

# Modeling Type One Diabetes with a Linear Response Function Based Model

REU Summary Report

Matthew Bauerle and Xinyan Han

Advisers: Prof. Yaniv Plan and Prof. Roman Vershynin

Summer 2014

## Abstract

Type one diabetes occurs when the pancreas stops producing glucose-reducing insulin. Regular injections of exogenous insulin are required for survival but the amount is difficult to determine to maintain optimal blood glucose levels. The goal of this research is to create a new blood glucose model that can predict future blood glucose levels in response to carbohydrate and insulin stimuli. We use a linear combination of kernel functions and use least squares to find the response functions. This simple model is able to fit the existing data as well as sophisticated biologically based models and has some predictive value. Future research could add biological knowledge and variable kernel functions to this model.

## 1 Introduction

### 1.1 The Blood Sugar Control System

The body possesses many systems that regulate chemical concentrations in the blood. One of the most important systems controls blood sugar levels. Most of the cells in the body use glucose as fuel to operate. [3]. When food is eaten and digested, blood glucose rises. Then the pancreas receives signals to release insulin into bloodstream, letting cells use the glucose for growth and energy. The liver also plays a role in glucose regulation by storing glucose

in glycogen when insulin is detected and releasing it between meals (e.g. at night). In a type-one diabetes patient, we simulate this by delivering a burst of insulin in the first bite of food. Then a steady stream of insulin is released to control blood glucose.

## 1.2 Diabetes Description

Diabetes is a disorder that drastically impacts the bodies ability to handle glucose with insulin. The causes of diabetes are various or unknown but the primary result of the disease is persistent elevated blood glucose levels. For people with diabetes, the pancreas does not automatically release the correct amount of insulin to keep control of blood sugar, instead, the pancreas either produces little or no insulin, or the cells do not have an appropriate response to released insulin. Thus, blood glucose will increase and is excreted through urine. Although the blood glucose level is high, cells cannot use blood glucose for growth and energy, and the body loses its main source of fuel. The classic symptoms of untreated diabetes are weight loss, polyuria (frequent urination), polydipsia (increased thirst), and polyphagia (increased hunger)[4]. Its long-term complications include to damage to blood vessels, eyes, nerves.

There is no known cure for diabetes and management is usually essential for survival. It is divided into three major categories: type 1, type 2, and gestational. Gestational diabetes is a temporary resistance to insulin during pregnancy that affects approximately 18 percent of pregnancies [1]. Type 2 diabetes is the most common kind of diabetes and occurs the cells in the body become resistant to insulin. Type 1 diabetes occurs when the body's immune system destroys the insulin producing  $\beta$  cells of the pancreas [4]. Injections of exogenous insulin is required to survive.

## 2 Model

Our model takes in tuning parameters, insulin administered, and carbohydrates and outputs a predicted blood sugar value at a certain time. It uses response functions to predict the impact of a certain quantity of insulin of carbohydrates on the blood sugar level. For example, the response function,  $F^C(t)$ , gives the change in blood sugar for one gram of carbs eaten. Each response function is in turn divided into multiple kernel functions. The tuning

parameters are the weights for each kernel function. Each individual kernel function should look like a step function that is perhaps smoothed. The total response function should look like a smoothed step function that is zero for times less than zero (no influence before stimulus). The transient behavior of the blood sugar level is determined by the intermediate values of the response function. This is the primary behavior that we would like to predict because low or high blood sugar spikes are uncomfortable and often dangerous.

## 2.1 Detailed Description

Each response function (for carbs, insulin, and liver) is in the following format.

$$F(t) = \sum_{i=1}^m a_i \phi^\sigma(t - t_i) \quad (1)$$

Where  $\phi^\sigma$  is a kernel function and  $m$  is the number of kernel functions used to make a response function (typically 20). We use the probability density function of the Gaussian as the Kernel estimation for the derivative of the response function  $F'(t)$ . Thus, the kernel function for response function  $F(t)$  is the cumulative distribution function of the Gaussian.

$$\phi^\sigma(t) = \frac{1}{2} \left( 1 + \operatorname{erf}\left(\frac{t}{\sqrt{2}\sigma}\right) \right) \quad (2)$$

$\sigma$  determines the time scale of the kernel function.

