GLUCOSE CONTROL IN TYPE I DIABETIC PATIENTS: A VOLTERRA MODEL–BASED APPROACH

Justin D. Rubb and Robert S. Parker

Department of Chemical and Petroleum Engineering
University of Pittsburgh
Pittsburgh, PA 15261 USA
rparker@pitt.edu

Abstract: Glucose concentration controllers for Type I diabetic patients are synthesized using model–based methods. A physiologically–based model of the insulin–dependent diabetic is employed as the patient. For modeling and control purposes, the patient is approximated using nonlinear Volterra series models. These data–driven models are then employed in two nonlinear controller synthesis strategies: internal model control using partitioned inverses (Doyle III *et al.*, 1995), and model predictive control.

Keywords: Biomedical systems, internal model control, model predictive control, nonlinear models, sampled-data systems, Volterra series

1. INTRODUCTION

In the normal patient, proper glucose control is maintained by the pancreatic β–cells; these cells alter their secretion of insulin, a potentiator for glucose removal from the bloodstream, in response to changing glucose levels. In the Type I (or insulin–dependent) diabetic patient, this control mechanism does not function properly, leading to sustained elevated blood glucose concentration and a condition known as hyperglycemia (defined as blood glucose $>$ 120 mg/dL). The Diabetes Complications and Control Trial (1993; 1996) has shown that this condition is responsible for many of the long–term effects of diabetes, such as blindness, kidney failure, and limb loss. While classical injection therapy can return the patient to near normoglycemic levels (glucose between 70 and 120 mg/dL), two primary drawbacks result from this treatment. First, the non–continuous nature of treatment often leads to wide variations in glucose concentration. In addition, over–delivery of insulin can result in significant drops in blood glucose concentration into the hypoglycemic range ($<$ 60 mg/dL). This low–glucose condition deprives the cells of fuel and can lead to coma and patient death. The maintenance of glucose within tight physiological limits is of supreme importance for the survival of diabetic patients.

The treatment of insulin–dependent diabetes currently employs insulin injection, inhalation delivery systems, or continuous infusion pumps. The inherent drawback of all of these approaches is their reliance on patient compliance to achieve long–term glucose control. The patient is normally required to adjust their insulin dose levels prior to meals, exercise, and sleep, and it is assumed that the patient delivers a correctly estimated dose at the proper time. The loss of glucose control may result from an incorrect estimate of insulin need or a missed dose. In an effort to remove the patient from the control loop, this work focuses on the development of a closed–loop insulin delivery system using periodic glucose measurements to calculate and deliver an insulin dose that will maintain the patient within the normoglycemic range in response to a variety of physiological disturbances.

Three primary components would compose a closed–loop insulin delivery system. Patient glucose measurements would be accomplished with an *in vivo* glucose concentration sensor; a significant effort is ongoing in this area (Jaremko and Rorstad, 1998). To

deliver variable amounts of insulin to the patient, there exist a variety of pump mechanisms (Cohen, 1993; Minimed Corporation, 1999). Linking the sensor and the delivery device is the control algorithm, on which this paper is focused. Classical feedback algorithms are inadequate for glucose control, due to the existence of system constraints and their interactions with the patient dynamics. The following work evaluates two candidate advanced control structures, nonlinear internal model control using partitioned model inverses (PNLI–IMC) and nonlinear model predictive control (NMPC), as well as their linear counterparts, with respect to their glucose control performance in response to glucose concentration challenges.

