

Predicting Subcutaneous Glucose Concentration in Humans: Data-Driven Glucose Modeling

Adiwinata Gani, Andrei V. Gribok, Srinivasan Rajaraman, W. Kenneth Ward, and Jaques Reifman*

Abstract—The combination of predictive data-driven models with frequent glucose measurements may provide for an early warning of impending glucose excursions and proactive regulatory interventions for diabetes patients. However, from a modeling perspective, before the benefits of such a strategy can be attained, we must first be able to quantitatively characterize the behavior of the model coefficients as well as the model predictions as a function of prediction horizon. We need to determine if the model coefficients reflect viable physiologic dependencies of the individual glycemic measurements and whether the model is stable with respect to small changes in noise levels, leading to accurate near-future predictions with negligible time lag. We assessed the behavior of linear autoregressive data-driven models developed under three possible modeling scenarios, using continuous glucose measurements of nine subjects collected on a minute-by-minute basis for approximately 5 days. Simulation results indicated that stable and accurate models for near-future glycemic predictions (<60 min) with clinically acceptable time lags are attained only when the raw glucose measurements are smoothed and the model coefficients are regularized. This study provides a starting point for further needed investigations before real-time deployment can be considered.

Index Terms—Diabetes, glucose regulation, inverse problems, mathematical model, prediction, regularization, system identification.

I. INTRODUCTION

RECENT developments in continuous glucose monitoring (CGM) devices open new opportunities for glycemia management of diabetes patients. Modern CGM devices provide a minimally invasive mechanism to measure and record a patient's current glycemic state as frequently as every minute. Although

such devices measure glucose concentrations in the subcutaneous tissue, instead of in the blood, the frequent measurements are clinically useful, as they provide information about a patient's temporal glucose variability [1]. This benefit is in contrast to the current and commonly used fingerstick blood glucose meter in which only limited intermittent daily measurements are performed. CGM devices, however, only monitor current glucose concentrations, providing alerts when a blood glucose excursion is already underway (i.e., the glucose concentration may already be at an unacceptable high or low level) rather than regulating glucose interventions in a proactive manner.

The idea, originally suggested by Bremer and Gough [2], is to use CGM time-series signals and data-driven models to predict near-future glucose concentrations from past signals and integrate this predictive capability with CGM devices to anticipate glucose excursions and adjust therapy before concentrations drift from the desired range. Data-driven models represent a class of modeling techniques where the relationships between input and output process variables that characterize the underlying phenomenon being modeled are learned, during the training phase, from existing input-output data. Once the relationships have been learned, given new, unseen input process data, the models can accurately predict, up to a certain prediction horizon, the corresponding output as long as these data are within the envelope of relationships learned in the training phase.

A few data-driven techniques that employ glucose measurements to derive a model or to adjust a model's coefficients have been proposed to predict and control glucose concentrations in humans. For example, Trajanoski [3] proposes a radial-basis-function neural network model to predict and control subcutaneous glucose concentrations, and Dua *et al.* [4] employ a Kalman filter to adjust the parameters of a first-principles model for the prediction and control of blood glucose. The performance of these techniques, however, was only tested on simulated data and their performance on real data is unknown. Moreover, the training of neural networks requires the solution of a nonconvex optimization and the resulting network weights, or model coefficients, lack interpretability. Conversely, Kalman filters pose different implementation challenges. They require the availability of a high-fidelity first-principles model capable of accounting for meals and physical activity, which, given the complexity of the underlying physiology of glucose regulation coupled with the nonlinear dynamics of insulin action and glucose kinetics, has been elusive and remains an active area of research [5]–[7].

In contrast, data-driven autoregressive (AR) models are simpler and possess a number of advantages over the aforementioned techniques. AR models are linear, which allows for an

Manuscript received March 21, 2008; revised May 27, 2008 and June 30, 2008. First published September 16, 2008; current version published March 25, 2009. This work was supported by the Telemedicine and Advanced Technology Research Center of the U.S. Army Medical Research and Materiel Command, Fort Detrick, MD, under the U.S. Army Medical Department, Advanced Medical Technology Initiative. *Asterisk indicates corresponding author.*

A. Gani is with the Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702 USA (e-mail: agani@ucla.edu).

A. Gribok is with the Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702, and also with the Nuclear Engineering Department, University of Tennessee, Knoxville, TN 37996 USA (e-mail: agribok@bioanalysis.org).

S. Rajaraman is with the Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702 USA (e-mail: srini@bioanalysis.org).

W. K. Ward is with iSense Corporation, Wilsonville, OR 97070 USA, and also with Oregon Health and Sciences University, Portland, OR 97239 USA (e-mail: kward@isensecorp.com).

*J. Reifman is with the Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702 USA (e-mail: Jaques.reifman@us.army.mil).

Digital Object Identifier 10.1109/TBME.2008.2005937

explicit analysis of their stability and guarantees that the training process involves the solution of a convex optimization problem with a unique minimum. The simplicity of AR models is also reflected in the interpretability of the model coefficients, which, akin to the coefficients of an autocorrelation function (ACF), describe the dependencies present in the time-series data. Moreover, if trained with sufficient and representative data, AR models can learn and generalize across the daily variations of an individual's glucose concentration.

Despite their potential benefits, to date, only Cobelli's [8], [9] and our group [10] have proposed AR models for predicting individual-specific glucose concentrations. In [8], Sparacino *et al.* use a first-order AR model, AR(1), in which the model's coefficient is dynamically computed at each time step through a weighted least squares. They preprocess (smooth) the raw CGM data to remove high-frequency noise with a low-pass first-order Butterworth filter and find AR models with order larger than one to be unstable and AR(1) models with fixed coefficients to yield unacceptable prediction lags with delays equal to the prediction horizon. To assess the quality of the predictions, Sparacino *et al.* use the mean-squared prediction error (MSPE) and the time lag between the predicted and the smoothed signals as metrics. For a 30-min prediction horizon with weight (or forgetting factor) equal to 0.8, they report a median MSPE of $1.3 \text{ mmol}^2/\text{l}^2$ ($413 \text{ mg}^2/\text{dl}^2$) and an average lag of 3.79 min (SD = 18.01) on nadir-to-peak trends and an average lag of 5.29 min (SD = 12.63) on peak-to-nadir trends. They report lags at only peaks and nadirs because prediction delays are longest at these points. However, these lags are computed at 75% of peak-to-nadir and nadir-to-peak distances, in effect masking and reducing the true time lags. Moreover, they do not report if additional lags (between the smoothed and the raw signals) are introduced by the Butterworth filter, which could further increase the underestimation of the true time lags.

