al. [26] for the 20 subjects used here, we set a minimum acceptable value for  $r_{fit}$  of 0.40. Using these three measures of performance, the Model Acceptability Criteria (MAC) for this study is given as:

$$r_{fit} = \begin{cases} 0.6 \\ or \max(0.4, 0.05 AAE - 0.4) \\ or \max(0.4, RAAE - 0.2) \end{cases} \quad (18)$$

As the MAC shows, a model with an  $r_{fit}$  of at least 0.6 is considered acceptable based on this value alone. However, if this value is between 0.4 and 0.6, the fitted

model must meet a certain level of performance on both  $r_{fit}$  and AAE or both  $r_{fit}$  and RAAE. For example, if AAE and RAAE are 18 mg/dL and 0.75, respectively, the AAE sub-criterion is  $r_{fit} \geq 0.5$  and the RAAE one is  $r_{fit} \geq 0.55$ . Thus, for this example, the fitted model will meet the MAC if, and only if,  $r_{fit} \geq 0.5$ . As this example illustrate, when  $r_{fit}$  is in this range, the MAC is written to require  $r_{fit}$  values to be higher than 0.4 with this requirement increasing as AAE and RAAE increase. Note that on the upper boundary of  $r_{fit} = 0.6$ , AAE must be  $< 20$  mg/dL or RAAE must be  $< 0.8$ , and on the lower boundary of  $r_{fit} = 0.4$ , AAE  $< 16$  mg/dL or RAAE  $< 0.6$ . The lower boundary was defined based on the results in Rollins, et al. [16] and the upper boundary was established from an examination of fitted models in this work. See Fig. 3 for plots of three fitted models as they compare with CGM measured responses at the limits and middle of the MAC.

## Results

The results of this study are given in Tables 3-5. Each table represents a different training period. There are two types of predictions in these tables;  $\hat{\eta}_{t^*}$ , the values of the fitted model at the time the lancet measurements were taken and  $\hat{y}_t$ , fitted model at the sampling rate of the CGMS, i.e., every five minutes. The model parameters are estimated using  $\hat{\eta}_{t^*}$ , but the MAC is applied to the testing results for  $\hat{y}_t$  to determine model acceptability on an individual basis which is shown in the tables. In addition, summary results for the performance measures are given for all the subjects and for the set of subjects meeting the MAC.

Table 3 gives the modeling results are for three (3) days of training under Eq. (13) (i.e., for food only). Since training stop at convergence under the least squares criterion given by Eq. (11), the remaining days consisted of the test set. In addition, for all these subjects,  $\zeta_1 = 0.2$ ,  $\tau_{a1} = 0$ ,  $\zeta_I = 0.8$ , and  $\tau_{aI} = 0$ . This was done to increase the degrees of freedom to estimate the more critical parameter  $\tau_1$ ,  $\tau_I$  and the four coupled parameters and to simplify the optimization. The best choice for these values is future research work. As Table 3 shows, the results indicate that 35% of the

cases met the MAC. This is really quite promising as a minimum initial calibration period given that the number of data points,  $n$ , used is only 12. In practice, it appears that a significant number of subjects could have successful calibration after three days and as more data are collected this number would grow. This conclusion is supported by the results in the next two tables.

