

# Modeling and Analysis of Glucose in Blood with Bayesian Inference

Erika Ibarra Campos

A project submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of  
Master of Science

Tyler Jarvis, Chair  
Jared Whitehead  
Benjamin Webb

Department of Mathematics  
Brigham Young University

Copyright © 2024 Erika Ibarra Campos  
All Rights Reserved

## ABSTRACT

Modeling and Analysis of Glucose in Blood with Bayesian Inference

Erika Ibarra Campos

Department of Mathematics, BYU

Master of Science

Studying and developing non-invasive technologies for healthcare, particularly in glucose tracking for diabetic patients, holds immense significance on multiple fronts. It promises to revolutionize the quality of life by eliminating the need for frequent finger pricks, offering a more comfortable and convenient method of monitoring sugar in blood. One alternative method is through a spectrometer device that operates in the near-infrared spectrum. However, in order to study physiological data gathered with spectroscopy, it is important to also obtain knowledge of the nature of glucose readings from blood samples.

Utilizing Bayesian inference, we aim to adapt the acceptance-rejection MCMC algorithm, Metropolis-Hastings [3] [2], to produce an adequate posterior distribution that can describe the behavior of glucose levels in blood measurements. This will allow the creation of more data without the medical intervention to obtain blood samples, and to accurately aid the analysis of spectrometer data.

Keywords: Blood Glucose, Parameter estimation, Curve Modeling, Priors Distributions, Metropolis-Hastings, Independent Sampler, Posterior Distribution.

## ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to my parents, with special thanks to my mother, for their unwavering support throughout my academic journey. Their encouragement and belief in my potential have been invaluable. I am also profoundly thankful for my dear research mates, whose collaboration and camaraderie have enriched this experience. This research would not have been possible without the generous funding and data collection support from TULA Health and Octavian Solutions. Finally, but most important, I am immensely grateful to my advisor, Dr. Tyler Jarvis, for his patient guidance and steadfast support, which have been instrumental in the completion of this research.

# CONTENTS

<b>Contents</b>	<b>iv</b>
<b>List of Tables</b>	<b>v</b>
<b>List of Figures</b>	<b>vi</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Modeling</b>	<b>2</b>
2.1 Function Selection . . . . .	2
2.2 Substitution and Analysis of Parameters . . . . .	4
2.3 Prior Distributions . . . . .	7
2.4 First Prior Selection . . . . .	8
2.5 First Prior Predictive Test . . . . .	9
2.6 Second Prior Selection . . . . .	12
2.7 Second Prior Predictive Test . . . . .	14
<b>3 Metropolis-Hastings</b>	<b>17</b>
3.1 Algorithm Overview . . . . .	17
3.2 Adaptation . . . . .	19
<b>4 Results</b>	<b>20</b>
4.1 First Set of Posterior Samples . . . . .	21
4.2 Second Set of Posterior Samples . . . . .	25
<b>5 Conclusion</b>	<b>29</b>
<b>Bibliography</b>	<b>31</b>

## LIST OF TABLES

2.1	First set of prior distributions for each parameter. . . . .	8
2.2	Second set of prior distributions for each parameter . . . . .	13

## LIST OF FIGURES

1.1 Three blood glucose runs for patient 130 . . . . .	2
2.1 Curve modeling . . . . .	3
2.2 Actual data vs generated data with the modeling equations . . . . .	4
2.3 Sample curves with the two modeling equations with randomly selected parameters . . . . .	5
2.4 Sample curves with the substitution of $W$ in $f_2$ and parameter variation . . . . .	7
2.5 First prior predictive test. Patient 130 at two different dates. Date 2: offset = $0, t_I = 6$ . Date 3: offset = $10, t_I = 7$ . . . . .	10
2.6 First prior predictive test. Patient 131 at three different dates. Date 1: offset = $15, t_I = 8$ . Date 2: offset = $10, t_I = 7$ , Date 3: offset = $7, t_I = 7$ . . . . .	11
2.7 First prior predictive test. Patient 131 at three different dates. Date 1: offset = $15, t_I = 8$ . Date 2: offset = $10, t_I = 7$ , Date 3: offset = $15, t_I = 8$ . . . . .	12
2.8 Second prior predictive test. Patient 130 at two different dates. Date 2: offset = $0, t_I = 6$ . Date 3: offset = $10, t_I = 7$ . . . . .	14
2.9 Second prior predictive test. Patient 131 at three different dates. Date 1: offset = $15, t_I = 8$ . Date 2: offset = $10, t_I = 7$ , Date 3: offset = $7, t_I = 7$ . . . . .	15
2.10 Second prior predictive test with. Patient 131 at three different dates. Date 1: offset = $15, t_I = 8$ . Date 2: offset = $10, t_I = 7$ , Date 3: offset = $15, t_I = 8$ . . . . .	16
4.1 Metropolis-Hastings posterior samples and densities approximations. Patient 130, date 2 . . . . .	22
4.2 Metropolis-Hastings posterior samples and densities approximations. Patient 131, date 1. . . . .	23
4.3 Metropolis-Hastings posterior samples and densities approximations. Patient 149, date 1. . . . .	24

4.4	Metropolis-Hastings posterior samples and densities approximations. Patient 130 date 2 . . . . .	26
4.5	Metropolis-Hastings posterior samples and densities approximations. Patient 131 date 1 . . . . .	27
4.6	Metropolis-Hastings posterior samples and densities approximations. Patient 149 date 1 . . . . .	28

## CHAPTER 1. INTRODUCTION

In the realm of medical research and patient care, accurate modeling of blood glucose measurements plays a crucial role in understanding and managing conditions like diabetes. This project delves into the intricate task of predicting the parameters of a blood glucose model using Bayesian inference. By leveraging the power of Bayesian methods, we aim to harness the data gathered from laboratory tests to develop a probabilistic distribution that accurately describes glucose levels for patients.

To achieve this, we employ the rejection-acceptance algorithm known as Metropolis-Hastings [3][2]. This algorithm, combined with carefully constructed prior distributions, allows us to generate samples that best represent the glucose measurements of patients by fitting the parameters in the modeling. The ultimate goal is to provide a robust and flexible framework that healthcare professionals can utilize to better understand and manage blood glucose fluctuations in patients. Through this approach, we aspire to contribute to the advancement of personalized and data-driven patient care strategies.

Blood glucose measurements were collected from a group of diabetic patients within a controlled laboratory environment. Upon arrival, these patients underwent testing after fasting, during which initial baseline glucose values were documented. After this fasting reading, each patient was instructed to orally consume a 50 mg glucose tablet. This intake led to a noticeable increase in glucose levels, as reflected in the subsequent readings. To counteract and stabilize the rising glucose levels before they reach potentially dangerous thresholds, each patient received a tailored insulin injection. This personalized insulin dosage was administered to regulate and bring their glucose levels back within a normal range.



Figure 1.1: Three blood glucose runs for patient 130

Figure 1.1 provides a visual representation of this entire procedure, focusing on the glucose response curve for a specific patient identified as '130'. Within this figure, each curve corresponds to the blood glucose measurements obtained during a distinct testing date, called 'runs'. The x-axis corresponds to indices of time, while the y-axis the actual glucose amounts. The curves depict a pattern where the initial glucose values remain relatively constant, indicative of the fasting state. Following the ingestion of the glucose tablet, there is a pronounced spike in glucose levels. However, as the administered insulin begins to take effect, the glucose values gradually decrease, returning to stable values. It is important to note that the time indices vary per run, since not always the highest reading is recorded at the same time as the others. Hence, the spiking occurring at different times in each run.

## CHAPTER 2. MODELING

### 2.1 FUNCTION SELECTION

To capture the underlying structure of the curves, two distinct functions were chosen to represent the first and second halves of the overall curve. The initial portion of the curve is best described by a logistic function, while the latter part resembles a negative exponential

function. Figure 2.1 visually represents the division of the curves and their corresponding functions. The mathematical representations for these functions are:

$$f_1(t_1, G) = x_{(i,j)0} + G \frac{a}{1 + e^{-b(t_1 - c)}} \quad (2.1)$$

$$f_2(t_2, t_I, I) = We^{I(t_2 - t_I)K} + Z. \quad (2.2)$$

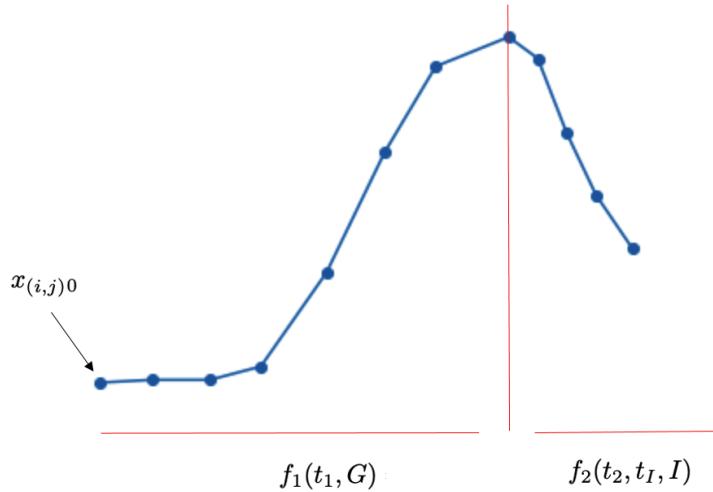


Figure 2.1: Curve modeling

In equation (2.1),  $x_{(i,j)0}$  represents the initial glucose value recorded,  $i$  is the patient number,  $j$  is the specific run. For simplicity, we'll refer it as simply  $x_0$ . Then  $G$  denotes the constant amount of glucose ingested, which is always 50 mg for all patients, all within the time interval  $t_1$  which encompasses the readings up to the ingestion of the tablet. Given the shape of the glucose increase graph, it is most appropriately modeled using a logistic function of time.