In our previous study [10], we use a tenth-order AR model, AR(10), with (fixed) coefficients determined by regularized least squares using the raw (unsmoothed) CGM time-series data. We employ the root-mean-squared error (RMSE) as a metric to assess the quality of the predictions, where the RMSE is defined as the square root of the mean-squared distance between the predicted and the raw signals. We report small prediction errors, with an average RMSE of 1.24 mmol/l (22.3 mg/dl) for a 30-min prediction horizon. However, we do not provide an assessment of the resulting prediction time lags, which we later observed to be approximately equal to the prediction horizon. These relatively large time lags reduce (or eliminate) the clinical benefits of the predictions.

In this paper, we propose a method to solve this problem and predict near-future glucose concentrations with acceptable time lags. We compared and contrasted three possible scenarios for near-future prediction of glucose concentrations with AR models: scenario I, which used raw glucose data to obtain unregularized AR coefficients; scenario II, which employed smoothed glucose data to compute unregularized AR coefficients; and scenario III, which used smoothed data to generate regularized AR coefficients. We showed that the first two scenarios generated models with unphysiologic coefficients, leading to models that

either possessed undesirable accuracy and large time lags or were unstable. In contrast, the method described by scenario III generated AR models with coefficients that reflected viable physiologic dependencies, leading to accurate near-future glucose concentration predictions with acceptable time lags. For a 30-min-ahead prediction, such models yielded negligible RMSEs (average of 0.1 mmol/l over nine subjects) and prediction time lags (average of 0.2 min), and clinically acceptable predictions for up to 60 min ahead with average RMSE of 0.7 mmol/l and average lag of 12.3 min.

II. METHODS

A. AR Modeling

In AR modeling, a predicted signal \hat{y}_n at time n ($n = m + 1, \dots, N$, where N denotes the total number of data samples available for modeling) is inferred as a linear combination of previously observed signals y_{n-i}

$$\hat{y}_n = \sum_{i=1}^m b_i y_{n-i} \quad (1)$$

where b denotes the vector of AR coefficients to be determined, and m denotes the order of the model (i.e., the number of previously observed glucose concentrations y_{n-i} used to predict a future glucose concentration \hat{y}_n). Accordingly, each AR coefficient b_i reflects the degree of dependency between the corresponding previous sample y_{n-i} and the predicted signal \hat{y}_n , providing a measure of the physiologic association of the time-series glucose data.

Training of an AR model consists of finding the coefficients b that best describe the dependencies in the entire time-series y . The coefficients are generally obtained by the ordinary least squares (OLS) method [11] in which b is estimated so that the functional $\|y - Ub\|^2$ is minimized, where U denotes the design matrix representing delayed values of y .

For glucose concentrations to be predictable with AR models, the CGM data must possess "detectable structure" and the dynamics of the time series data should, ideally, be stationary [2], [12]. By definition, a process is considered stationary when the sample mean and variance of the process measurements are constant with respect to time and the ACF is independent of absolute time [2], [12]. Indication of the stationary nature of the underlying process is therefore sought before applying AR models.

Fig. 1 shows three possible scenarios of how an AR model can be used to predict future glucose concentrations. In scenario I, AR models were obtained directly from the raw CGM time-series y . However, noise in the raw signal can make it difficult to compute AR coefficients that reflect viable physiologic dependencies. In scenario II, this problem was solved by smoothing the raw data to remove high-frequency noise before computing the AR coefficients. The additional correlations imposed on the data due to smoothing, however, can lead to severe ill-conditioning of the design matrix U . To overcome this problem, we proposed a third alternative, scenario III, where

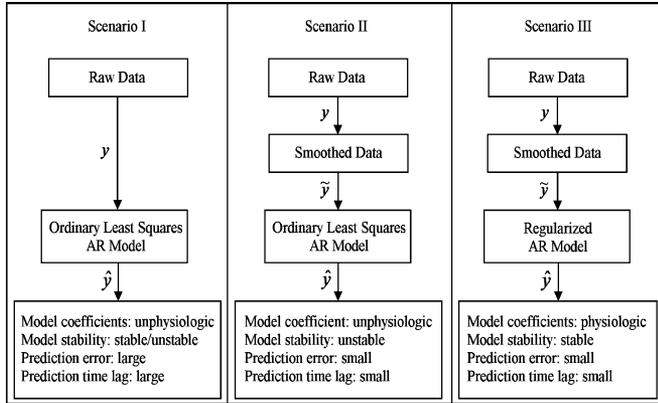


Fig. 1. Three different scenarios of how AR models can be used to predict glucose concentrations. Scenario I: OLS on the raw data; scenario II: OLS on the smoothed data; and scenario III: regularized least squares on the smoothed data.

we computed the AR coefficients through regularized least squares.

B. Smoothing the Raw CGM Data

There exists a number of smoothing techniques for filtering high-frequency noise [13]. In this paper, we used the Tikhonov regularization approach [14], which yields smoothed signals \tilde{y} by computing $\tilde{y} = U_d w$, where U_d denotes the integral operator and w denotes estimates of the glucose signals' first derivatives. This approach was chosen because the derivatives' estimates yield excellent data smoothing and do not introduce lag on the smoothed signal relative to the original raw signal. Through this approach, we chose the first derivative or the rate of change of glucose with time to impose smoothness constraints in the glucose signal [15]. In other words, we required the smoothed glucose signal \tilde{y} to vary minimally from one value to another, thereby ensuring regularity in the underlying signal that we wished to estimate.

To estimate the signal's derivatives w , we minimized the functional $f(w)$, given by

$$f(w) = \|y - U_d w\|^2 + \lambda_d^2 \|L_d w\|^2 \quad (2)$$

where y denotes the $N \times 1$ vector of the raw CGM time-series signal, U_d denotes the $N \times N$ integral operator, w represents the $N \times 1$ vector of first-order differences (the rate of change of glucose with time), λ_d represents the data regularization parameter, and L_d denotes a well-conditioned matrix chosen to impose smoothness constraints on the derivative of the glucose signal.