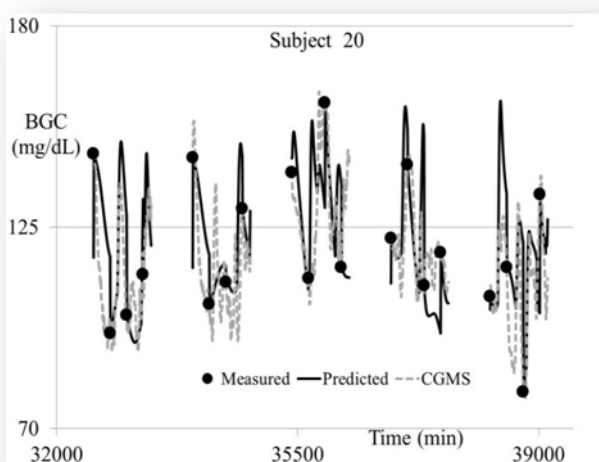
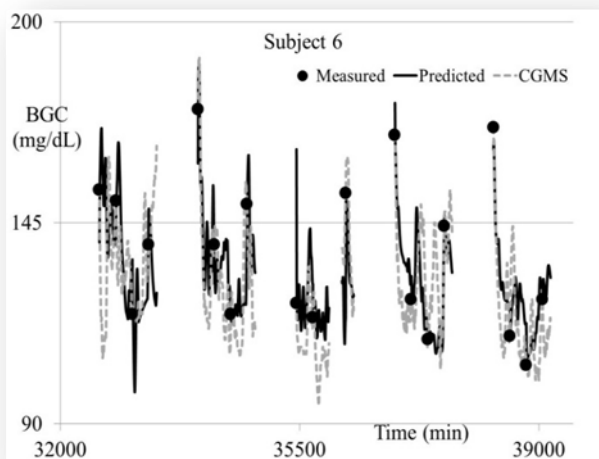
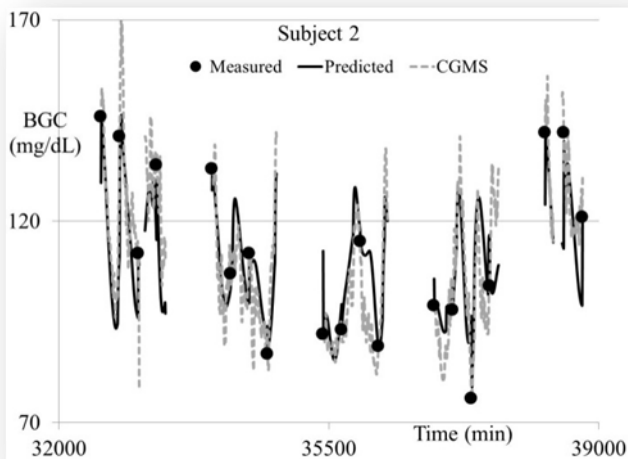
Table 4 contains results under Eq. (9) for one week of training, one week of validation and two weeks of testing. As shown, 55% of these cases met the MAC. In addition, for this group that meets MAC versus all the cases in Table 4 as whole, the average values of AAE and RAAE dropped considerably from 19.8 mg/dL and 0.76 to 13.5 mg/dL and 0.71, respectively, while  $r_{fit}$  increased from 0.47 to 0.55. These values are excellent. Table 5, also under Eq. (9), contains results for two weeks of training, one week of validation and one week of testing. As shown, the number meeting the MAC further to 65% (it is a promising result given the strict MAC) with very good average results for this group with AAE = 15.0 mg/dL, RAAE = 0.74 and  $r_{fit} = 0.54$ .

We found that using the armband inputs increases  $r_{fit}$  for  $\hat{\eta}_{t^*}$  by 0.1 over using just food alone. (These cases are not shown for space considerations). Thus, both food and the armband inputs are to obtain the results presented in this section. The robustness to poorer food quality is supported by similar  $r_{fit}$  values for this study as compared to the ones in Beverlin et al. [26] and Beverlin [32] where food quantities were used on these same data sets.

## Concluding Remarks

This article presented preliminary work on the development of a virtual sensor for BGC with the objective of developing a noninvasive CGM system that could increase CGM among non-insulin dependent people. This device would require users to wear a readily available armband monitor and manually entering meal sizes through the use of a button on the armband. This device would require four (4) lancet measurements per day as most current invasive CGMS

(Continued on page 30)



**Fig. 3.** Graphical examples during the testing period for three subjects meeting the MAC: Subject 2 (strongly meeting the MAC); Subject 6 (weakly meeting the MAC) and; Subject 20 (moderately meeting the MAC).

require.

The modeling methodology presented in this work is quite powerful. It takes on the challenge of modeling BGC in a highly complex, non-linear, multiple-input, highly underdetermined problem. As illustrated in this work, it is able to develop useful multiple-input dynamic models for BGC under free living, outpatient, data collection from just four glucose measurements per day and from as little as three days of data. In addition, these results are achieved with minimal food information of only three discrete levels. This ability stems from a number of innovative ideas to overcome several challenges in this complex modeling problem as follows. First, the use of the coupled structure allows for the inclusion of inferential blood insulin concentration and leads to insulin and glucose interaction in the blood. This structure is a significant advancement over a straight Wiener network and contributes significantly to the accuracy and ability to obtain adequate fitting for acceptable model usefulness. Secondly, the result in Appendix A provided the knowledge that produced the idea to decompose the modeling problem into a dynamic part and a static part. Added to this idea is the inspiration of determining the dynamic parameters for each input, one input at a time. Once the dynamic parameters are determined for each input, they are fixed. Note that, from the use of a validation set we are able to control over-fitting and by controlling  $r_{fit}$  to be about the same in the training set and validation set for each input separately, we have found that this helps the final  $r_{fit}$  in all the data sets (Training, Validation, and Testing) to be quite similar. After obtaining the dynamic parameters, the low number of static parameters is then obtained separately as a linear regression model. Thirdly, as the results show, the correction provided by Eq. (14) contributes strongly to the accuracy of the proposed method in the case of continuous glucose monitoring (CGM). While this correction does contribute significantly to the reduction in bias, it also contributes majorly in the reduction in AAE. This can only occur if there is a significant positive correlation for the fitted response, as the correction brings it