In equation (2.2), the function is dependent on  $I$ , representing the variable insulin amount administered to the patient at time  $t_I$ . As the insulin is injected, and considering the graph's shape, an exponential decay function is deemed the most suitable model, all within the time interval  $t_2$  which encompasses the readings after the insulin is injected. The variables  $W$ ,  $K$ ,

$c$ , and  $Z$  serve as scaling factors to appropriately adjust the curve's shape and magnitude. Figure 2.2 shows the simulation of data with default parameters against the real data.

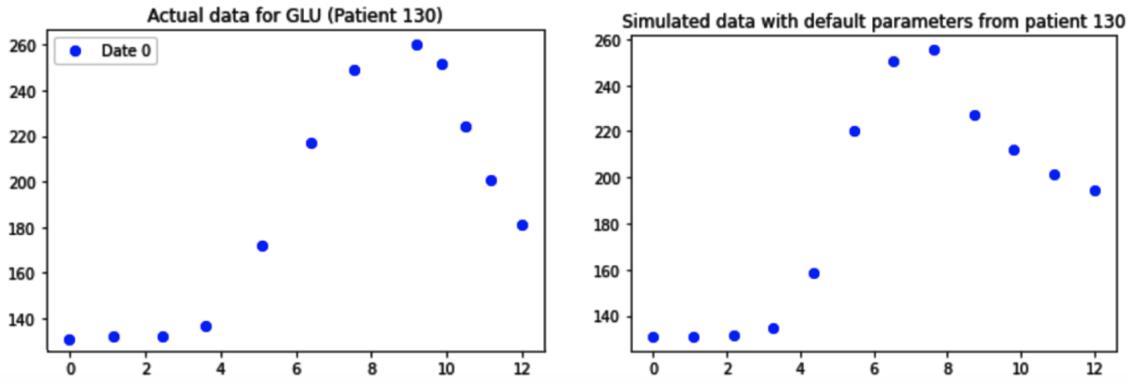


Figure 2.2: Actual data vs generated data with the modeling equations

## 2.2 SUBSTITUTION AND ANALYSIS OF PARAMETERS

Upon further examination of the curve formation for both equations, some unexpected behavior was observed. We selected parameters by observation, which are random yet suitable for both equations. Certain resulting curves exhibited an abrupt vertical line between the endpoints and start points of curves  $f_1$  and  $f_2$  respectively, as shown in the middle plot of Figure 2.3. Despite only minor variations in parameters, these changes significantly impacted the plotting, leading to inaccuracies in the curve rendering. Between curve 2 and curve 3, the only change was the initial glucose amount, i.e.  $x_{(i,j)_0}$ , which is a highly varied parameter since glucose is subjective to each patient and many biological factors. This finding suggests that even small parameter adjustments can drastically affect the overall shape and continuity of the curves, highlighting the sensitivity of the system to parameter selection.

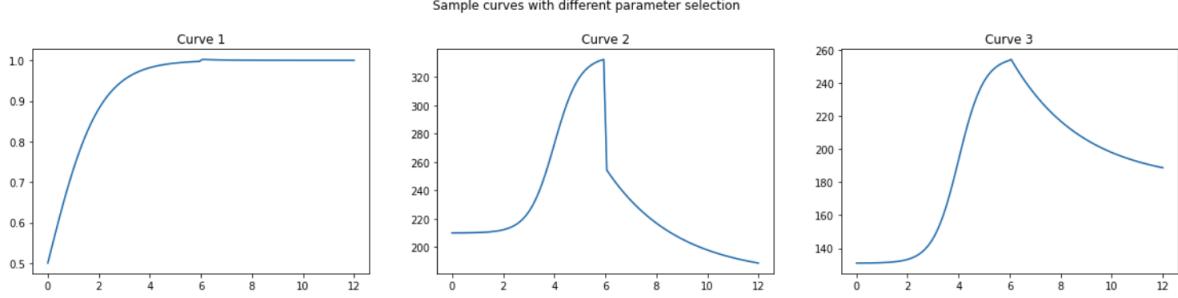


Figure 2.3: Sample curves with the two modeling equations with randomly selected parameters

Given the implications for the MCMC algorithm's predictive accuracy using these curves, we implemented a parameter substitution strategy. Specifically, we apply parameter substitution to ensure that the endpoint of the curve  $f_1$  always aligned precisely with the starting point of curve  $f_2$ . This was achieved by expressing  $f_2$  in terms of  $f_1$ , thereby reducing the number of independent parameters to just six.

Given that the values of both curves are the same at time  $t_I$ :

$$f_1(t_I) = f_2(t_I). \quad (2.3)$$

We solve for  $W$  from curve  $f_2$  in terms of  $f_1$ :

$$x_{(i,j)0} + G \frac{a}{1 + e^{-b(t_1 - c)}} = W + Z. \quad (2.4)$$

Resulting in the substitution of  $W$  in function  $f_2$  with:

$$x_{(i,j)0} + G \frac{a}{1 + e^{-b(t_1 - c)}} - Z = W. \quad (2.5)$$

Once this substitution was made, better results were observed. The choice of random parameters didn't affect the connectivity of the overall curve. Figure 2.4 shows the effect of these choices. Curve 1 depicts a desirable curve with selected parameters of  $\theta = [140, 1, 3, 4, 0.02, 150]$  for  $x_0, a, b, c, K, Z$  respectively, with standard amounts of 50 mg of ingested glucose and 18 mg of injected insulin. The starting amount of glucose is at 140 mg,

keeping steady levels around this mark until the time index of 3, where it is when the glucose tablet may be taken and the values increase rapidly. Later, at time index 6, the insulin is applied and the decay starts gradually, leveling off to an ending amount of 150 mg.

In curve 2, the parameter  $a$  was increased dramatically, which shows a curve with a very high peak value of glucose surpassing the 600 mg level. As the negative exponential ratio in the denominator of the function  $f_1$  goes to zero, the curve reaches the highest levels at  $x_0 + Ga$ , making  $a$  crucial to regulate the highest glucose amount as well as the overall increment of these values.

In curve 3, the parameter  $b$  was dramatically decreased, showing how it scales the curvature of the logistic portion, since a straight diagonal line is shown, suggesting that it reaches the highest point of glucose more linearly when decreased.

In curve 4, the parameter  $c$  was dramatically increased to show its effect at reaching the glucose level at time of insulin. A higher value leads to a flatter line in the logistic portion of the curve.

Lastly, in curve 5, the parameter  $K$  was dramatically increased. The high scaling for the negative exponential shows how sharp this portion of the overall curve becomes.

These modeling realizations provide crucial insights for selecting appropriate prior distributions for each parameter. Choosing suitable bounds and density functions that accurately reflect the nature of these scalars and amounts is essential for the success of the MCMC algorithm and for making accurate predictions.

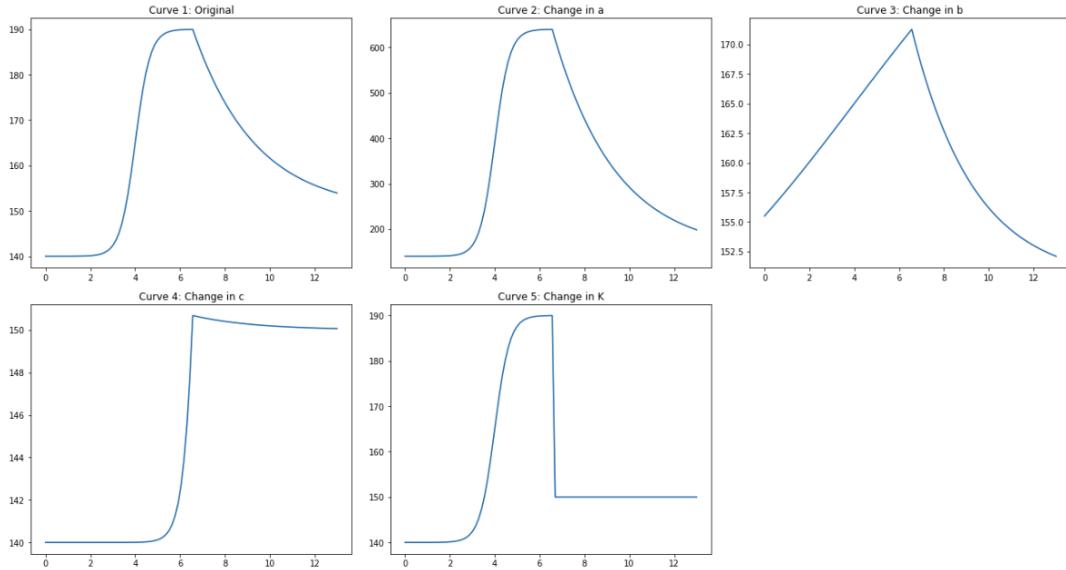


Figure 2.4: Sample curves with the substitution of  $W$  in  $f_2$  and parameter variation

### 2.3 PRIOR DISTRIBUTIONS

Choosing prior distributions for parameters is a fundamental aspect of Bayesian statistical analysis. In this framework, prior distributions represent the initial beliefs or knowledge about the parameters before observing any data. These priors encapsulate the uncertainty or assumptions about the parameters and play a crucial role in shaping the posterior distribution, which combines the prior information with the observed data. Selecting appropriate priors is critical as they influence the inference, model predictions, and the robustness of the conclusions drawn from the analysis.