For a chosen L_d , the quality of smoothing in the aforesaid formulation is determined solely by the regularization parameter λ_d . When $\lambda_d = 0$, no regularization is performed, resulting in the original raw CGM data y . As λ_d increases, the solution w (and hence \tilde{y}) increasingly satisfies the imposed smoothness constraint, resulting, at the same time, in larger deviations from the raw data.

C. Model Regularization

To construct stable AR models, we obtained AR coefficients through regularization. For a stationary process, the sequence of autocorrelation coefficients representing the ACF describes statistical dependencies between two measurements separated by fixed time intervals throughout the recorded observations [2]. To force the AR coefficients to follow the same statistical dependencies of the ACF, a smoothness constraint was imposed on the OLS solution of the coefficients in (1), resulting in the regularized least squares functional $g(b)$, given by

$$g(b) = \|\tilde{y} - U_m b\|^2 + \lambda_m^2 \|L_m b\|^2 \quad (3)$$

where \tilde{y} denotes the $(N - m) \times 1$ vector of smoothed data, U_m denotes the $(N - m) \times m$ design matrix, b represents the $m \times 1$ vector of regularized AR coefficients, λ_m represents the model regularization parameter, and L_m denotes a well-conditioned matrix chosen to impose smoothness on the AR coefficients. Accordingly, the minimization of (3) resulted in regularized coefficients b . Although (2) and (3) have the same Tikhonov regularization form, they solve two different problems for two different purposes, i.e., data regularization/smoothing and model regularization, respectively.

Similar to the smoothing of the raw data, for a chosen L_m , the stability of the AR model in the aforesaid formulation is determined solely by the regularization parameter λ_m . When $\lambda_m = 0$, no regularization is performed, resulting in an OLS solution of (1). As λ_m increases, the coefficients are constrained, resulting in more stable, regularized AR coefficients.

D. Metrics

To assess the quality of predictions, we used two metrics: RMSE and time lag. When calculating the RMSE in scenario I, we computed the difference between the predicted signal and the raw signal. In scenarios II and III, the RMSE was computed as the difference between the predicted signal and the smoothed signal. The time lag was calculated as the cross correlation between the smoothed signal and its predictions. The lag corresponding to the peak of the cross-correlation function provides an accurate estimate of the prediction time lag [12], provided the length of the time series being compared is longer than one 24-h period of the human circadian rhythm [16].

E. Parameter Selection

There are a number of methods [17] for estimating the optimal values of the regularization parameters, λ_d and λ_m , and the order m of the AR model. In this paper, we found the optimum value of λ_d by minimizing the sum of the RMSE of the smoothed signal (i.e., the RMSE between the raw and the smoothed signal) and the RMSE of the prediction (i.e., the RMSE between the smoothed signal and its predictions). The RMSE of the smoothed signal is a monotonically increasing function of λ_d because the smoother the signal, the more it deviates from the original raw data. Conversely, the RMSE of the prediction is a monotonically decreasing function of λ_d because the smoother the signal, the more predictable it becomes.

Therefore, by obtaining λ_d that minimized the sum of these two RMSEs, we were effectively imposing a tradeoff between smoothness and predictability, resulting in signals with good predictability without oversmoothing them. We selected λ_m empirically and m through cross validation [11].

F. Subject Selection and CGM Measurements

Deidentified data for this investigation were obtained from a previous independent study of nine subjects with type 1 diabetes collected over a 5-day period. Subjects gave their voluntary and written informed consent to participate in the study, which had received approval by the appropriate Institutional Review Board. Subjects were included if they were $18 \leq \text{years} \leq 70$, had been diagnosed with type 1 diabetes and treated with insulin for at least 12 months, had body mass index $< 35.0 \text{ kg/m}^2$, and had glycated hemoglobin (HbA1c) $> 6.1\%$. Subjects were excluded if they had acute and severe illness apart from diabetes mellitus, clinically significant abnormal electrocardiogram, hematology or biochemistry screening tests, or any disease requiring use of anticoagulants. In addition, subjects were excluded if they were pregnant or lactating. The nine subjects ate standard weight maintenance meals (three per day) designed to avoid gaining or losing weight, and performed usual daily living activities. Also, each subject received a bolus of regular or ultrarapid insulin immediately before each meal either by subcutaneous injection or via the subcutaneous catheter of the insulin pump. Additional information about this independent study can be found in [10].

Subcutaneous glucose measurements were collected on a minute-by-minute basis for each of the nine subjects for approximately 5 days with the iSense CGM system [18]. We used only the first 4000 min (i.e., the first 4000 data points) of the available data for each subject, in which the first 2000 min were used for training the AR model (training dataset) and the next 2000 min were used for testing the prediction (testing dataset). For illustration purposes, for the different simulations described herein, we show prediction results for the same representative diabetic patient (subject #7). However, to demonstrate consistency of the proposed approach across different subjects, we also show one set of results for the other eight subjects.

III. RESULTS AND DISCUSSION

A. Stationarity of the Time-Series Data

To assess the stationarity of the glucose time series data, we examined the temporal behavior of three statistics of the signal. Each of the three metrics indicated that the glucose data were stationary. We found the mean and variance of both the raw CGM signal and the smoothed signal to be time invariant. We also found the estimates of the ACF of the raw signal to be independent of absolute time for time lags that far exceeded the near-future prediction horizon for which the AR models were used. Fig. 2 shows estimates of three ACFs for subject #7 based on three arbitrarily selected segments of the data, 0–4000, 1000–4000, and 2100–4000 min, each starting at a different time. The ACFs seem practically indistinguishable for time lags up to about 120 min, implying that the signal was stationary for our

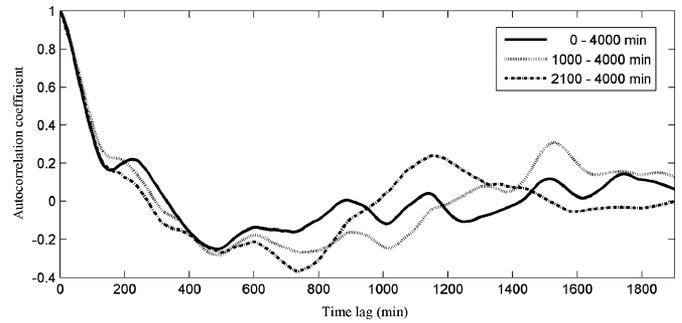


Fig. 2. Autocorrelation function of subcutaneous glucose measurements for iSense subject #7 calculated for three different starting times, 0, 1000, and 2100 min, corresponding to three overlapping segments, 0–4000, 1000–4000, and 2100–4000 min, respectively, of the time-series signal. The autocorrelation function is represented as a sequence of normalized autocorrelation coefficients.