close at the times infrequent measurements occur, but the correlation determines the direction from these points. If the correlation is not positive, the trend will be in the wrong direction and accuracy would suffer tremendously. This is why subjects must have a significant positive correlation for acceptability. In practice if this is not achieved, the device would simply give a calibration error and not report measurements. Lastly, we developed the new Model Acceptability Criteria (MAC) which instead of evaluating separate aspects of performance such as correlation or bias, and is able to evaluate all of them and also can give a summary statistics (MAC Passing Rate) on the sample

population.

This work applied and improved the methodology from Rollins et al. [16]. It was not the purpose of this work to improve on this approach by investigating the impact of other armband variables, or the time variant nature of model parameters, as well as other model improvement issues. To develop the best model for a specific subject, these issues could be considered but they add to the modeling overhead, which is already very high given the small amount of information. The value of this work lies in that the modeling methodology shows great potential in

**Table 3.** Modeling results of 3 days training and up to 25 days of testing data

Subject	$\hat{\eta}_{t^*}$								$\hat{y}_t$				Meeting Criteria
	Training				Testing				Testing				
	AE	AAE	RAAE	r <sub>fit</sub>	AE	AAE	RAAE	r <sub>fit</sub>	AE	AAE	RAAE	r <sub>fit</sub>	
1	0.0	11.6	0.48	0.78	6.1	35.3	1.32	0.03	5.9	39.2	1.38	0.23	No
2	0.0	10.9	0.55	0.64	-10.6	21.7	0.85	0.33	-3.8	16.4	0.79	0.41	No
3	0.0	7.5	0.66	0.47	-11.2	15.9	0.88	0.40	-3.5	12.1	0.69	0.53	Yes
4	0.0	6.8	0.45	0.79	0.7	19.9	1.67	0.19	1.6	17.8	1.36	0.28	No
5	0.0	10.2	0.57	0.63	-3.3	18.6	0.77	0.22	-0.6	16.1	0.77	0.50	Yes
6	0.0	14.7	0.99	0.30	-11.5	19.9	0.97	0.17	-12.9	18.0	1.08	0.24	No
7	0.0	20.2	0.51	0.76	9.7	27.7	1.14	0.20	-2.2	37.3	1.38	0.20	No
8	0.0	6.4	0.51	0.78	-11.8	25.2	0.83	0.23	-7.0	17.2	0.66	0.60	Yes
9	0.0	14.0	0.57	0.68	-24.5	30.6	1.29	0.24	-11.3	25.7	0.90	0.30	No
10	0.0	15.0	0.58	0.51	-32.5	39.4	1.19	0.27	-17.3	33.0	0.91	0.33	No
11	0.0	38.5	0.64	0.46	-15.6	51.7	0.93	0.29	-0.6	51.0	0.93	0.35	No
12	0.0	9.2	0.70	0.54	3.3	11.7	0.86	0.20	-0.2	12.4	0.88	0.27	No
13	0.0	20.5	0.75	0.35	10.7	19.9	0.72	0.27	7.0	15.8	0.69	0.41	Yes
14	0.0	19.3	0.69	0.58	-21.3	27.7	1.21	0.20	-16.5	21.7	1.05	0.39	No
15	0.0	33.3	0.88	0.03	-7.9	27.2	0.90	0.36	-12.6	29.5	0.99	0.33	No
16	0.0	9.0	0.71	0.04	7.1	14.6	0.63	0.22	1.2	12.9	0.51	0.58	Yes
17	0.0	37.5	0.45	0.86	13.4	91.0	1.51	0.39	9.8	85.4	1.37	0.17	No
18	0.0	19.9	0.70	0.33	-13.6	21.4	1.01	0.18	-9.7	15.9	0.84	0.44	Yes
19	0.0	14.8	0.55	0.65	-5.8	16.2	0.74	0.37	2.0	16.6	0.69	0.43	Yes
20	0.0	10.3	0.56	0.78	-3.4	17.1	0.88	0.27	-5.4	17.2	0.96	0.39	No
Mean	0.0	16.5	0.62	0.55	-6.1	27.6	1.01	0.25	-3.8	25.6	0.94	0.37	Criteria Passing Rate
Stdev	0.0	9.7	0.14	0.24	12.4	17.7	0.27	0.09	7.7	17.5	0.26	0.12	
For cases meeting the criteria													
Mean	0.0	12.6	0.64	0.47	-4.0	18.8	0.80	0.27	-1.5	15.2	0.69	0.50	35%
Stdev	0.0	5.8	0.09	0.25	9.5	3.7	0.12	0.08	5.7	1.9	0.10	0.07	

modeling BGC, and provided powerful tools for statistical learning in real life scenario. Nonetheless, these are issues that can be addressed in future research.