2. DIABETIC PATIENT CASE STUDY

The structure used as the diabetic patient in this work is the physiologically–based pharmacokinetic / pharmacodynamic model given in (Sorensen, 1985; Parker *et al.*, 2000). This model uses a compartmental technique to account for the connectivity and interaction of the various organs important to glucose and insulin metabolism and dynamics. From an input–output perspective, the steady state behavior is shown in Figure 1. In the local region of the nominal

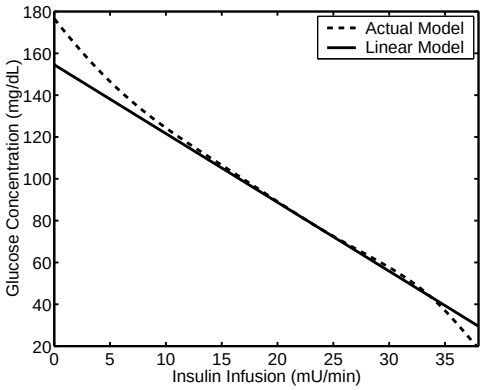

Fig. 1. Steady state locus for the diabetic patient. Nonlinear patient model (dashed), linear model (solid).

condition ($u_{ss} \approx 23.9$ mU/min, $y_{ss} \approx 81.1$ mg/dL), the diabetic patient displays linear behavior. However, more severe hypo– and hyperglycemic states elucidate the nonlinear character of the diabetic patient glucose–insulin response. The shape of the steady state locus motivates the use of polynomial empirical models to approximate the system behavior. Advantages of this approach include the relative ease with which empirical models can be identified and updated, as compared to more physiologically relevant models. Furthermore the calculation of the model inverse, for use in the control algorithm, can be facilitated by selecting certain model structures, as discussed in greater detail below (Doyle III *et al.*, 1995).

3. MODEL DEVELOPMENT

Model–based control systems require an accurate dynamic patient model. Given the significant variability observed in biomedical systems (Puckett and Lightfoot, 1995), an easily customizable model is preferable. These requirements lead to the use of empirical model structures for capturing patient dynamics. If a linear model is required, which facilitates controller synthesis, then discrete–time transfer functions can be employed. Alternatively, Volterra series, a member of the class of nonlinear moving average models, are effective in approximating nonlinear process dynamics (Boyd and Chua, 1985; Zheng and Zafiriou, 1994; Zheng and Zafiriou, 1995). Both structures are straightforward to update, and the decision about which model to employ depends on the control structure and desired performance.

3.1 *Volterra Series*

In an effort to capture the nonlinear characteristics of the diabetic patient, a nonlinear Volterra series model was selected to approximate the input–output behavior of the diabetic patient. Previous work (Florian and Parker, 2002) developed a Volterra model for the patient process described in Section 2. The remainder of this subsection will highlight those results as they form the basis for the model employed in the control studies.

The general Volterra series model has the form:

$$\hat{y}(k) = y_0 + \sum_{i=1}^{N} \sum_{j_1=1}^{M} \cdots \sum_{j_N=1}^{M} h_i(j_1, \ldots, j_N) \times u(k-j_1) \ldots u(k-j_N) \quad (1)$$

The diabetic patient can be approximated using the above discrete–time nonlinear model because the glucose–insulin dynamics display fading memory (Boyd and Chua, 1985); inputs further in the past have a lesser effect on the output than more recent input changes, up to a memory of M, beyond which the input effects are no longer significant. By selecting a model memory, M and model order, N, a truncated Volterra series can be employed to model a given system. Model coefficients ($h_i(j_1, \ldots, j_N)$) identified from patient data provide an empirical relationship (the Volterra model) between past insulin infusion rates ($u(k-i)$) and glucose concentration ($y(k)$) for a given patient at each sample time (k).

Starting from the general Volterra series model in equation (1), Florian and Parker (2002) showed that a third–order diagonal structure provides a good trade–off between identifiability (from limited clinical data) and predictive accuracy. The diagonal structure reduced the number of unknown model coefficients from 12,341 to 121, and decreased the

data requirements by orders of magnitude. This Volterra model can be decomposed as:

$$y(k) = h_0 + \mathscr{L}(k) + \mathscr{D}_2(k) + \mathscr{D}_3(k) \tag{2}$$
$$\mathscr{L}(k) = \sum_{i=1}^{M} h_1(i)u(k-i),$$
$$\mathscr{D}_2(k) = \sum_{i=1}^{M} h_2(i,i)u^2(k-i),$$
$$\mathscr{D}_3(k) = \sum_{i=1}^{M} h_3(i,i,i)u^3(k-i),$$

Here $\mathscr{L}$ denotes the linear terms, and diagonal terms of order N are given by $\mathscr{D}_N$. To facilitate the identification procedure, $y(k)$ and $u(k)$ are in scaled deviation form. In order to capture the dynamic response of the diabetic patient, the glucose sampling rate, T_s, and the model memory, M, were selected such that $\approx 99\%$ of the patient step response was captured by M coefficients, resulting in $T_s = 10$ min and $M = 40$. This is a reasonable memory length for application in controller synthesis, and the glucose sampling interval of 10 minutes is characteristic of current sensor development goals.