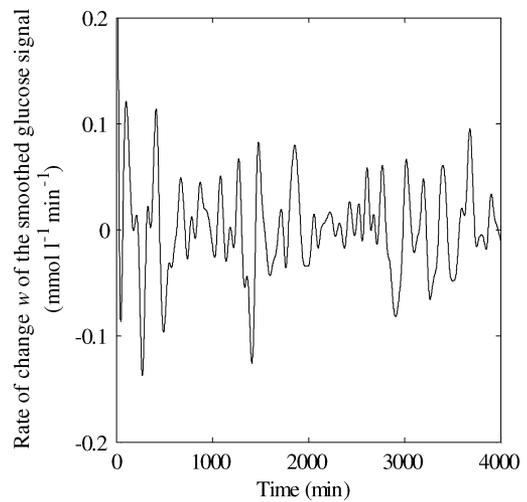


Fig. 3. Rate of change of the smoothed glucose data for subject #7.

modeling purposes. This also suggested that, for a fixed time lag, the dependency between any two series of glucose signals collected less than 120 min apart was the same, allowing for the construction of AR models with fixed coefficients and applicable to any time of the day. Fig. 2 also shows that the absolute values of the ACF became negligible and started to diverge for time lags larger than 200 min, indicating, as one might expect, that the glucose signal retained little, if any, interdependency between samples collected that far apart.

B. Smoothing of the Raw Data

We used Tikhonov regularization for smoothing the entire 4000 min of raw time-series data. We selected λ_d in (2) as the average of the optimum λ_d over the nine iSense subjects ($\lambda_d = 3000$), the integral operator U_d as described by Rajaraman *et al.* [19], and L_d to be a second-derivative operator [17]. Fig. 3 shows the estimated rate of change of glucose w in (2) as a function of time after imposing the smoothness constraint.

We noted that the smoothing yielded w values that lie within the expected limits of $-0.2 \text{ mmol l}^{-1} \text{ min}^{-1} \leq w \leq 0.2 \text{ mmol l}^{-1} \text{ min}^{-1}$ ($-4.0 \text{ mg dl}^{-1} \text{ min}^{-1} \leq w \leq$

4.0 mg dl⁻¹ min⁻¹) reported by Kovatchev *et al.* [20], whereas the raw CGM signal produced w in the range of $-1.7 \text{ mmol l}^{-1} \text{ min}^{-1} \leq w \leq 1.7 \text{ mmol l}^{-1} \text{ min}^{-1}$ (result not shown). For some situations, a scenario can be envisioned where glucose levels exhibit larger variations than the ones reported in [20]. In such cases, we expect the smoothing algorithm to oversmooth the glucose signal and the resulting predictions to underestimate the true glucose values. However, such cases are uncommon.

C. Selection of the AR Model Order m

We applied the cross-validation technique to select the order m of the AR model in (1). We could, therefore, select an optimal m for each of the nine subjects in each of the three scenarios. However, to ensure consistency in the comparisons of the simulation results across the different scenarios, we computed the optimal order of the AR model for scenario I and used this value throughout all simulations. Because the models in scenarios II and III were based on smoothed data, it is likely that such selection overestimated the order of the models for these two scenarios, potentially leading to overfitted models. Our results, however, indicated that the models generalized well, yielding small prediction errors, which suggested that the patterns in the testing data were similar to those in the training data. Furthermore, if λ_m in (3) is optimally selected, the imposed constraint prevents the regularized model in scenario III from overfitting.

The optimal model order for the nine subjects in scenario I ranged from 25 to 35. Because the differences in model performance for $25 \leq m \leq 35$ were negligible, we fixed the order of the AR models to a quasioptimal value of $m = 30$ for all simulations.

D. Scenario I: OLS AR Model Using Raw Data

We obtained the AR coefficients b by solving (1) with the OLS method using the training dataset consisting of raw signals. Fig. 4(a) shows that the coefficients of the 30th-order AR model, AR(30), obtained in this scenario are unphysiologic. Scenario I resulted in a random-walk model in which, incorrectly, the only significant coefficient was the first one and the glucose signal was predicted trivially by using only the current value as a predictor for future values.

We then investigated the predictive power of this AR(30) model on the testing dataset by determining the accuracy of the glucose concentration predictions—in this case, the RMSE between the predictions and the raw data—as a function of prediction horizon. The solid line in Fig. 5(a) shows the 30-min-ahead predictions (RMSE = 1.3 mmol/l and time lag = 30 min), where, for example, the first prediction at 2001 min was performed at 1971 min. We also checked the stability of the AR model by adding white Gaussian noise of zero mean and an arbitrarily small variance [$5.6 \times 10^{-13} \text{ mmol/l}$ (10^{-11} mg/dl)] to the raw signal and verified that the model was stable (not shown).

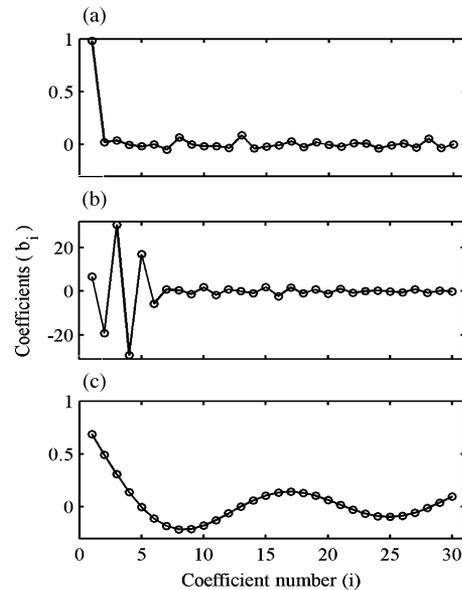


Fig. 4. Coefficients of a 30th-order AR model, AR(30), for iSense subject #7 obtained through: (a) scenario I, (b) scenario II, and (c) scenario III.