Future work will involve running clinical studies

sufficient number of lancet measurements have been collected, the model will be built from these data automatically for calibration of the device. After successful calibration, the armband will collect input data, infrequent output data, and display BGC

**Table 4.** Modeling results of 1 week training, 1 week validation and 2 weeks testing data

Subject	$\hat{\eta}_{t^*}$												$\hat{y}_t$				Meeting Criteria
	Training				Validation				Testing				Testing				
	AE	AAE	RAAE	$r_{fit}$	AE	AAE	RAAE	$r_{fit}$	AE	AAE	RAAE	$r_{fit}$	AE	AAE	RAAE	$r_{fit}$	
1	-1.0	14.9	0.64	0.60	-5.8	17.3	0.80	0.42	7.9	24.3	0.74	0.40	10.3	23.4	0.77	0.39	No
2	-1.1	12.1	0.62	0.60	-2.0	12.7	0.57	0.63	7.7	17.6	0.65	0.54	1.6	11.9	0.54	0.72	Yes
3	-2.8	8.6	0.59	0.64	-2.5	11.3	0.64	0.57	13.1	15.7	0.94	0.45	2.1	10.7	0.65	0.60	Yes
4	-0.8	9.0	0.69	0.42	0.6	7.8	0.71	0.47	7.2	11.1	0.93	0.32	6.8	9.8	0.73	0.50	Yes
5	2.9	16.1	0.63	0.50	-4.2	18.8	0.71	0.37	1.6	15.6	0.75	0.33	-2.1	13.4	0.72	0.54	Yes
6	-0.5	17.2	0.72	0.45	-10.1	14.7	1.10	0.36	-7.8	15.8	0.81	0.50	-7.3	15.0	0.93	0.31	No
7	0.3	25.2	0.74	0.42	1.9	13.7	0.60	0.52	2.6	15.1	0.71	0.51	0.0	23.0	0.84	0.23	No
8	0.0	13.7	0.67	0.38	6.1	23.6	0.71	0.44	0.0	22.0	0.72	0.40	-3.9	16.5	0.61	0.61	Yes
9	0.2	16.3	0.67	0.63	-2.8	14.4	0.70	0.68	-0.3	18.6	0.70	0.49	7.4	22.9	0.72	0.41	No
10	1.3	15.2	0.40	0.80	-4.8	23.2	0.69	0.49	-7.9	24.1	0.69	0.53	8.4	28.9	0.78	0.43	No
11	0.0	42.3	0.65	0.55	-16.1	37.8	0.97	0.54	-24.3	45.3	0.80	0.43	-3.8	41.2	0.70	0.50	Yes
12	0.0	8.7	0.64	0.38	9.7	15.5	1.01	0.29	3.6	11.2	0.89	0.27	0.4	10.7	0.90	0.47	Yes
13	3.1	19.1	0.91	0.49	14.5	19.8	1.30	0.53	16.9	24.4	0.72	0.52	6.9	21.4	0.80	0.40	No
14	-1.1	14.4	0.54	0.74	-7.6	17.0	0.83	0.43	-7.6	18.0	0.77	0.53	-11.6	20.1	0.94	0.40	No
15	0.3	21.0	0.78	0.40	21.1	24.3	1.01	0.45	0.7	23.3	0.68	0.43	-0.9	20.3	0.66	0.49	Yes
16	-0.2	21.4	0.70	0.39	-26.1	26.1	1.68	0.52	-24.3	25.1	1.56	0.42	0.2	13.0	0.85	0.43	Yes
17	4.4	51.9	0.64	0.62	-4.5	36.1	0.66	0.59	16.6	41.1	0.73	0.59	-5.5	47.6	0.73	0.47	No
18	0.1	13.0	0.55	0.69	8.5	14.8	0.64	0.63	-5.4	18.8	0.95	0.33	-0.6	13.3	0.75	0.51	Yes
19	-9.2	14.9	0.69	0.56	-4.3	16.1	0.65	0.50	-5.7	19.7	0.90	0.14	0.7	17.9	0.74	0.47	No
20	0.0	13.9	0.56	0.64	-3.7	11.0	0.78	0.43	-12.4	16.9	0.98	0.38	-9.6	14.7	0.85	0.49	Yes
Mean	-0.2	18.4	0.65	0.55	-1.6	18.8	0.84	0.49	-0.9	21.2	0.83	0.42	0.0	19.8	0.76	0.47	Criteria
Stdev	2.7	10.8	0.10	0.13	10.4	7.8	0.28	0.10	11.4	8.6	0.20	0.11	6.0	9.9	0.11	0.11	Passing Rate
For cases meeting the criteria																	
Mean	-0.2	12.9	0.64	0.52	3.8	15.5	0.75	0.47	1.8	16.9	0.83	0.38	-0.7	13.5	0.71	0.55	55%
Stdev	1.5	4.0	0.07	0.13	8.4	5.7	0.16	0.12	7.5	4.2	0.13	0.08	4.5	3.3	0.11	0.08	