In model identification, tailored input sequences can dramatically reduce the amount of data necessary for model development while simultaneously offering improved coefficients (Parker *et al.*, 2001*a*). An input sequence of at least four discrete levels is required to identify a third–order Volterra model (Nowak and Van Veen, 1994). As in (Florian and Parker, 2002), a tailored five–level sequence was constructed that excited only the diagonal terms. The input sequence:

$$u(k) = \begin{cases} \gamma_1 & k = 0 \\ 0 & 1 \le k \le M \\ -\gamma_1 & k = M+1 \\ 0 & M+2 \le k \le 2M+1 \\ \gamma_2 & k = 2M+2 \\ 0 & 2M+3 \le k \le 3M+2 \\ -\gamma_2 & k = 3M+3 \\ 0 & 3M+4 \le k \le 4M+3 \end{cases} \tag{3}$$

is a special case of a continuous–switching–pace symmetric random sequence, where the sequence levels represent deviations from the nominal input value. In the previous identification study (Florian and Parker, 2002), it was established that $|\gamma_1| < |\gamma_2|$ provided superior estimates of model coefficients due to the lesser effect of the finite memory assumption for smaller inputs. Furthermore, the second–order diagonal coefficients were identified only from the γ_2 magnitude pulse responses, as these provided better stimulation of nonlinear behaviors. The sequence used to identify the Volterra model in equation (2) is shown in Figure 2

The identification objective was chosen to be minimization of model prediction error:

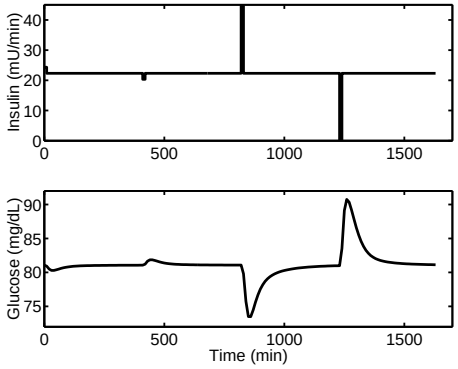

Fig. 2. Input sequence (top) and output response (bottom) for linear plus nonlinear diagonal coefficient identification.

$$J = \sum_{k=0}^{4M+3} e^2(k) = \sum_{k=0}^{4M+3} [y(k) - \hat{y}(k)] \tag{4}$$

In combination with the input sequence above, coefficient estimators can be analytically derived as in (Florian and Parker, 2002):

$$\hat{h}_0 = \frac{y(0)+y(M+1)}{4} + \frac{y(2M+2)+y(3M+3)}{4} \tag{5}$$

$$\hat{h}_1(k) = \frac{\gamma_2^3(y(k)-y(k+M+1))}{2\gamma_1\gamma_2(\gamma_2^2-\gamma_1^2)} - \frac{\gamma_1^3(y(k+2M+2)-y(k+3M+3))}{2\gamma_1\gamma_2(\gamma_2^2-\gamma_1^2)} \tag{6}$$

$$\hat{h}_2(k,k) = \frac{(y(k+2M+2)-y(2M+2))}{2\gamma_2^2} + \frac{(y(k+3M+3)-y(3M+3))}{2\gamma_2^2} \tag{7}$$

$$\hat{h}_3(k,k,k) = \frac{\gamma_2(y(k)-y(k+M+1))}{2\gamma_1\gamma_2(\gamma_2^2-\gamma_1^2)} - \frac{\gamma_1(y(k+2M+2)-y(k+3M+3))}{2\gamma_1\gamma_2(\gamma_2^2-\gamma_1^2)} \tag{8}$$

These estimators were updated from (Parker *et al.*, 2001*a*) to include third–order diagonal model estimation effects and the use of a partial sequence for second–order coefficient estimation. These estimators show superior performance in the absence of measurement noise; by repeating the input sequence in Figure 2, noise effects can be averaged over the number of repeats, as in (Parker *et al.*, 2001*a*).