In the following subsections, we propose two different sets of prior distributions for the model parameters. The first set was initially considered, but after further analysis, a second set of distributions was developed and is believed to better represent the overall behavior of the curve. Nevertheless, it is valuable to present the testing, analysis, and results for both sets.

## 2.4 FIRST PRIOR SELECTION

The insights gained from observing how the parameters interact and affect the overall curve emphasize the importance of careful prior distribution selection. Properly defining the bounds ensures that the parameter values remain within realistic and meaningful ranges, preventing unrealistic or non-physical results. Furthermore, selecting density functions that capture the true variability and distribution of the parameters helps in accurately modeling the underlying processes.

We first experimented with the following priors in Table 2.1 for each parameter:

Parameter	Prior Distribution
$x_0$	$ \mathcal{N}(150, 35) $
$a$	$ \mathcal{N}\left(\frac{350-x_0}{G}, \frac{100}{G}\right) $
$b$	$\text{Exp}\left(\frac{2}{\ln(2)}\right)$
$c$	$\text{Exp}\left(\frac{t_I}{2}\right)$
$K$	$ \mathcal{N}(0.01, 0.01) $
$Z$	$ \mathcal{N}(x_0, 35) $

Table 2.1: First set of prior distributions for each parameter.

We observed that patients' initial glucose values (fasting), denoted as  $x_0$ , typically averaged to 150 mg. To represent this parameter, we chose a positive-normal distribution, by taking the absolute value, with a mean of 150 mg and a standard deviation of 35 mg to encompass reasonable positive starting values.

In conjunction with  $x_0$ , we considered the value of the ending glucose level, denoted as  $Z$ . For simplicity, we also modeled  $Z$  with a positive-normal distribution centered at the initial value  $x_0$  with the same standard deviation of 35 mg.

For the parameters  $b$  and  $c$ , since they are scaling ratios corresponding to the negative

exponential equation in the denominator for the logistic portion, we ensure to bound them as much as possible to not affect the shape. From figure 2.4 we note that  $b$  determines the steepness of the logistic curve. Therefore, we chose an exponential distribution to ensure positive values, centered at  $\frac{2}{\ln(2)}$ , since we assume the increase to happen around the time index of 2. As where  $c$  represents the midpoint of the logistic curve, where the curve transitions from low to high values. Which is exponentially distributed around the midpoint of half-time of insulin  $t_I$ .

Similarly,  $K$  is the decay rate for the exponential term. The values for such parameter must be very small and positive. If the value is too big, as seen in figure 2.4, the decay sharpens unrealistically.

The highest glucose value per patient and even per run will differ. The parameter  $a$  allows the scaling of the glucose constant  $G$ , so it must be bounded so it does not reach an unnatural level of glucose. However, certain patients will acquire a dangerous reading after the glucose tablet ingestion, which happens at  $x_0 + Ga$ . We found appropriate to use 350 mg as an averaged high amount of glucose. We use this amount subtracted with the initial amount and normalized with the milligrams of the glucose tablet as the center of the distribution, with also a normalized standard deviation of 100. We keep a high variation due to the uncertainty of glucose in blood.

## 2.5 FIRST PRIOR PREDICTIVE TEST

Prior predictive tests or checks involve generating data based on the prior distributions to assess the suitability of the prior assumptions. This process helps to ensure that the chosen prior does not lead to implausible or unrealistic data outcomes. By doing so, we can refine their priors to better reflect their understanding and expectations about the parameters of interest, as we will discuss in the next section.

Using random sampling, we draw the parameter vector  $\theta$  from the respective prior distributions and evaluate the overall curve  $f(\theta)$  at each time point. This process generates a

range of possible curves based on the prior assumptions. As illustrated in the figures 2.5, 2.6, and 2.7, the resulting sampled curves encompass a broad spectrum of plausible values. Each individual curve differs on the time that glucose readings were captured. For instance, for patient 130 in figure 2.5, at date 2 the highest glucose reading occurred at time index 6, while on date 3 it occurred at time index 7. We make sure to incorporate an offset value that will shift the time domain of the samples generated according to the time of max glucose, which is the time where the insulin injection is administered, i.e.  $t_I$ .

When tested on three different patients across multiple runs, the priors produce curves that exhibit a good shape and variety of curves, effectively representing the uncertainty in glucose measurements according to the laboratory procedures. This demonstrates that the chosen priors are capable of capturing the variability inherent in the glucose measurement process.

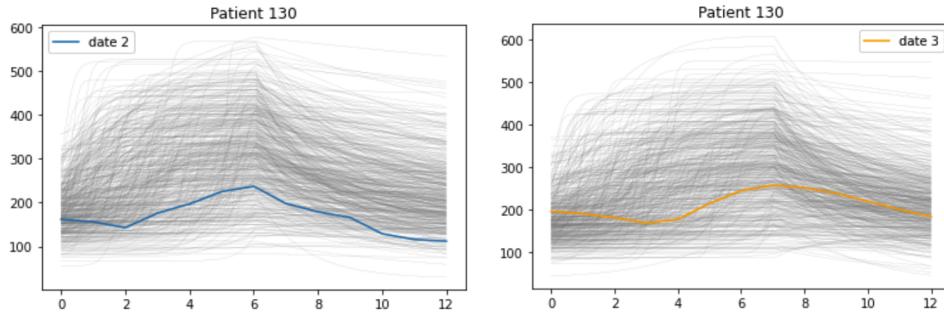


Figure 2.5: First prior predictive test. Patient 130 at two different dates. Date 2: offset = 0,  $t_I = 6$ . Date 3: offset = 10,  $t_I = 7$ .

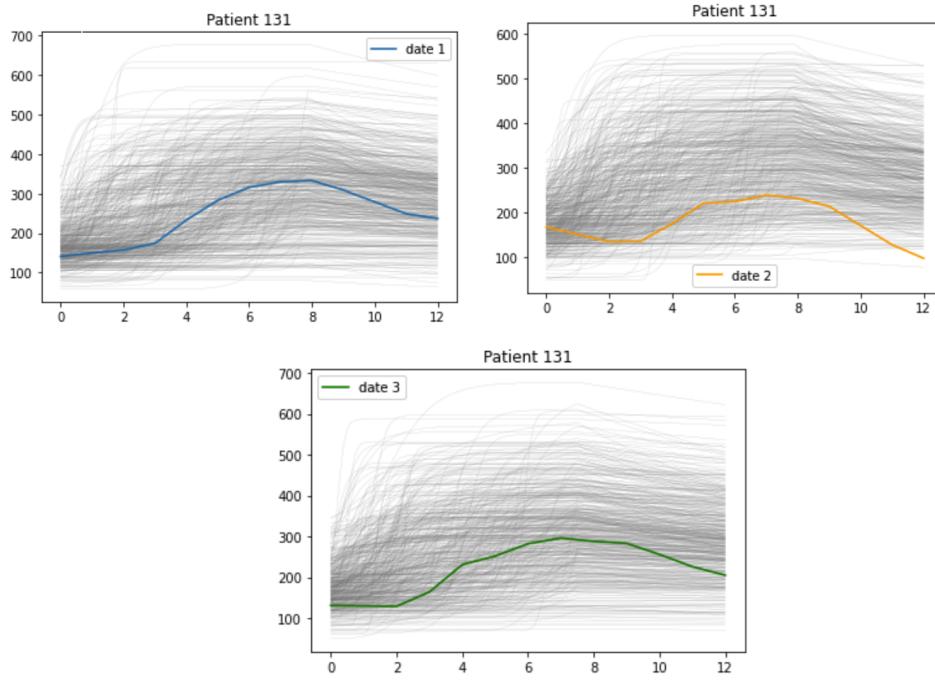


Figure 2.6: First prior predictive test. Patient 131 at three different dates. Date 1: offset = 15,  $t_I = 8$ . Date 2: offset = 10,  $t_I = 7$ , Date 3: offset = 7,  $t_I = 7$ .

It is crucial to observe that the samples generated with these choices of priors exhibit a noticeable lack of curvature. Specifically, the transition between the initial upward slope and the flattening in the first half of the curve, which corresponds to the logistic phase, is quite abrupt. Furthermore, in some samples, the second half of the curve shows little to no decay, with the lines extending almost horizontally. This flatness indicates a failure to capture the expected downward trend. Both cases potentially due to the limitations imposed by the selected priors for the scaling parameters.