under the protocol that subjects will follow when wearing the device such as time stamping for meal size and using only their glucose meter to collect data. If these studies are successful, we plan to develop a prototype armband and evaluate it on several subjects. We envision this device collecting input and output data into the armband where the model will reside. After a

continuously over time on a watch type display or smart phone. Transmission of data from the armband to the display monitor may utilize Bluetooth technology.

We have overcome many challenges such as the use of a food index, the lack of initial conditions, frequent and long term removal of the armband and

**Table 5.** Modeling results of 2 week training, 1 week validation and 1 week testing data

Subject	$\hat{\eta}_{t^*}$												$\hat{y}_t$					Meeting Criteria
	Training				Validation				Testing				Testing					
	AE	AAE	RAAE	$r_{fit}$	AE	AAE	RAAE	$r_{fit}$	AE	AAE	RAAE	$r_{fit}$	AE	AAE	RAAE	$r_{fit}$		
1	1.7	16.2	0.72	0.47	14.6	30.3	0.79	0.47	11.9	19.5	0.75	0.42	17.5	24.7	0.82	0.33	No	
2	0.4	11.6	0.56	0.66	12.6	19.2	0.71	0.36	7.6	17.3	0.64	0.70	0.2	10.3	0.50	0.78	Yes	
3	-1.8	10.0	0.61	0.59	12.5	17.4	1.08	0.29	7.2	11.1	0.64	0.62	-0.2	10.1	0.74	0.46	Yes	
4	-0.5	8.5	0.70	0.45	5.8	11.4	0.90	0.18	7.7	10.2	0.90	0.50	6.2	9.9	0.68	0.50	Yes	
5	6.7	17.6	0.68	0.41	9.2	14.8	0.74	0.49	9.6	17.9	0.82	0.28	-1.5	14.1	0.74	0.48	Yes	
6	-3.7	16.1	0.80	0.41	-6.2	13.8	0.69	0.63	-5.5	14.3	0.74	0.53	-2.8	13.0	0.77	0.41	Yes	
7	-2.3	22.0	0.70	0.43	3.1	14.5	0.57	0.53	1.2	12.9	0.65	0.55	3.2	20.3	0.80	0.25	No	
8	2.5	18.4	0.67	0.45	-2.7	20.4	0.70	0.44	-0.2	22.7	0.71	0.42	-7.8	17.9	0.70	0.56	Yes	
9	0.7	14.0	0.62	0.61	2.5	18.5	0.88	0.36	3.4	21.2	0.69	0.59	8.5	22.3	0.69	0.49	Yes	
10	-5.1	19.7	0.55	0.66	-10.5	23.4	0.81	0.44	-11.5	26.7	0.67	0.58	3.0	29.7	0.78	0.41	No	
11	-2.6	35.7	0.65	0.51	-9.1	30.0	0.57	0.58	-28.7	50.5	0.86	0.49	6.5	38.9	0.61	0.62	Yes	
12	0.8	12.3	0.85	0.10	3.3	11.3	0.85	0.27	-3.1	8.9	0.74	0.21	1.3	8.9	0.88	0.46	Yes	
13	9.1	17.0	0.90	0.54	21.3	23.0	0.84	0.56	12.8	23.8	0.59	0.57	7.4	17.7	0.64	0.58	Yes	
14	-2.9	15.2	0.62	0.64	-7.6	15.9	0.69	0.64	-6.4	18.3	0.81	0.40	-11.9	20.7	1.04	0.33	No	
15	10.1	21.2	0.78	0.48	6.2	22.6	0.69	0.62	-4.7	23.0	0.65	0.51	-10.3	23.7	0.75	0.44	No	
16	9.9	18.1	0.73	0.27	-2.7	16.4	0.85	0.26	-4.4	7.7	1.12	0.26	3.9	9.0	1.09	0.43	Yes	
17	2.3	48.1	0.69	0.67	3.5	39.8	0.76	0.58	26.9	45.9	0.76	0.67	3.4	48.8	0.80	0.42	No	
18	1.3	12.2	0.52	0.69	-9.0	18.5	0.90	0.46	-9.3	19.5	1.03	0.25	-0.8	14.3	0.82	0.46	Yes	
19	-1.0	18.7	0.81	0.28	2.3	16.7	1.06	0.28	-0.8	25.9	0.98	0.15	-5.8	21.3	0.78	0.46	No	
20	6.5	13.2	0.65	0.59	-3.7	12.1	0.71	0.58	-4.4	16.6	0.93	0.21	-5.6	12.7	0.81	0.53	Yes	
Mean	1.6	18.3	0.69	0.50	2.3	19.5	0.79	0.45	0.5	20.7	0.78	0.45	0.7	19.4	0.77	0.47	Criteria	
Stdev	4.6	9.1	0.10	0.15	8.8	7.2	0.14	0.14	11.4	10.9	0.15	0.17	7.0	10.3	0.13	0.11	Passing	
For cases meeting the criteria																	Rate	
Mean	3.1	16.2	0.70	0.46	4.7	17.6	0.79	0.40	0.4	18.7	0.80	0.43	1.1	15.0	0.74	0.54	65%	
Stdev	4.6	7.7	0.11	0.17	9.4	5.9	0.14	0.15	12.0	12.5	0.16	0.18	5.1	9.1	0.16	0.10		