4. CONTROLLER SYNTHESIS

While a significant amount of work has been performed in the area of controller synthesis for insulin–dependent diabetic patients (see (Parker *et al.*, 2001*b*) for a survey) the majority of these controllers employ linear patient models. The key

Table 1. Controller and model structure evaluation chart. The cells include the abbreviation used for the particular controller–model pairs.

		Controller	
		IMC	MPC
Model	Linear	LIMC	LMPC
	Nonlinear	PNLI-IMC	NMPC

contribution in the present study is the evaluation and comparison of two control structures, internal model control (IMC) using a partitioned model inverse and model predictive control (MPC), when the models employed are clinically–relevant empirical models of the Volterra type. Table 1 provides the abbreviations used for the controllers under evaluation. In all cases, the controllers were designed to accommodate a sampling rate of 1 measurement per 10 min. Based on current sensor research (Jaremko and Rorstad, 1998), this sampling rate provides a reasonable trade–off between the capabilities of sensing technology and the controller performance needs.

4.1 *IMC Synthesis*

Synthesis of a linear discrete–time IMC controller can be accomplished using a variety of techniques (discretization of a continuous–time controller, discrete–time synthesis, etc.) (Ogunnaike *et al.*, 1994). In the present work, a discrete–time transfer–function model ($G(z)$) was constructed to approximate the linear Volterra series model. By appending a first–order filter (time constant ϕ) to the linear inverse, the discrete–time IMC controller was synthesized. The resulting transfer function model and LIMC controller were:

$$G(z) = \frac{-0.458}{z - 0.849} \tag{9}$$

$$Q(z) = \frac{(1-\phi)z - 0.849(1-\phi)}{-0.458z + 0.458\phi} \tag{10}$$

The quality of model fit can be seen in Figure 3.

Partitioned nonlinear inverse controller synthesis was accomplished using the approach of Doyle III *et al.* (1995), in discrete–time with the third–order diagonal Volterra model. If the linear inverse exists, then a nonlinear system $\mathscr{P}$ that can be partitioned as

$$\mathscr{P} = (\mathscr{L} + \mathscr{N}) = \mathscr{L}(\mathscr{I} + \mathscr{L}^{-1}\mathscr{N})$$

can be analytically inverted yielding:

$$\mathscr{P}^{-1} = (\mathscr{I} + \mathscr{L}^{-1}\mathscr{N})^{-1}\mathscr{L}^{-1}$$

In block diagram form, this can be constructed as shown in Figure 4. Here the linear controller ($\mathscr{L}^{-1}$) is the transfer function in Equation (9), written in difference equation form. The nonlinear controller component ($\mathscr{N}$) is comprised of the $\mathscr{D}_2$ and $\mathscr{D}_3$

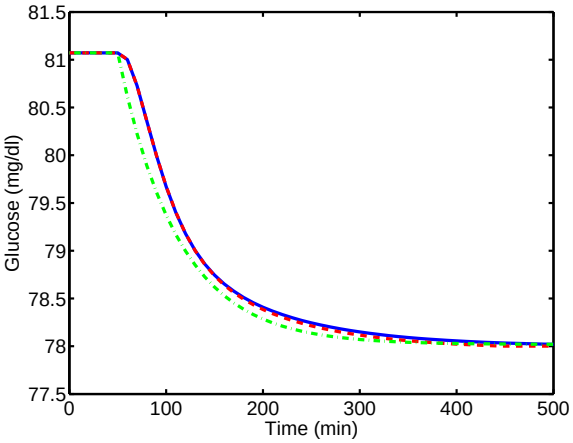

Fig. 3. Comparison of model dynamics. Solid: actual patient (continuous); dashed: linear Volterra; dash–dot: transfer function.