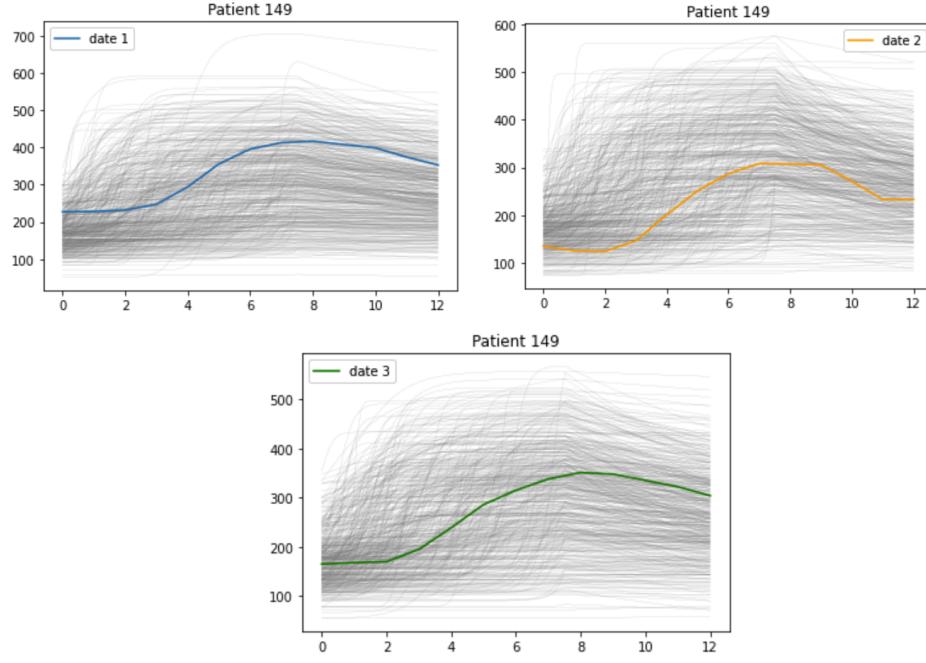


Figure 2.7: First prior predictive test. Patient 131 at three different dates. Date 1: offset = 15,  $t_I = 8$ . Date 2: offset = 10,  $t_I = 7$ , Date 3: offset = 15,  $t_I = 8$ .

## 2.6 SECOND PRIOR SELECTION

To address the limitations observed in the initial samples, we opted to use a different set of distributions for the curve parameters by selecting distributions that better capture the expected curvature in the data. This adjustment allowed us to generate curves with smoother transitions and more realistic behavior in both the logistic portion and the decay phase. The new parameter distributions were chosen to ensure that the model could more accurately reflect the underlying patterns in the data, thereby producing results that are not only more representative but also more aligned with theoretical expectations.

In light of this, we propose the following priors:

Parameter	Prior Distribution
$x_0$	$\mathcal{N}(150, 35)$
$a$	$ \mathcal{N}\left(\frac{250-x_0}{G}, \frac{100}{G}\right) $
$b$	$\Gamma\left(shape = \frac{100}{3} \cdot \frac{4}{G \cdot a}\right)$
$c$	$\mathcal{N}\left(\frac{t_I+2}{2}, 0.25\right)$
$K$	$\Gamma\left(shape = \frac{\ln(2)}{0.2} \cdot 2, scale = 0.2\right)$
$Z$	$\Gamma\left(shape = \frac{x_0}{0.3}, scale = 0.3\right)$

Table 2.2: Second set of prior distributions for each parameter

Note that the main and most important alterations of priors is for the parameters  $b$ ,  $K$ , and  $Z$ , with Gamma distributions ensuring positivity. As well as some minor edits to the rest of the parameters. The parameter  $b$  controls the steepness of the logistic curve. Its prior is a Gamma distribution, which is always positive and skewed. The shape parameter  $\frac{100}{3} \cdot \frac{4}{G \cdot a}$  was determined by imposing an approximation of a slope of  $\frac{100}{3}$  and setting it equal to the ratio  $\frac{4}{G \cdot a}$  which is the derivative of slope at time  $c$ , and solving for  $b$ . The skewness of the Gamma distribution indicates a tendency toward smaller values, especially when  $G$  or  $a$  are large, leading to a less steep logistic curve.

For the parameter  $K$ , we set a slope of  $\frac{1}{2}$ , which we equaled it to the decay term  $e^{-IK(2-t_I)}$  and solved for  $K$  by taking the log. Note that  $t_2 = 2$  since we suggest that at the highest glucose level at time  $c$ , the exponential decay is two time indices after, such that  $(t_2 - t_i) \rightarrow 0$ . Resulting in a mean of  $\frac{\ln(2)}{I \cdot 2}$  and a choice of scale of  $\frac{1}{4}$ . These values are subject to vary however.

Lastly, for the parameter  $Z$  which we know represents a constant offset. The Gamma distribution with a shape parameter dependent on  $x_0$  suggests that  $Z$  varies based on the baseline of the initial glucose value. This relationship between  $Z$  and  $x_0$  suggests that the constant offset adjusts according to the initial glucose levels, potentially reflecting the underlying physiological state or the initial condition that the model is trying to capture.

The Gamma distribution's skewness also means that while higher values of  $Z$  are more probable as  $x_0$  increases, there is still some probability of observing lower values, reflecting a balance between expectation and variability.

## 2.7 SECOND PRIOR PREDICTIVE TEST

After running simulations using the second set of prior distributions, we observed that the choice of Gamma distributions for the scaling parameters significantly influenced the curvature in the transitions of both halves of the overall curve. Specifically, the logistic portion of the curve, now exhibits a smoother and more gradual transition compared to the sharper change seen with the previous prior settings. This suggests that the Gamma distribution, with its oblique nature and dependency on parameters has reduced the steepness of the logistic curve, leading to a more gradual rise or fall around the midpoint.

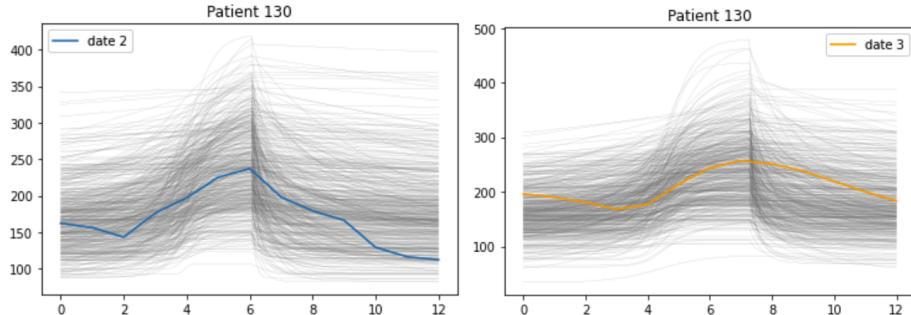


Figure 2.8: Second prior predictive test. Patient 130 at two different dates. Date 2: offset = 0,  $t_I = 6$ . Date 3: offset = 10,  $t_I = 7$ .

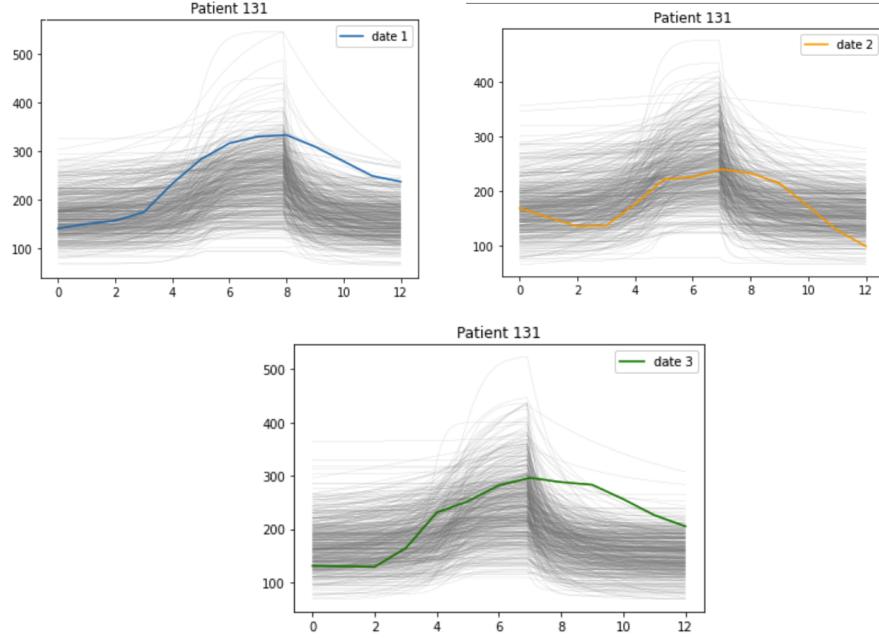


Figure 2.9: Second prior predictive test. Patient 131 at three different dates. Date 1: offset = 15,  $t_I = 8$ . Date 2: offset = 10,  $t_I = 7$ , Date 3: offset = 7,  $t_I = 7$ .

Similarly, in the exponential decay function, we notice that the number of samples producing nearly horizontal lines has decreased. This change indicates that the Gamma distribution used for scaling parameters like  $K$  has introduced more variability into the rate of decay, preventing the function from flattening out too quickly. As a result, the decay is more dynamic, with fewer instances of near-constant behavior, reflecting a more nuanced and realistic modeling of the exponential process.