multiple inputs, subject-specific, modeling under infrequent sampling. However, as this work is only preliminary, there are still several challenges to overcome. This includes finding novel ways to improve the accuracy that leads to a higher percent of users meeting the MAC. In addition, the model procedure is quite complex as it requires advanced modeling experience and consists of several steps. One way we plan to improve accuracy is by gaining a better understanding on the bounds of each parameter. To address the model identification issue, we plan to development an estimation algorithm that identifies parameters automatically. This program will reside in the armband and will be used to calibrate the virtual sensor from on-line data. These are areas of future research that we have begun and the results are quite promising.

The authors thank BodyMedia for funding much of the data collection and allowing us to use their equipment and Jeanne Stewart for assisting in data collection.

### Appendix A

The purpose of this appendix is to provide a mathematical proof that  $r_{fit}$  under the simple linear regression model given by Eq. 12, does not depend on the model coefficients  $\gamma_o$  and  $\gamma_i$  but only one the explanatory variable,  $v_{i,t}$  in this case.

With  $\eta_t = \gamma_o + \gamma_i v_{i,t}$ , in this context,  $r_{fit}$  is mathematically given by

### Acknowledgements

$$\begin{aligned}
 r_{fit} = r_{y_i, \eta_i} &= \frac{\sum_{j=1}^n (y_j - \bar{y})(\eta_j - \bar{\eta})}{\sqrt{\sum_{j=1}^n (y_j - \bar{y})^2} \cdot \sqrt{\sum_{j=1}^n (\eta_j - \bar{\eta})^2}} \\
 &= \frac{\sum_{j=1}^n (y_j - \bar{y})(\gamma_0 + \gamma_i v_{i,t} - \gamma_0 - \gamma_i \bar{v}_i)}{\sqrt{\sum_{j=1}^n (y_j - \bar{y})^2} \cdot \sqrt{\sum_{j=1}^n (\gamma_0 + \gamma_i v_{i,t} - \gamma_0 - \gamma_i \bar{v}_i)^2}} \\
 &= \frac{\gamma_i \sum_{j=1}^n (y_j - \bar{y})(v_{i,t} - \bar{v}_i)}{\sqrt{\gamma_i^2} \sqrt{\sum_{j=1}^n (y_j - \bar{y})^2} \cdot \sqrt{\sum_{j=1}^n (v_{i,t} - \bar{v}_i)^2}} \\
 &= \frac{\gamma_i}{|\gamma_i|} r_{y_i, v_{i,t}}
 \end{aligned}$$

(A1)

Thus, with  $\gamma_i > 0$ ,  $r_{fit} = r_{y_i, v_{i,t}}$  and for  $\gamma_i < 0$ ,  $r_{fit} = -r_{y_i, v_{i,t}}$ . This result means that if the correlation of blood glucose concentration (BGC) and  $v_{i,t}$  is positive,  $\gamma_i$  can be set at any positive value and  $r_{fit}$  which will be  $> 0$ , will depend only of the behavior of  $v_{i,t}$  which is independently controlled by the values of the dynamic parameters associated with  $v_{i,t}$ . Conversely, if the correlation of BGC and  $v_{i,t}$  is negative,  $\gamma_i$  can be set at any negative value and  $r_{fit}$  will be  $> 0$  and independently controlled by the values of the dynamic parameters associated with  $v_{i,t}$ .