E. Scenario II: OLS AR Model Using Smoothed Data

After smoothing the glucose data, we retrained the AR(30) model using the same training dataset but now consisting of smoothed signals. Fig. 4(b) illustrates the coefficients of the AR(30) model, which show an oscillatory, unphysiologic behavior. When we added the same white noise of zero mean and $5.6 \times 10^{-13} \text{ mmol/l}$ variance to the smoothed data, we observed that the predictions became unstable. For 30-min-ahead predictions, illustrated by the light gray line in Fig. 5(b), the predictions became erratic and unbounded for small additions of white noise.

The instability in AR models of order higher than one is also reported by Sparacino *et al.* [8]. In contrast, the noise-free predictions [solid line in Fig. 5(b)] were very accurate (RMSE = 0.02 mmol/l). This demonstrated that models obtained through scenario II, albeit accurate in terms of predictions, were sensitive to small variations in the test data.

F. Scenario III: Regularized Least-Squares AR Model Using Smoothed Data

To obtain smooth-varying coefficients for the AR (30) model, we regularized the AR coefficients by empirically selecting $\lambda_m = 0.28$ in (3) and choosing the *a priori* constraint L_m to be a second-derivative operator similar to L_d in (2). Fig. 4(c) shows the coefficients of the resulting AR(30) model captured from the training dataset. These coefficients seem to reflect viable physiological dependencies in the glucose signal.

We used this regularized AR(30) model and investigated its predictive power on the smoothed testing dataset. The solid line in Fig. 5(c) shows that the 30-min-ahead predictions for the testing dataset of subject #7 were accurate (RMSE = 0.1 mmol/l) and had no time lag. We also checked the stability of the AR model by adding the same white noise

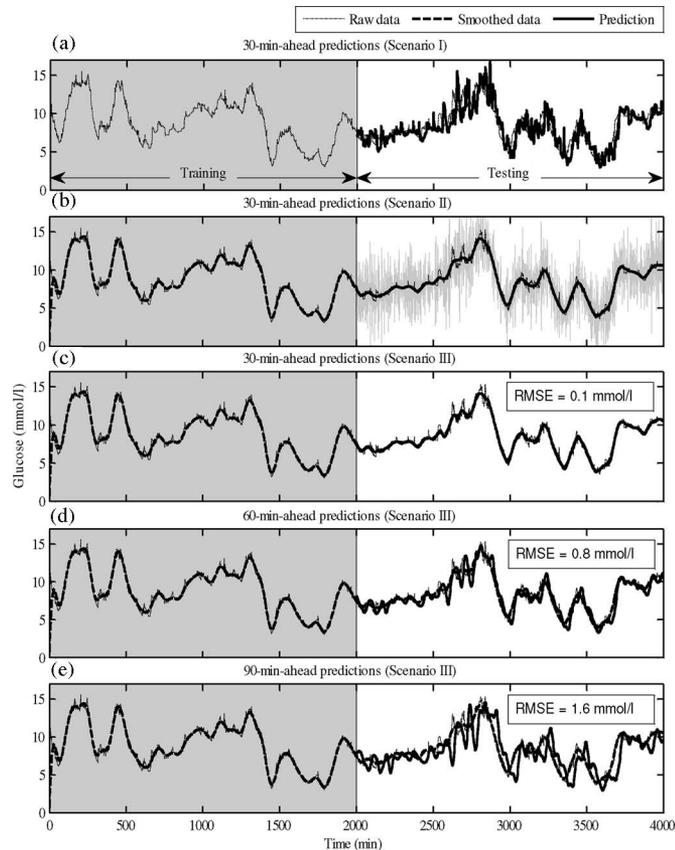


Fig. 5. Glucose concentration predictions using an AR(30) model for iSense subject #7 for (a) 30-min-ahead predictions obtained through scenario I (RMSE = 1.3 mmol/l and time lag = 30 min), (b) 30-min-ahead predictions obtained through scenario II without (solid line) and with (light gray line) added white noise of zero mean and 5.6×10^{-13} mmol/l variance to the smoothed data, (c) 30-min-ahead predictions obtained through scenario III (RMSE = 0.1 mmol/l and no time lag) with and without added noise, (d) 60-min-ahead predictions using scenario III (RMSE = 0.8 mmol/l and time lag = 12 min), and (e) 90-min prediction horizon using scenario III (RMSE = 1.6 mmol/l and time lag = 38 min).

of zero mean and 5.6×10^{-13} mmol/l variance to the smoothed data and verified that the predictions were stable [predictions with and without noise are superimposed in Fig. 5(c)].

We applied the same AR(30) model to make 60- and 90-min-ahead predictions. The solid line in Fig. 5(d) shows that the 60-min-ahead predictions were accurate (RMSE = 0.8 mmol/l) and had a clinically acceptable prediction time lag (12 min). The solid line in Fig. 5(e) shows that the 90-min-ahead predictions were less accurate (RMSE = 1.6 mmol/l) and had a larger prediction time lag (38 min). The time lag in the predictions reduced the effective prediction horizon by the corresponding number of minutes.

Fig. 6 illustrates 30-min-ahead predictions for the testing dataset of each of the remaining eight subjects using scenario III. The predictions for each subject demonstrated the same accuracy and consistency as the predictions for subject #7 shown in Fig. 5(c).