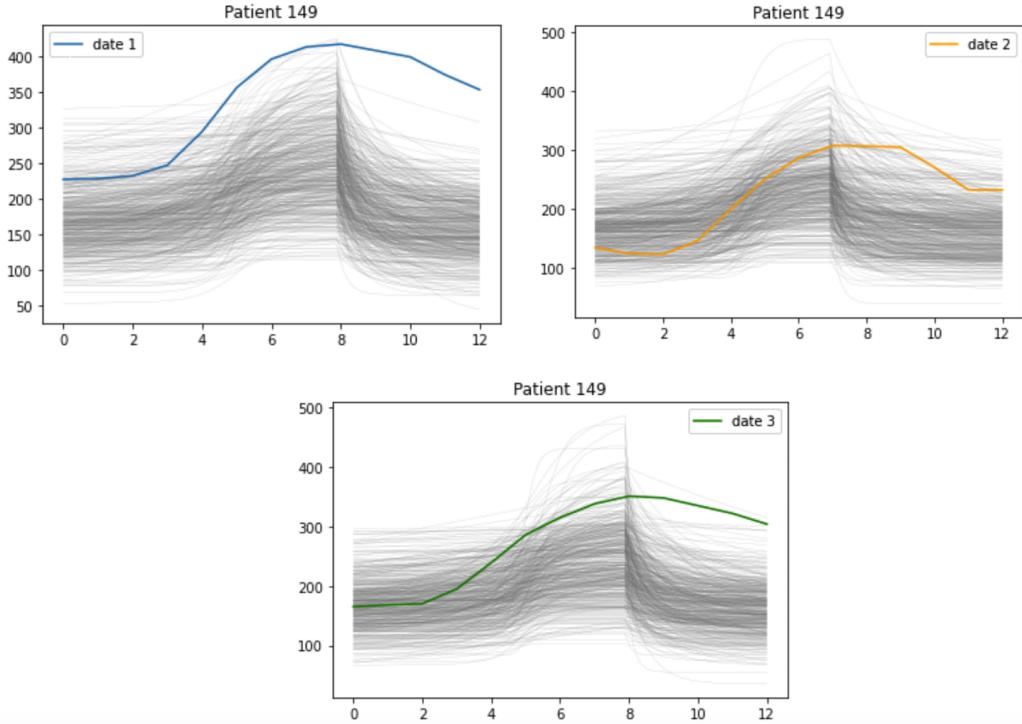


Figure 2.10: Second prior predictive test with Patient 131 at three different dates. Date 1: offset = 15,  $t_I = 8$ . Date 2: offset = 10,  $t_I = 7$ , Date 3: offset = 15,  $t_I = 8$ .

Although the new set of prior distributions improved the overall curvature of the sampled functions, we have encountered a new challenge, particularly in the second half of the curve. The exponential decay in this portion now tends to decay too rapidly, which has led to an unintended consequence—the function no longer levels out horizontally as it should in certain scenarios as observed in figure 2.9 and 2.10 for data for patients 131 and 149 respectively.

In the previous prior with an Exponential distribution, we observed that some samples exhibited a nearly horizontal behavior, allowing the exponential decay to slow down and eventually stabilize at a certain value. This characteristic is crucial for accurately modeling processes where a variable decreases rapidly at first but then stabilizes over time, such as the settling of a physical system or the gradual tapering of an effect.

However, with the new Gamma-distributed priors for the scaling parameters, the exponential decay has become much steeper. As a result, the function tends to drop off too quickly, and we no longer see the horizontal leveling out that would indicate a slow approach

to the ending glucose amount. This sharp decay might be suitable for modeling certain types of data where a quick drop is expected, as shown in 2.8 in date 2 for patient 130, such as in runs with a sudden decline or depletion.

On the other hand, for datasets that require a more gradual decline and eventual stabilization, this rapid decay could lead to inaccuracies. The model may struggle to capture the long-term behavior of the system, particularly when it is expected to approach a steady state or maintain a residual level over time. Therefore, while the new priors have enhanced the overall curvature, they have introduced a limitation in the flexibility of the transition to represent slow-decaying processes, making it less suitable for datasets requiring a more nuanced approach to exponential decay.

## CHAPTER 3. METROPOLIS-HASTINGS

The Metropolis-Hastings algorithm [3][2] is a fundamental technique in computational statistics, particularly within the field of Bayesian inference. Developed as an extension of the Metropolis algorithm, it serves as a Markov Chain Monte Carlo (MCMC) method for obtaining a sequence of random samples from a probability distribution for which direct sampling is difficult. By constructing a Markov chain that has the desired distribution as its equilibrium distribution, the algorithm iteratively proposes candidate samples and decides whether to accept or reject them based on a specific acceptance criterion. This algorithm is powerful and flexible, allowing for efficient exploration of complex models. Its ability to generate samples from posterior distributions makes it an essential tool for Bayesian analysis.

### 3.1 ALGORITHM OVERVIEW

The algorithm starts with an initial parameter value and iteratively proposes new states based on a proposal distribution. Each proposed state is evaluated using an acceptance ratio that compares the likelihood of the proposed state to the current state, adjusting for the

proposal distribution

- $p(\theta')$ : The target or posterior probability density of the proposed state  $\theta'$ .
- $p(\theta_{t-1})$ : The target or posterior probability density of the current state  $\theta_{t-1}$ .
- $q(\theta' | \theta_{t-1})$ : The proposal distribution used to generate  $\theta'$  from  $\theta_{t-1}$ .
- $q(\theta_{t-1} | \theta')$ : The proposal distribution used to generate  $\theta_{t-1}$  from  $\theta'$ .

A random number is then used to decide whether to accept or reject the proposed state.

This process continues for a specified number of iterations, generating a sequence of samples that approximate the target distribution. The algorithm effectively explores the parameter space and provides samples for estimating properties of the target distribution. See algorithm 1.

---

**Algorithm 1** Metropolis-Hastings Algorithm

---

- 1: Initialize  $\theta_0$  randomly
  - 2: **for**  $t = 1, 2, \dots, T$  **do**
  - 3:     Propose a new state  $\theta'$  based on the current state  $\theta_{t-1}$  using a proposal distribution  $q(\theta' | \theta_{t-1})$
  - 4:     Compute the acceptance ratio:
- $$\alpha = \min \left( 1, \frac{p(\theta') \cdot q(\theta_{t-1} | \theta')}{p(\theta_{t-1}) \cdot q(\theta' | \theta_{t-1})} \right)$$
- 5:     Generate a uniform random number  $u$  from  $U(0, 1)$
  - 6:     **if**  $u < \alpha$  **then**
  - 7:         Accept the new state:  $\theta_t = \theta'$
  - 8:     **else**
  - 9:         Reject the new state:  $\theta_t = \theta_{t-1}$
  - 10:   **end if**
  - 11: **end for**
  - 12: **Return** the sequence of states  $\{\theta_0, \theta_1, \dots, \theta_T\}$
- 

However, we will be adapting the log form of the Metropolis-Hastings algorithm [3][3], which is a variation that computes the acceptance ratio in logarithmic terms to enhance numerical stability and efficiency. By using logarithms, we avoid potential issues with very

small probabilities or large numerical values that can arise when working with raw probabilities. Specifically, the acceptance ratio is calculated as the difference between the logarithms of the probabilities, which is computationally more stable and prevents overflow or underflow errors. Additionally, comparing logarithmic values of a uniform random number and the log of the acceptance ratio simplifies the calculations and improves precision. This approach ensures that the algorithm remains robust and accurate even when dealing with extreme or very small probability values, making it particularly useful in high-dimensional or complex models where numerical stability is crucial. See algorithm 2.

---

**Algorithm 2** Metropolis-Hastings Algorithm (Logarithmic Form)

---

- 1: Initialize  $\theta_0$  randomly
- 2: **for**  $t = 1, 2, \dots, T$  **do**
- 3:     Propose a new state  $\theta'$  based on the current state  $\theta_{t-1}$  using a proposal distribution  $q(\theta'|\theta_{t-1})$
- 4:     Compute the log acceptance ratio:

$$\log \alpha = \log \left( \frac{p(\theta') \cdot q(\theta_{t-1}|\theta')}{p(\theta_{t-1}) \cdot q(\theta'|\theta_{t-1})} \right) = \log p(\theta') - \log p(\theta_{t-1}) + \log q(\theta_{t-1}|\theta') - \log q(\theta'|\theta_{t-1})$$

- 5:     Generate a uniform random number  $u$  from  $U(0, 1)$
- 6:     Compute the log of the random number:

$$\log u$$

- 7:     **if**  $\log u < \log \alpha$  **then**
  - 8:         Accept the new state:  $\theta_t = \theta'$
  - 9:     **else**
  - 10:         Reject the new state:  $\theta_t = \theta_{t-1}$
  - 11:     **end if**
  - 12: **end for**
  - 13: **Return** the sequence of states  $\{\theta_0, \theta_1, \dots, \theta_T\}$
- 

### 3.2 ADAPTATION

The three elements used in Metropolis-Hastings [3][2] are the proposal probability, the prior probability, and the likelihood. If we want to estimate the posterior distribution of  $\theta$  after observing some data  $\mathbf{x}$ , that is,  $p(\theta|\mathbf{x})$  given that we know the prior distribution  $p(\theta)$ . We

define our proposal distribution to draw proposals directly from the prior distributions. This is known as the independent sampler. This approach introduces variability into the parameter vector  $\theta$ . We also initialize  $\theta_0$  with the prior distributions. By assuming that  $\theta'$  and  $\theta_{t-1}$  are independent, we simplify the acceptance criterion, as described below.