Our function accepts data in  $m \times 2$  matrices. Each data point is a row with a time and a value. There are six data matrices: the insulin matrix, the carb matrix, the liver matrix, and their corresponding coefficient matrix. Consider the carb response function.  $A^C$  is the coefficient matrix for carbs shown below.

$$\begin{bmatrix} \tilde{t}_1^C & a_1^C \\ \tilde{t}_2^C & a_2^C \\ \vdots & \vdots \\ \tilde{t}_m^C & a_m^C \end{bmatrix}$$

This makes a response function in the following way.

$$F^C(t, A^C) = \sum_{i=1}^m a_i \phi^\sigma(t - \tilde{t}_i^C) \quad (3)$$

It is possible to reformulate it as a convolution of distributions. Convert the data into a distribution with the following identification.

$$A^C \rightarrow \mathbb{A}^C = \sum_{i=1}^m a_i^C \delta_{\tilde{t}_i^C} \quad (4)$$

Plugging 4 into 3 we get

$$F^C(t, A^C) = \phi^\sigma * \mathbb{A} = \sum_{i=1}^m \phi^\sigma * (a_i^C \delta_{\tilde{t}_i^C}) = \sum_{i=1}^m a_i^C \phi^\sigma(t - \tilde{t}_i^C) \quad (5)$$

The total response function is a combination of the individual response functions.  $c_i$  is the amount of carbs given at time  $t_i^C$

$$G(t, C, A^C) = \sum_{i=1}^n c_i F^C(t - t_i^C, A^C) \quad (6)$$

Formatting  $C$  as a distribution  $\mathbb{C}$  as in 4 gives

$$G(t, C, A^C) = \mathbb{C} * F^C(t, A^C) = \mathbb{C} * \mathbb{A}^C * \phi^\sigma \quad (7)$$

Evaluating the convolutions gives

$$G(t, C, A^C) = \sum_{i=1}^n \sum_{j=1}^m c_i a_j \phi^\sigma(t - t_i^C - \tilde{t}_j^C) \quad (8)$$

The total response function is

$$G(t, I, A^I, C, A^C, L, A^L) = \mathbb{C} * \mathbb{A}^C * \phi^{\sigma_C} - \mathbb{I} * \mathbb{A}^I * \phi^{\sigma_I} + \mathbb{L} * \mathbb{A}^L * \phi^{\sigma_L} \quad (9)$$

We would like to evaluate the model function at the blood sugar measurement times and minimize the difference between the model and the measurements. Define a function

$$M(A) = \|G(t, I, A^I, C, A^C, L, A^L) - B(t)\|_2^2 = \sum_k (G(t_k, I, A^I, C, A^C, L, A^L) - B(t_k))^2 \quad (10)$$

Our current goal is to find  $A$  that minimizes  $M$

$$\min_A M(A) \quad (11)$$

A more efficient way to evaluate the blood sugar response is to use Matlab's matrix operations to speed up computation. There are  $m$   $\tilde{t}_i$ 's and  $n$   $t_i$ 's. We make a weighting matrix by calculating the outer product of the  $\mathbf{a}$  and  $\mathbf{c}$  matrices.

$$W = \mathbf{c} \mathbf{a}^T \quad (12)$$

$$W_{ij} = c_i a_j \quad (13)$$

Now we need to create a time matrix so that we can evaluate the kernel function at many times at once.

$$T = t \mathbf{1}_n \mathbf{1}_m^T - \mathbf{t} \mathbf{1}_m^T - \mathbf{1} \tilde{\mathbf{t}}^T \quad (14)$$

The blood sugar response at time  $t$  is given by

$$W : \phi(T) \quad (15)$$

where  $:$  is the "dot product" of two matrices, i.e.  $A : B = \sum_i \sum_j a_{ij} b_{ij}$ . Another efficiency improvement is to treat the basal insulin as a continuous stream rather than a collection of impulse functions. The response is calculated by convolving the step functions from the basal data with the insulin response function.

### 3 Goals for Modeling

The goals of this research is to create a modular model that can be easily expanded and altered to better model data. The objectives are: First, create a model that fits existing data well. Second, see that the response functions are in fact reasonable. Third, predict when blood sugar will spike below or above healthy values.