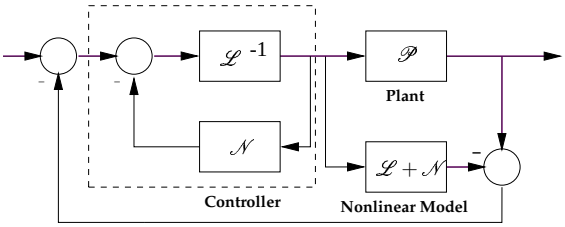

Fig. 4. PNLI–IMC schematic

components of the Volterra model (2). The difference equation formulation facilitates the "while" loop structure required for the controller loop to converge, an effect of the direct–feed nature of the controller and the static nonlinearities in the feedback loop. The convergence criterion is a difference in consecutive insulin infusion rate calculations of 1×10^{-4} mU/min.

The IMC controller (in both linear and PNLI–nonlinear forms) contains a single tuning parameter: the filter time constant. Since the controller is designed primarily to reject meal disturbances from a pseudo–steady state, the value of the filter constant could be selected to be small (in this case $\phi = 0.1$) to allow aggressive disturbance rejection. Should setpoint changes become a concern, or if sharp discontinuities were to occur in the measurement signal, the controller would have to be detuned significantly (to approximately $\phi = 0.7$) to maintain stability.

4.2 *MPC Synthesis*

Model predictive control is an algorithm that employs a process model to predict future dynamic behavior based on past inputs. This is an optimization–based control algorithm, executing at each sample time, and it uses the following objective function:

$$\min_{\Delta\mathscr{U}(k|k)} \|\Gamma_y[\mathscr{R}(k+1) - \mathscr{Y}(k+1|k)]\|_2^2 + \|\Gamma_u \Delta\mathscr{U}(k|k)\|_2^2 \tag{11}$$

Over a future prediction horizon of p steps, a series of $m \leq p$ manipulated variable moves is calculated in order to minimize the objective in Equation (11). The matrices Γ_y and Γ_u are used to trade off setpoint tracking error and manipulated variable movement, respectively. Additional constraints on the manipulated and controlled variables can be implemented in a straightforward fashion, as this is an optimization problem. In the case of a linear process model, the resulting problem is a quadratic programming problem. This changes to a nonlinear programming problems when nonlinear process models are employed. To solve the quadratic and nonlinear programming problems, the *fmincon* optimization routine of MATLAB (©2002, The Mathworks, Natick, MA) was employed.

Selecting the tuning parameters for a model predictive control algorithm is typically done on an *ad hoc* basis; there is no optimal tuning algorithm available. In general, the move and prediction horizons are adjusted to provide sufficient aggressiveness in control action, as well as adequate model prediction. The tuning matrices are used to alter the setpoint tracking performance (Γ_y) and to suppress noise–induced manipulated variable adjustment (Γ_u).

4.3 *Results and Discussion*

The results of using nonlinear compensation within the IMC framework to reject a meal disturbance of 50 grams (glucose) at time $t = 50$ min can be seen in Figure 5. A slightly more aggressive controller

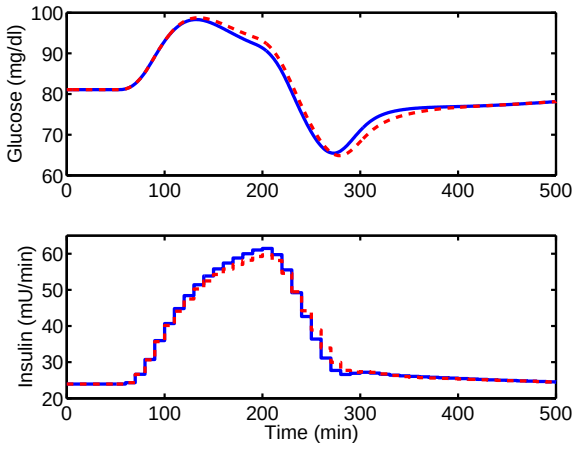

Fig. 5. 50 g meal disturbance simulation comparing LIMC (dashed) and PNLI–IMC (solid). The filter time constant was selected as $\phi = 0.1$ in both cases.