Given the acceptance criterion:

$$\alpha = \min \left( 1, \frac{p(\theta' | \mathbf{x}) \cdot q(\theta_{t-1} | \theta')}{p(\theta_{t-1} | \mathbf{x}) \cdot q(\theta' | \theta_{t-1})} \right). \quad (3.1)$$

By applying Bayes rule [1], the the posterior probability becomes:

$$\alpha = \min \left( 1, \frac{p(\theta') \cdot p(\mathbf{x} | \theta') \cdot q(\theta_{t-1} | \theta')}{p(\theta_{t-1}) \cdot p(\mathbf{x} | \theta_{t-1}) \cdot q(\theta' | \theta_{t-1})} \right). \quad (3.2)$$

Since the current and proposed parameters are independent then:

$$\alpha = \min \left( 1, \frac{p(\theta') \cdot p(\mathbf{x} | \theta') \cdot p(\theta_{t-1})}{p(\theta_{t-1}) \cdot p(\mathbf{x} | \theta_{t-1}) \cdot p(\theta')} \right). \quad (3.3)$$

Leaving only the ratio of likelihoods:

$$\alpha = \min \left( 1, \frac{p(\mathbf{x} | \theta')}{p(\mathbf{x} | \theta_{t-1})} \right). \quad (3.4)$$

And by taking the log:

$$\log \alpha = \log \left( \frac{p(\mathbf{x} | \theta')}{p(\mathbf{x} | \theta_{t-1})} \right) = \log p(\mathbf{x} | \theta') - \log p(\mathbf{x} | \theta_{t-1}). \quad (3.5)$$

We define the log likelihood probability density function to be:

$$\log_{\mathbf{x}_i} p(\mathbf{x} | \theta, \sigma) = -n \log \sigma - \frac{n}{2} \log(2\pi) - \frac{1}{2\sigma^2} \sum_{t=1}^n (f_\theta(t) - \mathbf{x}_i(t))^2. \quad (3.6)$$

Where  $f_\theta$  is a vector of values of the overall curve  $f$  evaluated at the vector of parameters  $\theta$ . And  $\mathbf{x}_i(t)$ , which serves as our observed data, is distributed as  $N(f_\theta(t), \sigma)$ . Here  $\sigma$  is just a global standard deviation over each patient's data. Resulting in algorithm 3.

---

**Algorithm 3** Metropolis-Hastings Algorithm with Independent Sampler (Logarithmic Form)

---

- 1: Initialize  $\theta_0$  randomly from the prior  $p(\theta)$
- 2: **for**  $t = 1, 2, \dots, T$  **do**
- 3:     Propose a new state  $\theta'$  from the prior distribution  $p(\theta')$
- 4:     Compute the log acceptance ratio:

$$\log \alpha = \log \left( \frac{p(\mathbf{x}|\theta')}{p(\mathbf{x}|\theta_{t-1})} \right) = \log p(\mathbf{x}|\theta') - \log p(\mathbf{x}|\theta_{t-1})$$

- 5:     Generate a uniform random number  $u$  from  $U(0, 1)$
- 6:     Compute the log of the random number:

$$\log u$$

- 7:     **if**  $\log u < \log \alpha$  **then**
  - 8:         Accept the new state:  $\theta_t = \theta'$
  - 9:     **else**
  - 10:         Reject the new state:  $\theta_t = \theta_{t-1}$
  - 11:     **end if**
  - 12: **end for**
  - 13: **Return** the sequence of states  $\{\theta_0, \theta_1, \dots, \theta_T\}$
- 

## CHAPTER 4. RESULTS

After integrating all the previously discussed steps, we observe some promising, yet deviating results with both set of prior distributions as shown in the following subsections.

### 4.1 FIRST SET OF POSTERIOR SAMPLES

Figures 4.1, 4.2, and 4.3 present four chains with generated samples of the overall curve derived from the posterior distribution of the parameters, along with their corresponding density plots. For clarity, we have chosen to display the results for only one date per patient. Although the approximations are not perfect, they are sufficiently accurate and capture the necessary variability, which aligns with our primary objective of achieving a realistic representation of the data.

One notable observation is that the decay in the second half of the curves is steeper

than what was observed during the initial data testing. This steeper decay indicates that the priors we initially selected for the second half of the curve—particularly in modeling the exponential decay—are well-suited for certain patient cases, specifically patients 130 and 131. For these patients, the sharper decline matches the expected behavior, suggesting that the model effectively captures the dynamics of their data.

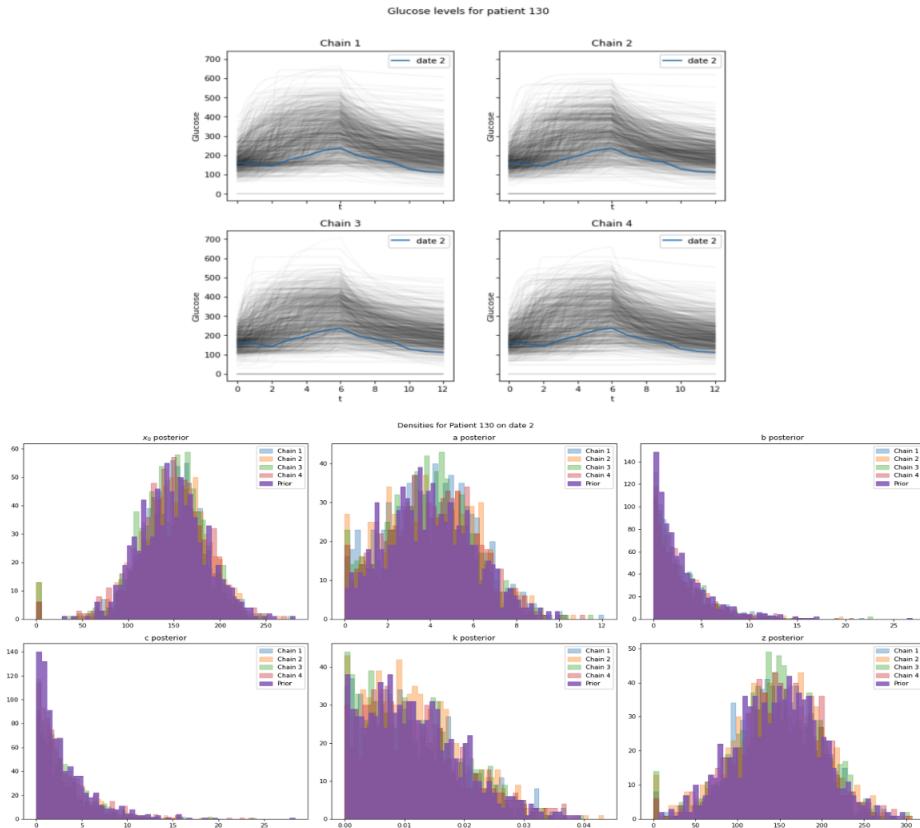


Figure 4.1: Metropolis-Hastings posterior samples and densities approximations. Patient 130, date 2

However, this is not the case for all patients. For patient 149, the decay remains too sharp, deviating significantly from what we would expect. In fact, the curve's behavior is so steep that the exponential decay almost resembles a logistic function rather than a negative exponential. This suggests that our chosen priors, while effective for some patients, may not be universally applicable. The overly steep decay for patient 149 indicates that the model may be overly constrained or that the priors need further adjustment to better accommodate the variations in decay rates among different patients.

The parameters appear to converge to adequate values after going through the algorithm. The density plots depict the number of appearance in the y-axis of a certain value in the x-axis. For the starting glucose value  $x_0$ , shows a Gaussian bell centered at 150, where the convergence seems to mainly range from 50 to 250, which is a good sign since this accounts for uncertainty in blood glucose. Similarly for  $Z$ , the convergence ranges from 50 to 300, which is a little higher than the starting value, but it is a correct approximation since the final glucose reading don't usually level off back to the initial amount of fasting.