Table I tabulates the performance for 30-, 60-, and 90-min-ahead predictions for each of the nine iSense subjects. We employed the training dataset of each subject to develop an

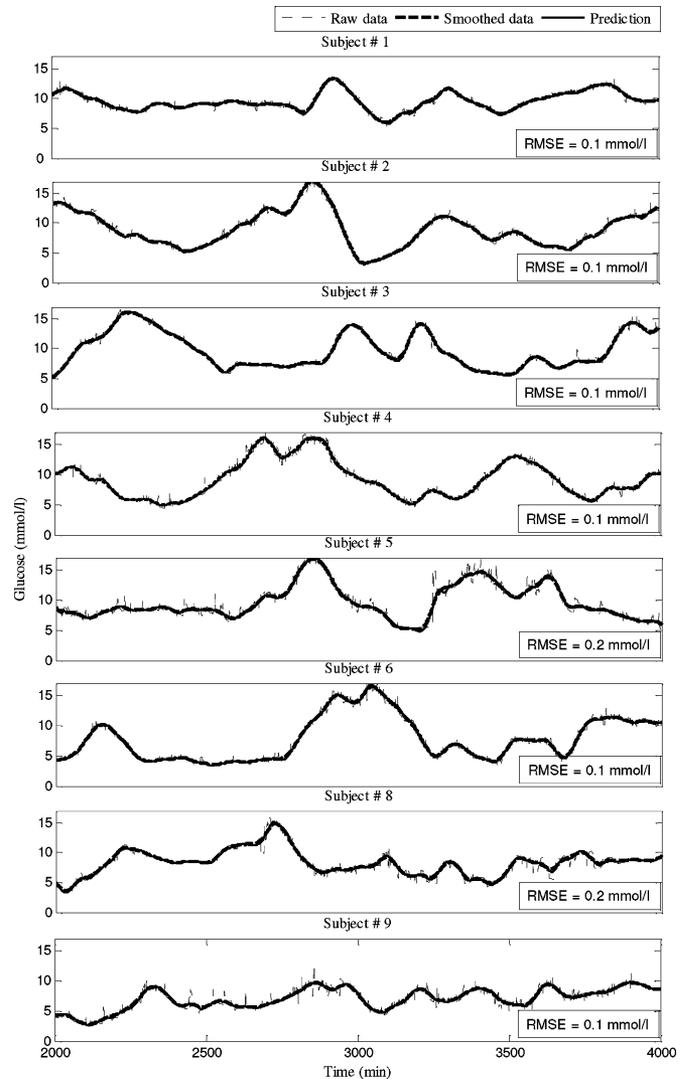


Fig. 6. Thirty-minute-ahead glucose concentration predictions for the remaining eight iSense subjects obtained with scenario III.

AR(30) model for that subject and tested the model performance on the corresponding testing dataset. The 30-min-ahead glucose concentration predictions were very accurate (average RMSE = 0.1 mmol/l), resulting in almost no prediction time lag (average of 0.2 min). The 60-min-ahead predictions had good accuracy (average RMSE = 0.7 mmol/l) and a clinically acceptable prediction time lag (average of 12.3 min).

IV. DISCUSSION

This paper attempts to address the question Bremer and Gough [2] posed nearly a decade ago: whether near-future glucose concentrations can be predicted from past glucose history. We showed that accurate 30-min-ahead predictions of subcutaneous glucose concentrations were achievable by smoothing raw CGM glucose data using Tikhonov regularization and deriving stable AR models of order 30 by applying the regularized least-squares technique on the smoothed data.

TABLE I
INDIVIDUALLY TUNED MODEL PERFORMANCE FOR NINE SUBJECTS

| Subject # | Prediction Horizon | | | | | |
|-----------------|--------------------|-------------------|------------------|-------------------|------------------|-------------------|
| | 30 min | | 60 min | | 90 min | |
| | RMSE (mmol/l) | Time Lag (min) | RMSE (mmol/l) | Time Lag (min) | RMSE (mmol/l) | Time Lag (min) |
| 1 | 0.1 | 1.0 | 0.6 | 15.0 | 1.2 | 42.0 |
| 2 | 0.1 | 0.0 | 0.7 | 7.0 | 1.6 | 28.0 |
| 3 | 0.1 | 0.0 | 0.8 | 9.0 | 1.8 | 34.0 |
| 4 | 0.1 | 1.0 | 0.7 | 14.0 | 1.6 | 39.0 |
| 5 | 0.2 | 0.0 | 0.9 | 14.0 | 2.0 | 44.0 |
| 6 | 0.1 | 0.0 | 0.8 | 10.0 | 1.7 | 33.0 |
| 7 | 0.1 | 0.0 | 0.8 | 12.0 | 1.6 | 38.0 |
| 8 | 0.2 | 0.0 | 0.9 | 14.0 | 1.7 | 40.0 |
| 9 | 0.1 | 0.0 | 0.6 | 14.0 | 1.3 | 42.0 |
| Average (SD) | 0.1 (0.02) | 0.2 (0.4) | 0.7 (0.1) | 12.3 (2.8) | 1.6 (0.2) | 38.4 (5.2) |

These 30-min-ahead predictions with no time lag provide ample time for proactive intervention and adjustment of therapy before glucose concentrations drift from desired ranges, hence, potentially avoiding possible hypo- and hyperglycemic episodes.

We discussed three possible scenarios (Fig. 1) of how an AR model can be used to predict future glucose concentrations. Deriving AR models directly from raw CGM data as in scenario I resulted in unphysiologic AR coefficients [Fig. 4(a)], where the first coefficient ($i = 1$) incorrectly reflected the entire dependency of the glucose signal. That is, future glucose signals were dependent only on the current glucose value. This behavior is caused by the high-frequency noise in the raw CGM data, which made it difficult to infer any but the first AR coefficient and yielded stable predictions with time lag equal to the prediction horizon (unobservable in Fig. 5(a) due to the scale of the figure). To remove the high-frequency noise, in scenario II, we applied Tikhonov regularization, which placed regularity constraints on the noisy signal. Retraining the AR models using the obtained smoothed data resulted in unphysiologic AR coefficients with alternating positive and negative correlations [Fig. 4(b)] and highly unstable predictions [Fig. 5(b)]. Fig. 4(b) also indicates that the third and fifth coefficients were larger than the first coefficient, incorrectly implying that future glucose values were more dependent on previous, further away values than on closer ones. If some high-frequency noise were retained, resulting in an undersmoothed signal, unstable AR models similar to the one observed in scenario I might be obtained. We speculate that this is the case in the study presented by Sparacino *et al.* [8] and Zanderigo *et al.* [9], where the Butterworth filter is used to preprocess the raw signals. If the raw signals were sufficiently smoothed, implying scenario II, the resulting AR model, albeit

unstable and containing unphysiologic coefficients, could yield very accurate predictions with almost no prediction time lag.