results from using the PNLI–IMC framework, as the nonlinear controller both increases and decreases the insulin delivery rate more rapidly than the linear controller. The sum–squared error (SSE) was reduced by 6.5%, with a small (3.7%) improvement in glucose concentration undershoot to a minimum of 65.5 mg/dl. Simulation results evaluating the model predictive control algorithm for the same disturbance as above are shown in Figure 6. The increased aggressiveness

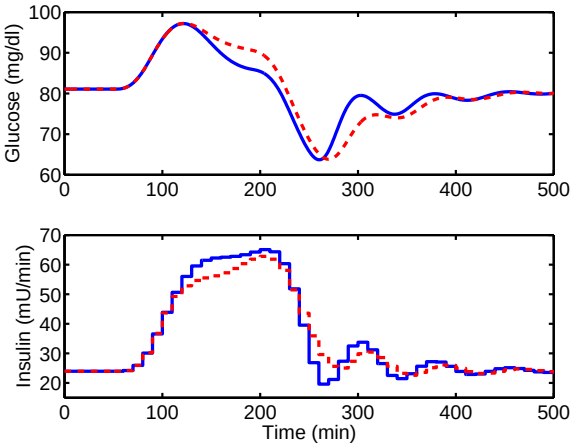

Fig. 6. 50 g meal disturbance simulation comparing MPC (dashed) and NMPC (solid). The tuning parameters for both controllers were $\Gamma_y = 3$, $\Gamma_u = 1$, $m = 3$, $p = 8$.

of the nonlinear controller can be clearly observed in the manipulated variable profile, where the insulin delivery rate is elevated more quickly and to a higher maximum delivery rate. This leads to the glucose concentration decrease observed between $t = 150$ and $t = 200$ min. Furthermore, the nonlinear controller decreases its delivery rate more quickly, thereby compensating more efficiently for the hypoglycemic excursion around $t = 270$ min. For an 0.3% increase in total insulin delivery, the nonlinear controller provides a 13% decrease in sum–squared error and a more rapid reaction to hypoglycemic excursions.

The increased ability to tailor the performance objective in MPC leads to a marked increase in performance as defined by sum–squared error, as shown in Table 2. This improvement in error is a

Table 2. Controller performance comparison. Absolute (mg^2/dl^2) and comparative (%) metrics versus linear IMC shown.

	Linear		Nonlinear	
	SSE	%	SSE	%
IMC	213	0	199	6.5
MPC	187	12	163	23.3

result of decreasing the magnitude and duration of the hyperglycemic excursion (between $t = 100$ and $t = 200$), as well as the return to steady state after the hypoglycemic excursion. The NMPC algorithm is particularly responsive to the depressed glucose concentrations between $t = 220$ and $t = 270$ min, and its aggressive response leads to superior performance. The fact that this nonlinear control algorithm responds so aggressively to the hypoglycemic excursion is imperative for diabetic patients in whom dramatically suppressed glucose levels can lead to coma and death. One minor penalty for improving the SSE is increased

undershoot. Both MPC controllers lead to minimum glucose concentrations of about 2 mg/dl less than the IMC controllers. However, the difference is well within the noise band of present sensors.

5. CONCLUSIONS

This paper presents an analysis of linear and nonlinear model–based control algorithms as employed in simulation on insulin–dependent diabetic patients. By employing a previously–developed parsimonious nonlinear Volterra series model (Florian and Parker, 2002) in the control structure, nonlinear control algorithms (IMC with partitioned inverses and MPC) were synthesized. These algorithms were then tested with respect to their capabilities in meal disturbance rejection. Nonlinear compensation proved beneficial, especially as the patient deviated from the nominal condition of 81 mg/dl. Given the performance metric of SSE, the NMPC controller is superior; however, if constraints are not imposed and a closed–form controller solution is required then linear MPC would be the best choice. In all cases, the utility of empirical model structures identified from patient data have proven effective in controlling glucose concentration in the presence of meal disturbances.