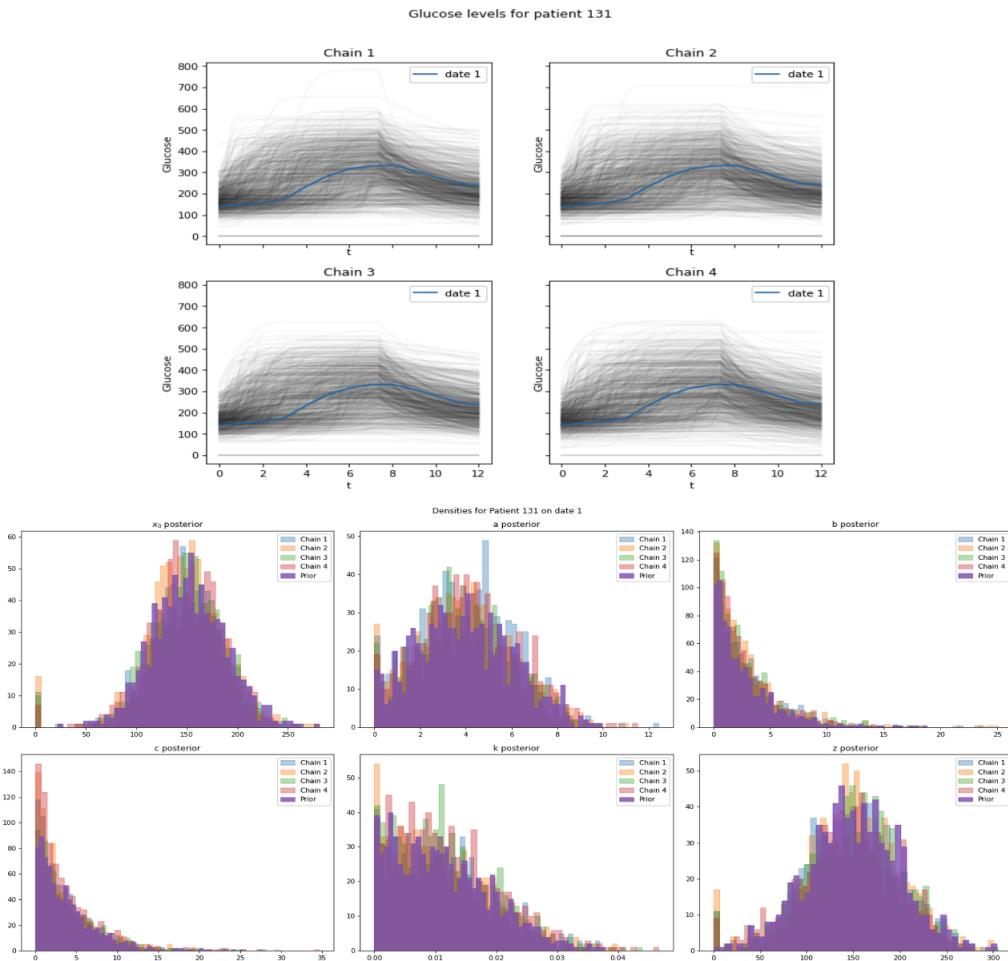


Figure 4.2: Metropolis-Hastings posterior samples and densities approximations. Patient 131, date 1.

For  $a$ , the values only range up to 12 and centered around 4, which adequately scales the highest glucose amount in the samples, giving room for high levels, but not too high such that the readings become unnatural.

Both parameters,  $b$  and  $c$ , have almost identical convergence. With  $c$  reaching some higher values. This is reasonable since both parameters maintain an equilibrium on scaling the logistic portion of the curve. They both ensure that the first half of the overall curve meets the highest glucose amount at insulin time correctly.

Finally, the convergence of  $K$  reaches values up to 0.04, which are a little higher than expected. Usually for this parameter, 0.02 is a good bound to stop such value. However, the choice of a prior distribution for this parameter allows it to have slightly bigger values. We can notice this discrepancy in the resulting approximations of the posterior as previously mentioned. The sharpness of the second half of the curves indicates an abrupt instead of smooth decay on glucose amounts after insulin.

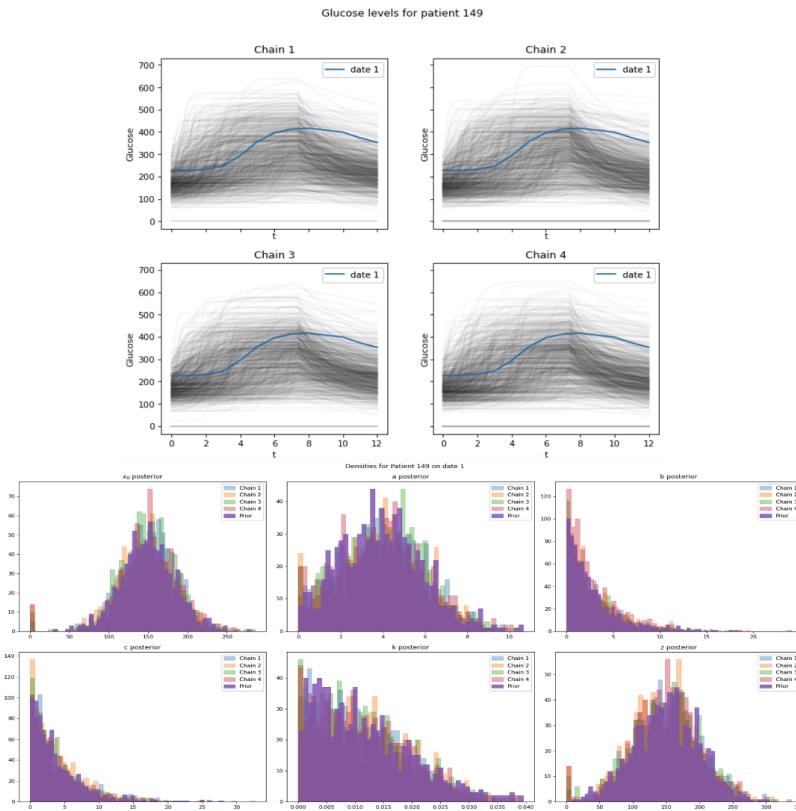


Figure 4.3: Metropolis-Hastings posterior samples and densities approximations. Patient 149, date 1.

The density plots also include the prior distributions overlaid on the posterior distributions of each parameter. This visualization provides insight into how the prior beliefs about

the parameters have influenced the final posterior estimates after observing the data.

What is particularly noteworthy is that the resulting posterior distributions closely align with the initial prior distributions. This close correspondence suggests that the data was not highly informative or that the priors were particularly strong, dominating the posterior outcomes. If the posterior distribution looks like the prior, it could mean that the new information hasn't significantly changed the prior.

This alignment can have both positive and negative implications. On the positive side, it indicates that our prior knowledge was consistent with the data, which may validate the assumptions made about the parameters' behavior. The model effectively captured the expected variability and trends as initially hypothesized, which can be beneficial in scenarios where we have strong prior knowledge that we trust to guide the inference process.

However, the fact that the posteriors remain so close to the priors also raises concerns. It suggests that the model might be under-responsive to the actual data, relying too heavily on the prior distributions, suggesting our likelihood function could be not informative enough.

## 4.2 SECOND SET OF POSTERIOR SAMPLES

Similarly, we present in this subsection in figures 4.4, 4.5, and 4.6 four chains with generated samples and the posterior parameter densities for the same patients.

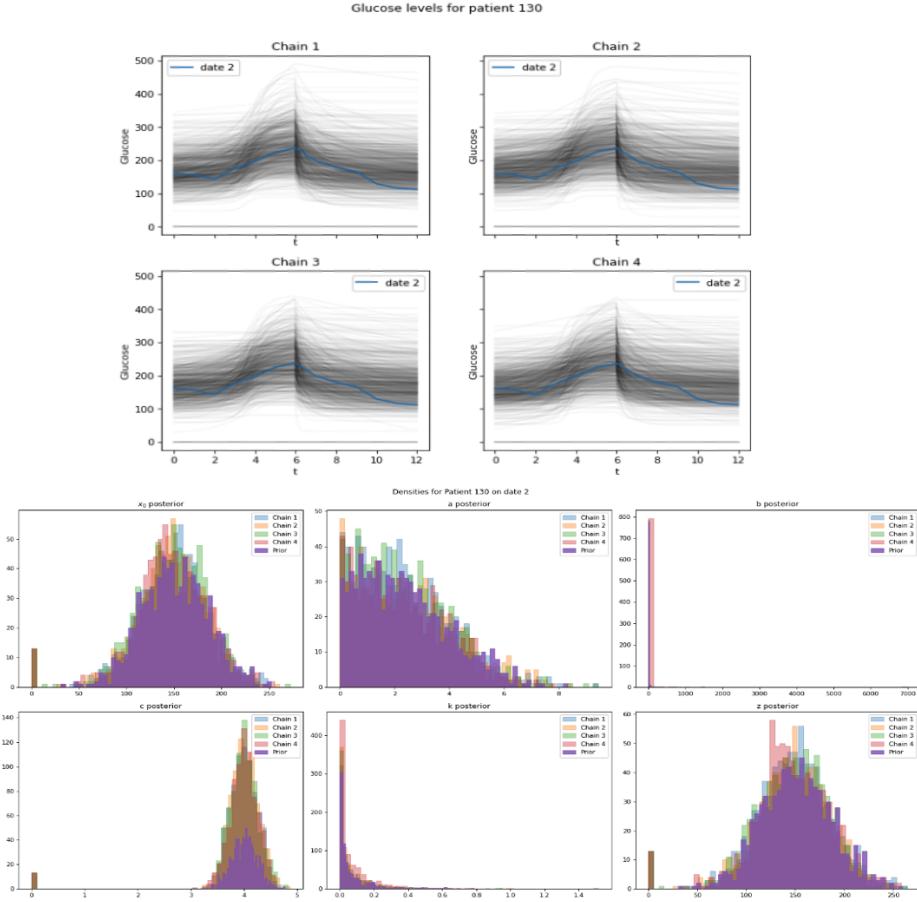


Figure 4.4: Metropolis-Hastings posterior samples and densities approximations. Patient 130 date 2.