## References

- 1) Leinung, M., Thompson, S., and Nardacci, E., (2010) Benefits of continuous glucose monitor use in clinical practice, *Endocrine Practice*, vol. 16, issue 3, pp. 371-375.
- 2) UK Prospective Diabetes Study Group. (1997) Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33), *Lancet*, vol. 18, pp.716-738.
- 3) Battelino, T., Phillip, M., Bratina, N., Nimri, R., Oskarsson, P., Bolinder, J., (2011) Effect of continuous glucose monitoring on hypoglycemia in type 1 diabetes, *Diabetes Care*, vol. 34, issue 4, pp. 795–800
- 4) Davey, R.J., Jones, T.W., Fournier, P.A., (2010) Effect of short-term use of a continuous glucose monitoring system with a real-time glucose display and a low glucose alarm on incidence and duration of hypoglycemia in a home setting in type 1 diabetes mellitus, *Journal of Diabetes Science and Technology*. Vol. 4, issue 6, pp. 1457–1464.
- 5) Gandhi, G.Y., Kovalaske, M., Kudva, Y., Walsh, K., Elamin, M.B., Beers, M., Coyle, C., Goalen, M., Murad, M.S., Erwin, P.J., Corpus, J., Montori, V.M., Murad, M.H., (2011) Efficacy of continuous glucose monitoring in improving glycemic control and reducing hypoglycemia: a systematic review and meta-analysis of randomized trials, *Journal of Diabetes Science and Technology*. vol. 5, issue 4, pp. 952–965.
- 6) Anderson, J., Attvall, S., Sternemalm, L., Pivodic, A., Fahlen, M., Hanas, R., Ekeroth, G., Lind, M., (2011) Effect on glycemic control by short- and long-term use of continuous glucose monitoring in clinical practice, *Journal of Diabetes Science and Technology*. vol. 5, issue 6, pp. 1472–1479.
- 7) Yoo, H.J., An, H.G., Park, S.Y., et al, (2008) Use of a