6. ACKNOWLEDGMENTS

The authors would like to acknowledge partial funding for this work from the Fulton C. Noss Faculty Fellowship of the University of Pittsburgh and the Pittsburgh Tissue Engineering Initiative.

7. REFERENCES

Boyd, S. and L. O. Chua (1985). Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Trans. Cir. Sys.* **CAS-32**(11), 1150–1161.

Cohen, A. (1993). New disposable electronic micropump for parenteral drug delivery. In: *Pulsatile Drug Delivery: Current Applications and Future Trends* (R. Gurny, H.E. Junginger and N.A. Peppas, Eds.). pp. 151–161. Wissenschaftliche, Stuttgart.

DCCT - The Diabetes Control and Complications Trial Research Group (1993). The effect of intensive treatment of diabetes on the development and progression of long–term complications in insulin–dependent diabetes mellitus. *N. Engl. J. Med.* **329**, 977–986.

DCCT - The Diabetes Control and Complications Trial Research Group (1996). The absence of a glycemic threshold for the development of long–term complications: The perspective of the diabetes control and complications trial. *Diabetes* **45**, 1289–1298.

Doyle III, F. J., B. A. Ogunnaike and R. K. Pearson (1995). Nonlinear model-based control using second-order Volterra models. *Automatica* **31**, 697–714.

Florian, J. A. and R. S. Parker (2002). A nonlinear data–driven approach to type I diabetic patient modeling. 15th IFAC World Congress on Automatic Control. IFAC.

Jaremko, J. and O. Rorstad (1998). Advances toward the implantable artificial pancreas for treatment of diabetes. *Diabetes Care* **21**, 444–450.

Minimed Corporation (1999). The minimed 507c, products and specifications. URL: http://www.minimed.com/files/mm_098.htm.

Nowak, R. D. and B. D. Van Veen (1994). Random and pseudorandom inputs for Volterra filter identification. *IEEE Trans. Signal Processing* **42**(8), 2124–2135.

Ogunnaike, B. A., R. K. Pearson, N. Samardzija and J. D. Bomberger (1994). Low order emprical modeling for nonlinear systems. In: *IFAC Symposium on Advanced Control of Chemical Processes*. Kyoto, Japan. pp. 41–46.

Parker, R. S., D. Heemstra, F. J. Doyle III, R. K. Pearson and B. A. Ogunnaike (2001*a*). The identification of nonlinear models for process control using tailored "plant–friendly" input sequences. *J. Proc. Control* **11, Sp. Issue SI**, 237–250.

Parker, R. S., F. J. Doyle III and N. A. Peppas (2001*b*). The intravenous route to blood glucose control. *IEEE Eng. Med. Biol.* **20**, 65–73.

Parker, R. S., J. H. Ward, N. A. Peppas and F. J. Doyle III (2000). Robust H_∞ glucose control in diabetes using a physiological model. *AIChE Journal* **46**, 2537–2549.

Puckett, W.R. and E.N. Lightfoot (1995). A model for multiple subcutaneous insulin injections developed from individual diabetic patient data. *Am. J. Physiol* **269 (Endocrinol. Metab. 32)**, E1115–E1124.

Sorensen, J.T. (1985). A Physiologic Model of Glucose Metabolism in Man and its Use to Design and Assess Improved Insulin Therapies for Diabetes. PhD thesis. Department of Chemical Engineering, MIT.

Zheng, Q. and E. Zafiriou (1994). Control-relevant identification of Volterra series models. In: *Proc. American Control Conf.*. Baltimore, MD. pp. 2050–2054.

Zheng, Q. and E. Zafiriou (1995). Nonlinear system identification for control using Volterra-Laguerre expansion. In: *Proc. American Control Conf.*. Seattle, WA. pp. 2195–2199.