To achieve stable models and physiologically plausible AR coefficients, in scenario III, we regularized the AR coefficients. Deriving AR models through regularized least squares resulted in AR coefficients that reflected viable physiological dependencies in the glucose signal [Fig. 4(c)] and stable, accurate predictions with the prediction accuracy decreasing with increasing prediction horizon [Fig. 5(c)–(e)].

Several comments regarding the physiologic interpretation of the AR model coefficients are in order. Equation (1) shows that the predicted values of a glucose signal are obtained as a linear combination of antecedent glucose samples y multiplied by the AR coefficients b . Accordingly, each AR coefficient b_i reflects the degree of dependency between the corresponding previous sample y_{n-i} and the predicted value \hat{y}_n . Because this degree of dependency between samples separated by i observations apart is also reflected in the coefficients of the ACF used to compute b_i , the physiologic interpretation of the AR coefficients b_i can be gleaned from the analysis of the ACF. Fig. 2 shows the temporal dependency of the ACF coefficients as a function of time delay, or time lag, between glucose samples for subject #7. As the figure indicates, for up to about 500 min, the value of the ACF coefficients gradually decreases as a function of time lag, reflecting a graduated decrease in the dependency between samples as a function of time lag between them. Another important observation is that the values of the ACF coefficients are close to the maximum value of one for time lags fewer than 30 min, reflecting that there is a strong dependency between glucose samples that are 30 or fewer minutes apart.

Analysis of the AR coefficients in Fig. 4(a) revealed that scenario I produced coefficients that drop to zero too fast, effectively implying that only the most immediate predecessor sample affects future glucose values. This observation is not supported by the ACF of the signal, which shows that a strong dependency between samples is retained well after 1 min. This supports the conclusion that coefficients obtained under scenario I are unphysiologic. Conversely, scenario II produced coefficients that imply that samples 2–5 min apart from the predicted value influence the predicted value more than an immediate predecessor sample [Fig. 4(b)]. This observation is contradicted by Fig. 2, which suggests a monotonically decreasing dependency of the coefficients as a function of time lag up to 120 min. Furthermore, from a physiologic point of view, it would be difficult to justify why samples collected farther apart should have stronger interdependencies than closer ones. Finally, scenario III produced coefficients [Fig. 4(c)] that, in general, reflect the temporal behavior of the ACF, implying that AR models with such coefficients correctly describe the underlying dependencies in the glucose time-series signal.

In smoothing the raw signal, we found that a range of parameter values for λ_d , spanning from 2000 to 5000, did not significantly affect the smoothness of the data. However, when λ_d was too small, the raw data became undersmoothed, resulting in models with behavior similar to those in scenario I. When λ_d was too large, the raw data became oversmoothed and resulted in signals that had lost some, if not all, of the major peaks and

nadir of the original time-series data. Similarly, in regularizing the AR model, we found that a range of parameter values for λ_m , spanning from 0.06 to 2.8, did not significantly affect the accuracy of the predictions. Similar observations regarding these regularization parameters are also noted by Freeland and Bonnacaze [21].

Due to the robust nature of our algorithm with respect to both λ_d and λ_m , we hypothesize that the approach proposed here following scenario III may yield “universal,” or “portable,” models for glucose concentration predictions. Such universal models, developed based on data from just one individual and applicable to all other individuals, would considerably reduce the burden associated with model tuning and data collection for model development. Our preliminary results support the feasibility of such universal model hypothesis, as the AR model coefficients seem to be indistinguishable among the nine individuals studied here. This hypothesis, however, assumes that we have accurate measurements of glucose concentration (in mmol/l, for example), with the appropriate calibration and mapping from the raw CGM subcutaneous measurements (in nanoamperes, for example) for each individual.

The real-time implementation of the proposed method is an important issue and needs to be investigated. In this paper, we smoothed the entire glucose time series signals (4000 min) at once and then used the smoothed signal to obtain the AR coefficients and make predictions. Such a “global” smoothing did not duplicate real-time conditions, where future glucose values are unknown and cannot be used to generate improved smoothing. Hence, for real-time predictions, the smoothing algorithm needs to be modified to use data up to the current time and not the entire dataset. The choice of an asymmetric (purely causal) or a symmetric (noncausal) smoothing algorithm also needs to be investigated. If an asymmetric algorithm is used, it is likely to yield *end-effect* problems because smoothing is calculated only up to the current time [12]. A symmetric smoothing algorithm, however, could be used to offset the end-effect problem by “borrowing” some future data for smoothing purposes. The net effect of this borrowing is a reduction of the true prediction horizon, which needs to be properly quantified to determine whether it still provides ample time for proactive interventions.

The continuous measurement of glucose concentration via CGM devices together with data-driven AR models provides a potential, practically useful combination of technologies for accurate near-future prediction of glucose concentrations. With improvements in CGM sensor technology to extend duration of use and future research to address model universality and real-time implementation, near-future predictions can give diabetes patients adequate time to intervene and adjust therapy before glucose concentrations drift from the desired range and avoid hypo- and hyperglycemia episodes.

ACKNOWLEDGMENT

The authors wish to thank Colonel R. A. Vigersky for providing expertise on diabetes mellitus, and to Dr. L. Heinemann and his staff at the Profil Clinic, Neuss, Germany, for collection of clinical study data. The opinions and assertions contained herein are the private views of the authors and are not to be construed

as official or as reflecting the views of the U.S. Army or of the U.S. Department of Defense. This paper has been approved for public release with unlimited distribution.