### 4 Model fitting

We attempted many tests of the model to predict blood glucose behavior but most were unsuccessful. The base problem was a conflict between the model being able to fit the data and the model having realistic intermediate values. The metric we used to test the model's ability to fit data was the following.

$$E(M) = \frac{\|M(A)\|_2}{\|B - \bar{B}\|} \quad (16)$$

Where  $\bar{B}$  is the average blood sugar value. Thus, a value of one for the error metric indicates a garbage model that is no better than a simple average while a value of zero indicates a perfect fit. Our error metrics ranged from about 0.4 to 1 and will be given in the data section. We also try to use non-negative least squares (NNLS) to make the response functions more realistic. However, this destroyed the ability of the model to fit the data in all but one case (lunch-time data with daytime offsets).

## 5 Data

We tried several different ways to fit the data.

Error rate	No processing	Take Log	Take Log and Root
Low spikes(BS values <70), straight ls	0.9	0.7326	0.7356
High spikes(BS values >180), straight ls	0.1387	0.1701	0.1694
Low spikes(BS values <70), nnls	0.9401	0.7682	0.7713
High spikes(BS values >180), nnls	0.1474	0.1753	0.1747

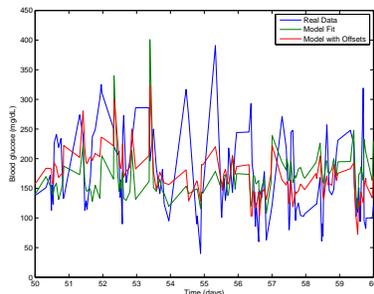


Figure 1: Actual blood sugar values compared with fit data.

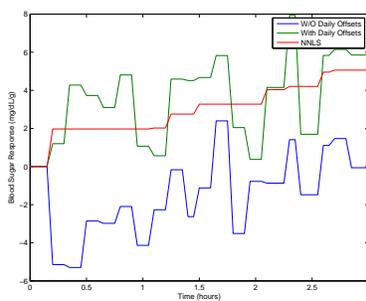


Figure 2: Lunchtime blood sugar response to glucose.

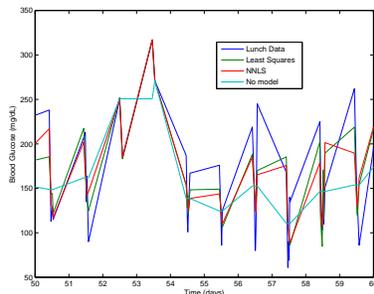


Figure 3: Lunchtime blood sugar data compared to model.

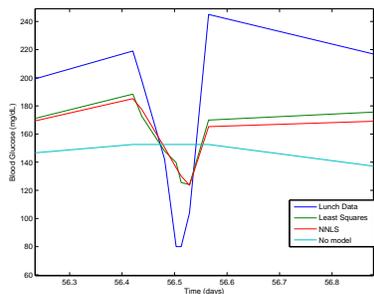


Figure 4: Lunchtime blood sugar data compared to model. (Detail)

## 6 Conclusion

The algorithm was able to fit the data fairly well compared to a random model. We noticed several factors that tended to improve the plausibility of the model. One problem is that the calculated dose response of both glucose and insulin tended to be the opposite of what biology suggests (i.e. glucose *decreased* blood sugar and insulin *increased* it.) This appears to be due to the fact that large insulin and glucose doses are given at meal times. This makes the data clustered and more difficult to work with. However, we were able to fit the data as well as existing models. Our model was able to fit approximately 22 percent of the data which is in the midrange in the models created in [2]. However, our model is much simpler and does not use as much biological information and has greater potential for improvement.

## References

- [1] ADA, *What is gestational diabetes?* <http://www.diabetes.org/diabetes-basics/gestational/what-is-gestational-diabetes.html>, 2014.
- [2] T. BRIEGEL AND V. TRESP, *A nonlinear state space model for the blood glucose metabolism of a diabetic*, 2002.
- [3] B. JM, T. JL, AND S. L., *Biochemistry*, W H Freeman, New York, 5 ed., 2002, ch. Section 30.2.
- [4] WEBMD, *Type 1 diabetes*. <http://www.webmd.com/diabetes/guide/type-1-diabetes>, May 2014.