With the implementation of the new set of prior distributions, the sampled curves for the logistic portion of the function show a marked improvement. The curvature is now more refined and precise, closely matching the expected behavior. This adjustment has resulted in a smoother and more accurate representation of the logistic transition, aligning well with the underlying data patterns we intended to model. The enhancement in the logistic portion is a clear indication that the revised priors have successfully captured the dynamics of this part of the curve.

However, the improvements in the logistic section have come with a trade-off. In the second half of the curve, which represents the exponential decay, the decay has become even steeper than in previous iterations. This sharper decline is problematic because it deviates further from the desired behavior of the second portion of the curve.

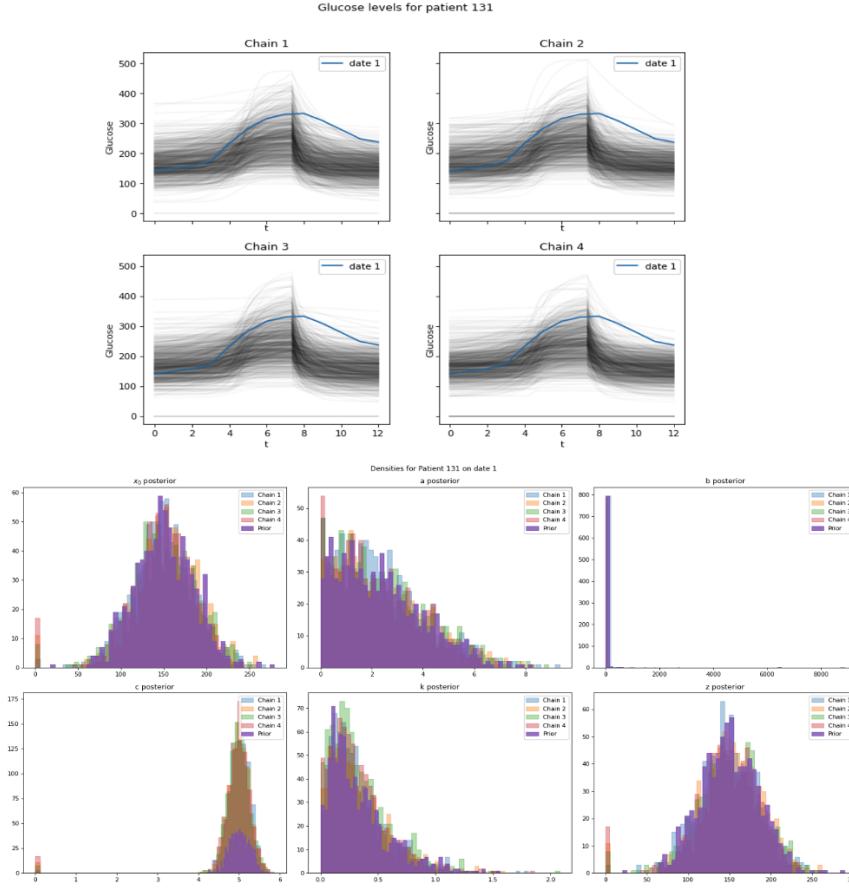


Figure 4.5: Metropolis-Hastings posterior samples and densities approximations. Patient 131 date 1.

Ideally, the exponential decay should taper off more gradually, reflecting a slower approach to an asymptote or a stabilization point. Instead, the current priors cause the curve to drop too quickly, reducing the flexibility of the model to accurately represent scenarios where a more moderate decay is expected.

The histogram plots still show a close density between the posteriors and the priors of the parameters. With similar values of convergence, the noticeable parameters that differ are the now higher values for  $b$  and  $k$ , which are represented on the steepness of the curve. The parameter  $b$  mainly stays at a constant range of values, as the histograms show one bin, while  $k$  shows more variability. The posterior densities for  $c$  seemed to increase in occurrence in contrast to the prior, with relatively similar values as before.

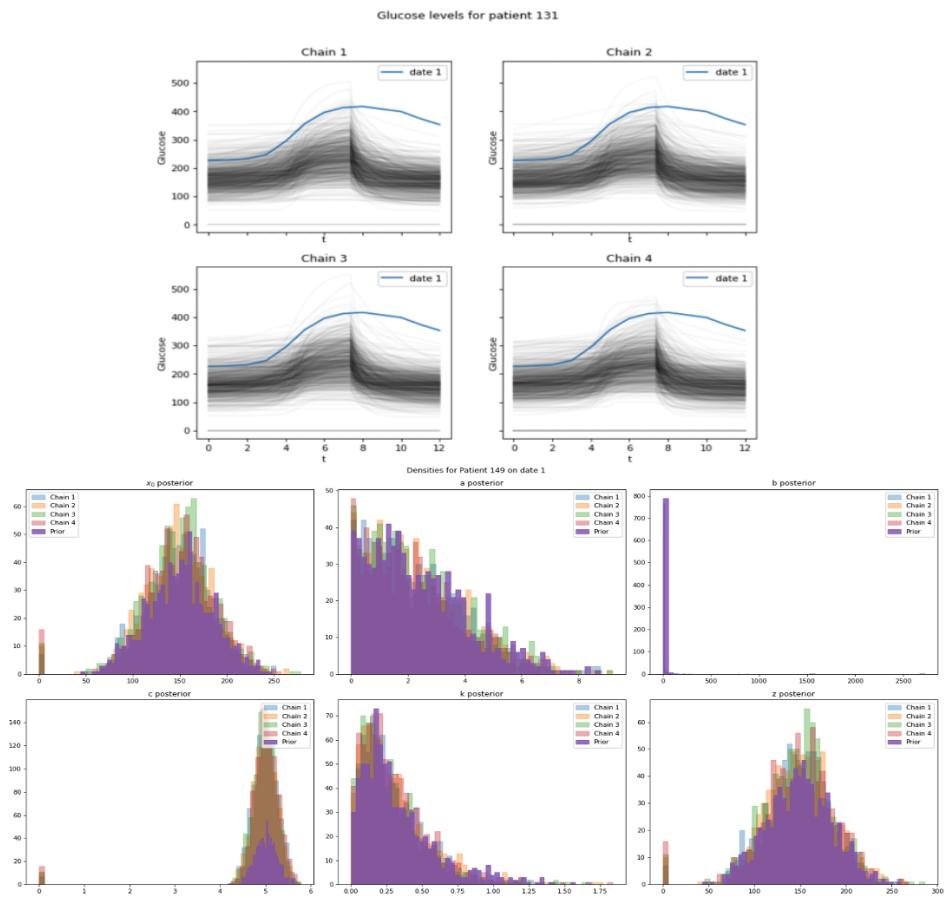


Figure 4.6: Metropolis-Hastings posterior samples and densities approximations. Patient 149 date 1.

## CHAPTER 5. CONCLUSION

Generating prior distributions for parameters that describe glucose levels in the blood is particularly challenging due to the complex and dynamic nature of glucose metabolism. Glucose levels are influenced by a multitude of factors including diet, physical activity, insulin sensitivity, and individual metabolic responses, all of which introduce significant variability. To accurately reflect this complexity in the priors, one must have a deep understanding of both physiological processes and the specific characteristics of the population under study. Moreover, the parameters themselves, such as initial glucose concentration, rate of glucose absorption, and insulin response, require careful consideration to ensure they encapsulate realistic ranges and behaviors. Inadequate or poorly chosen priors can lead to biased predictions and fail to capture the true variability seen in blood glucose levels, thereby reducing the effectiveness of the Metropolis algorithm in this context.

While both of the sets of chosen priors have proven effective for certain patients, yielding curves that accurately capture the intended dynamics, it is clear that these priors do not universally apply across all cases. For some patients, the priors produce posterior distributions that align well with the observed data, resulting in precise and realistic curve behaviors. However, for others, particularly where the decay or transition dynamics differ from the expected norm, the priors lead to deviations that reduce the model's accuracy. This variability highlights the need for a more tailored approach to prior selection, ensuring that the model remains flexible and adaptable to the specific characteristics of each patient's data. Going forward, it may be necessary to consider patient-specific priors or a more adaptive prior framework to achieve consistent and accurate modeling outcomes across diverse patient populations.

Finally, the high closeness between the posterior distributions and the priors may suggest strong priors but also overestimated ones. The data provided might not be sufficiently informative or varied, resulting in little change from the initial assumptions encoded in the

priors. The likelihood function may be not as informative as we expected since we only give the observed data a small variation from the obtained data. This also reflects on how overconfident the model is, with an average of 60% acceptance rate. Another limitation is also the lack of data as we only have 2-3 runs per patient. Additionally, the priors might be overly strong or narrowly defined, dominating the posterior distributions and preventing them from reflecting new insights from the data. Another possibility is that the model may not be complex enough to capture the intricacies of the data, leading to posteriors that remain close to the priors.

Further work could involve refining the acceptance criterion by adopting a more robust method, such as a random walk approach. This would entail taking steps of suitable sizes in the parameter space to improve the efficiency and accuracy of the exploration. Additionally, standardizing or resampling the data could help address the high variability observed between runs. Standardization techniques can reduce discrepancies in scale, while resampling methods, can create more stable subsets of data. These enhancements would facilitate more consistent modeling of the prior distributions and improve the overall reliability of the MCMC model.

## BIBLIOGRAPHY

- [1] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. CRC Press, 3rd edition, 2013.
- [2] W.K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [3] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.