REFERENCES

- [1] D. C. Klonoff, “Continuous glucose monitoring-roadmap for 21st century diabetes therapy,” *Diabetes Care*, vol. 28, pp. 1231–1239, 2005.
- [2] T. Bremer and D. A. Gough, “Is blood glucose predictable from previous values? A solicitation for data,” *Diabetes*, vol. 48, pp. 445–451, 1999.
- [3] Z. Trajanoski, “Simulation studies on neural predictive control of glucose using the subcutaneous route,” *Comput. Methods Programs Biomed.*, vol. 56, pp. 133–139, 1998.
- [4] P. Dua, F. J. Doyle, III, and E. N. Pistikopoulos, “Model-based blood glucose control for type 1 diabetes via parametric programming,” *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, pp. 1478–1491, Aug. 2006.
- [5] A. Boutayeb and A. Chetouani, “A critical review of mathematical models and data used in diabetology,” *Biomed. Eng. Online*, vol. 5, p. 43, 2006.
- [6] A. Makroglou, J. X. Li, and Y. Kuang, “Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: An overview,” *Appl. Numer. Math.*, vol. 56, pp. 559–573, 2006.
- [7] G. Nucci and C. Cobelli, “Models of subcutaneous insulin kinetics. A critical review,” *Comput. Methods Programs Biomed.*, vol. 62, pp. 249–257, 2000.
- [8] G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, A. Facchinetti, and C. Cobelli, “Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series,” *IEEE Trans. Biomed. Eng.*, vol. 54, no. 5, pp. 931–937, May 2007.
- [9] F. Zanderigo, G. Sparacino, B. Kovatchev, and C. Cobelli, “Glucose prediction algorithms from continuous monitoring data: Assessment of accuracy via continuous glucose error-grid analysis,” *J. Diabetes Sci. Technol.*, vol. 1, pp. 645–651, 2007.
- [10] J. Reifman, S. Rajaraman, A. Gribok, and W. K. Ward, “Predictive monitoring for improved management of glucose levels,” *J. Diabetes Sci. Technol.*, vol. 1, pp. 478–486, 2007.
- [11] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ: Prentice-Hall, 1999.
- [12] C. Chatfield, *The Analysis of Time Series: An Introduction*. Boca Raton, FL: Chapman & Hall/CRC, 2004.
- [13] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing*. Upper Saddle River, NJ: Prentice-Hall, 1999.
- [14] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*. Washington, DC: Winston, 1977.
- [15] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*. New York: Springer, 2005.
- [16] G. V. Seaman, R. Engel, R. L. Swank, and W. Hissen, “Circadian periodicity in some physicochemical parameters of circulating blood,” *Nature*, vol. 207, pp. 833–835, 1965.
- [17] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Philadelphia, PA: SIAM, 1998.
- [18] iSense Corporation Portland (2008, Feb. 29). Portland, OR [Online]. Available: <http://www.isensecorp.com>
- [19] S. Rajaraman, A. V. Gribok, N. J. Wessensten, T. J. Balkin, and J. Reifman, “Individualized performance prediction of sleep-deprived individuals with the two-process model,” *J. Appl. Physiol.*, vol. 104, pp. 459–468, 2008.
- [20] B. P. Kovatchev, W. L. Clarke, M. Breton, K. Brayman, and A. McCall, “Quantifying temporal glucose variability in diabetes via continuous glucose monitoring: Mathematical methods and clinical application,” *Diabetes Technol. Ther.*, vol. 7, pp. 849–862, 2005.
- [21] A. C. Freeland and R. T. Bonnacaze, “Inference of blood glucose concentrations from subcutaneous glucose concentrations: Applications to glucose biosensors,” *Ann. Biomed. Eng.*, vol. 27, pp. 525–537, 1999.



Adiwinata Gani received the B.S., M.S. and Ph.D. degrees in chemical engineering from the University of California, Los Angeles, in 2002, 2003, and 2007, respectively.

He is currently a Postdoctoral Fellow in the Department of Bioinformatics Cell, Telemedicine and Advanced Technology Research Center, Fort Detrick, MD. His current research interests include inverse and ill-posed problems in engineering, fault-detection, and fault-tolerant control systems.



Andrei V. Gribok received the B.S. degree in systems science and the M.S. degree in nuclear engineering from Moscow Institute of Physics and Engineering, Moscow, Russia, both in 1987, and the Ph.D. degree in biological physics from Moscow Institute of Biological Physics, Moscow, in 1996.

He is currently a Research Assistant Professor in the Department of Nuclear Engineering at the University of Tennessee, Knoxville. He is currently on an interpersonnel agreement assignment with Telemedicine and Advanced Technology Research Center, Fort Detrick, MD.

He is the author or coauthor of three book chapters and numerous journals and conference papers. His current research interests include inverse and ill-posed problems in engineering, statistical learning, and model misspecification in statistics.



Srinivasan Rajaraman received the B.S. degree in chemical engineering from Indian Institute of Technology, Roorkee, India, in 2000, and the Ph.D. degree in chemical engineering from Texas A&M University, College Station, in 2005.

He is currently a Research Scientist in the Bioinformatics Cell, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD. His current research interests include the areas of biomathematical modeling, estimation and prediction of dynamic systems, time-series analysis, process monitoring and

fault diagnosis, and inverse problems.



W. Kenneth Ward received the Medical degree from the University of Colorado, Boulder, in 1978. Thereafter, he was awarded a Research Fellowship in diabetes and endocrinology at the University of Washington.

He is currently a Diabetes Specialist and Biosensor Scientist for iSense Corporation, Wilsonville, OR, where he is engaged in collecting data used in the study while collaborating with the staff at the Profil Institute, Neuss, Germany. His current research interests include development of minimally invasive and

fully implantable biosensors, artificial pancreas algorithms, biomaterials, and the foreign body response to implanted devices.

Dr Ward has authored/coauthored several books and numerous articles in the area of biosensors.



Jaques Reifman received the B.S. degree in civil engineering from Rio de Janeiro State University, Rio de Janeiro, Brazil, in 1980, the B.B.A. degree in business administration from Rio de Janeiro Federal University, Rio de Janeiro, in 1985, and the M.S.E. and Ph.D. degrees in nuclear engineering from the University of Michigan, Ann Arbor, in 1985 and 1989, respectively.

He is currently a Senior Research Scientist in the Department of the Army, U.S. Army Medical Research and Materiel Command (USAMRMC), Fort

Detrick, MD, where he is also the Director of the Department of Defense Biotechnology High Performance Computing Software Applications Institute for Force Health Protection and the USAMRMC Bioinformatics Cell at the Telemedicine and Advanced Technology Research Center. His current research interests include the areas of physiological signal processing, statistical pattern recognition, artificial intelligence, data mining, biomathematical modeling, bioinformatics, genomics, and proteomics.