- real time continuous glucose monitoring system as a motivational device for poorly controlled type 2 diabetes, *Diabetes Res Clin Pract*, issue 82, pp. 73–79.
- 8) Hermanides, J., Phillip, M., DeVries, J.H., (2011) Current Application of Continuous Glucose Monitoring in the Treatment of Diabetes, *Diabetes Care*, vol. 34, Suppl. 2, pp.197–201.
- 9) Chamberlain, J., Dopita D., Gilgen, E., (2013) Persistence of Continuous Glucose Monitoring Use in a Community Setting 1 Year After Purchase, *Clinical Diabetes*, vol. 31, no. 3, pp. 106-109.
- 10) Cagampang, F. R., Bruce, K. D. (2012) The role of the circadian clock system in nutrition and metabolism, *British Journal of Nutrition*. Vol. 108, No. 3, pp. 381-392.
- 11) Van Cauter, E., Polonsky, K.S., Scheen, A.J., (1997) Roles of circadian rhythmicity and sleep in human glucose regulation, *Endocr Rev*. vol. 18, no. 5, pp. 716.
- 12) Hansen, A.P. and Johansen, K. , (1970) Diurnal patterns of blood glucose, serum free fatty acids, insulin, glucagon and growth hormone in normals and juvenile diabetics, *Diabetologia*, vol. 6, no. 1, pp. 27-33.
- 13) Berlin, K.S., Rabideau, E.M., Hains, A.A., (2012) Empirically Derived Patterns of Perceived Stress Among Youth With Type 1 Diabetes and Relationships to Metabolic Control, *Journal of Pediatric Psychology*, vol. 37, no. 9, pp. 990-998.
- 14) Mommersteeg, P.M.C., Herr, R., Zijlstra, W.P., Schneider, S. and Pouwer, F. , (2012) Higher levels of psychological distress are associated with a higher risk of incident diabetes during 18 year follow-up: results from the British household panel survey, *BMC Public Health*, issue 12, pp. 1109.
- 15) Johannsen, D. , Calabro, M. A. , Stewart, J. , Franke, W. , Rood, J. C. , et al. , (2010) Accuracy of Armband Monitors for Measuring Daily Energy Expenditure in Healthy Adults, *Journal of Medicine & Science in Sports & Exercise*, vol. 42, issue 11, pp.2134-2140.
- 16) Rollins, D. K. , Bhandari, N. , Kleinedler, J. , Kotz, K. , Strohbehn, A. , et al. , (2010) Free-living inferential modeling of blood glucose level using only noninvasive inputs, *Journal of process control*, vol. 20, issue 1, pp. 95-107.
- 17) Fortuna, L. , Graziani, S. , Rizzo, A. , and Xibilia, M. G. , (2007) Soft sensors for monitoring and Control on Industrial Processes, *Soft sensors for monitoring and Control on Industrial Processes*. New York, Springer-Verlag.
- 18) Macias-Hernandez, J. J. , Angelov, P. , and Zhou, X. , (2007) Soft sensor for predicting crude oil distillation side streams using Tagaki Sugeno evolving fuzzy model, In *Proceedings of the 2007 IEEE International Conference on Systems, Man and Cybernetics*, pp. 3305-3310.
- 19) Eren-Oruklu, M. , Cinar, A. , Quinn, L. , and Smith, D. , (2009) Estimation of future glucose concentrations with subject-specific recursive linear models, *Diabetes Technology and Therapeutics*, vol. 11, issue 4, pp. 243-253.
- 20) Finan, D. A. , Palerm, C. C. , Doyle the Third, F. J. , Seborg, D. E. , Zisser, H. , et al. , (2009) Effect of input excitation on the quality of empirical dynamic models for type 1 diabetes, *AICHE Journal*, vol. 55, issue 5, pp.1135-1146.
- 21) Pappada, S. M. , Cameron, B. D. , and Rosman, P. M. , (2008) Development of a neural network for prediction of glucose concentration in type 1 diabetes patients, *Journal of diabetes Science and Technology*, vol. 2, issue 5, pp. 792-801.
- 22) Pedersen, S. N. , and Hansen, L. B. , (2010) System identification for prediction of glucose levels in people with type 1 diabetes, *System identification for prediction of glucose levels in people with type 1 diabetes*. Bachelor of Science thesis, Technical

- University of Denmark, Informatics and Mathematical Modeling.
- 23) Sparacino, G. , Zanderigo, F. , Corazza, S. , Maran, A. , Facchinetti, A. , and et al. (2007) Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series, *IEEE Transactions on Biomedical Engineering*, vol. 54, issue 5, pp. 931-937.
- 24) Stahl, F. , and Johansson, R. , (2009) Diabetes mellitus modeling and short-term prediction based on blood glucose measurements, *Mathematical Biosciences*, vol. 217, pp. 101-117.
- 25) Hastie, T. , Tibshirani, R. , and Friedman, J. , (2009) *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York, Springer, Second edition.
- 26) Beverlin, L. P. , Rollins, D. K. , Kotz, K. , Vyas, N. , and Andre, D. , (2011) An algorithm for optimally fitting a wiener model, *Mathematical problems in Engineering*, vol. 10, pp. 1155-1170.
- 27) Andre, D. , Teller, A. , (2005) Health Care Anywhere Today, *Stud Health Technol Inform*, 118, pp 89-110.
- 28) Welk, G. J. , Blair, S. N. , Woof, K. , Jones, S. , Thompson, R.W. , (2000) A comparative evaluation of three accelerometry-based physical activity monitors, *Med Sci Sports Exer*, 32(9 Suppl), pp 489-497.
- 29) Georga, E.I. , Protopappas, V.C. , Polyzos, D. , Fotiadis, D.I. (2010) Predictive modeling of glucose metabolism using free-living data of type 1 diabetic patients, *Engineering in Medicine and Biology Society*, Aug. Sept., pp.589-592.
- 30) Izenman, A.J. , (2008) *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*, Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning, Springer.
- 31) Cobelli, C. , Dalla Man, C., Sparacino, G. , Magni, L. , Nicolao, G.D. , (2009) *Diabetes: Models, Signals, and Control*, *IEEE Reviews in Biomedical Engineering*, Vol 2, pp 54-96.
- 32) Beverlin, L.P. , (2011) *The Exploitation of Advanced Statistical Methods and Analysis to Improve Nonlinear Dynamic Glucose Modeling*, The Exploitation of Advanced Statistical Methods and Analysis to Improve Nonlinear Dynamic Glucose Modeling, PhD Thesis , Iowa State University.