Research & Reviews: Journal of of Statistics and Mathematical Sciences

Bayesian Hierarchical Functional Data Analysis for Healthy Menstrual Cycles

Calvin L Williams[1]* and Trevor Partridge[2]

Department of Mathematical Sciences, Clemson University, Clemson, SC 29634-1907, USA

Research Article

Received date: 04/09/2015

Accepted date: 09/11/2015

Published date: 25/11/2015

***For Correspondence**

Calvin L Williams, Department of Mathematical Sciences, Clemson University, Clemson, SC 29634-1907, USA

E-mail: calvinw@g.clemson.edu

Keywords: Hierarchical model, Gibbs sampling, Simulation, Bayesian specification, Menstrual cycles.

ABSTRACT

In almost every field of study, studying the distribution of a random curve can give needed information. In the field of *in vitro* fertilization, particular interest is given to modeling the temperature curves of women across their menstrual cycle. Fitting the curve can lead to categorizing cycles as healthy or unhealthy and can help identify the most fertile days of a cycle.

INTRODUCTION

In almost every field of study, studying the distribution of a random curve can give needed information. In the field of in vitro fertilization, particular interest is given to modeling the temperature curves of women across their menstrual cycle. Fitting the curve can lead to categorizing cycles as healthy or unhealthy and can help identify the most fertile days of a cycle. A healthy menstrual cycle consists of three phases, the follicular phase, ovulation, and the luteal phase. During the follicular phase, the basal body temperature (the lowest temperature reached during rest, referred to as bbt for the remainder of this paper) holds constant at a low plateau. During ovulation, the bbt increases until it hits the high plateau. The bbt then holds constant at this high plateau through the luteal phase. At the end of the cycle, the bbt drops back down to the low plateau before menstruation and the cycle repeats [1]. Figure 1 gives a few examples of a healthy menstrual cycle's bbt levels [2]. Section 2 outlines the general hierarchical model, then gives a parametric modeling of the bbt curves. The end of Section 2 completes the Bayesian specification and specifies the hyperparameters needed for the prior distributions. Section 3 briefly describes how the Gibbs sampling algorithm works and then applies it to model. Section 4 gives a short description of the change-point problem and defines the stopping rule needed in order to identify the two change-points in the model. Section 5 details a simulation of the data and then applies the Gibbs sampling algorithm with intentionally skewed hyperparameters. The results of the simulation are discussed at the end of Section 5. Finally, Section 6 provides a short discussion of the model and conclusions **(Figure 1)**.

THE HIERARCHICAL MODEL

First, we will define terms necessary to describing the model. A prior distribution of an unknown quantity X is the probability distribution that would express one's certainty about X before the data is taken into account. Similarly, a posterior distribution of an unknown quantity X is the probability distribution of that quantity conditional on the observed data. A hyperparameter is parameter of a prior distribution that is not treated as random. For example, consider a random variable X that came from A $N_2(\boldsymbol{\mu},\boldsymbol{\Sigma})$ distribution, where Σ is known and we can use a $N_2(\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0)$ distribution to model μ. Then μ is a parameter, $N_2(\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0)$ is a prior distribution, and μ_0 and Σ_0 are hyperparameters.

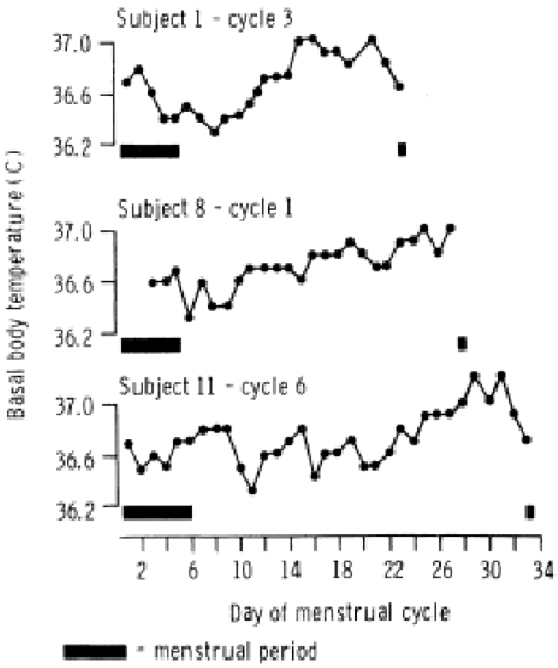

Figure 1. Examples of bbt levels for healthy cycles[2].

The General Model

Let $i \in \{1,2,\ldots,n_i\}$ refer to woman $i, j \in \{1,2,\ldots,n_i\}$ refer to cycle j of woman i, and $t \in \{1,2,\ldots,n_{ij}\}$ refer to day t in cycle j of woman i. These indices will retain these definitions throughout the remainder of the paper. We can establish the following model for bbt curves [3]:

$$y_{ij}(t) = \eta_{ij}(t) + \varepsilon_{ij}(t),$$

where ε_{ij} is the measurement error and η_{ij} is a smooth function from $\{1,2,\ldots,n_{ij}\}$ to □. Since cycles of the same woman are expected to be correlated, we can establish a prior distribution for the distribution of functions for woman $i\left(\{\eta_{ij}\}_{j=1}^{n_i}\right)$ call it G_i. However, healthy menstrual cycles will follow a similar pattern among women, so we can establish a prior distribution for the collection of distributions for the different women $\left(\mathcal{G} = \{G_i\}_{i=1}^{n}\right)$, call it P. Thus if we assume our error is normally distributed, we have the following model [3]:

$$y_{ij}(t) = \eta_{ij}(t) + \varepsilon_{ij}(t),$$
$$\varepsilon_{ij}(t) \sim N\left(0,\sigma^2\right), \quad \eta_{ij} \sim G_i, \quad \mathcal{G} = \{G_i\}_{i=1}^{n} \sim P.$$

Basis representation

Now that we have our general model, there are a number of different representations that could be used for η_{ij}. A common strategy is to consider a basis representation [3]:

$$\eta_{ij}(t) = \sum_{i=1}^{k} \theta_{ijh} b_h(t),$$

where $\{b_h(t)\}_{h=1}^{k}$ are pre-specified basis functions and $\boldsymbol{\theta}_{ij} = \left(\theta_{ij1},\ldots,\theta_{ijk}\right)'$ are cycle-specific basis coefficients. Since $\{b_h(t)\}_{h=1}^{k}$ is pre-specified, our major objective is to model the θ_{ij}s. The approach we use is to give the θ_{ij}s a hierarchical normal model [3]:

$$\boldsymbol{\theta}_{ij} \sim N_k\left(\boldsymbol{\alpha}_i, \Omega^{-1}\right), \quad \boldsymbol{\alpha}_i \sim N_k\left(\boldsymbol{\alpha}, \Omega_1^{-1}\right),$$

Where $N_k(\mu,\Sigma)$ is the multivariate normal distribution with mean vector μ and covariance matrix Σ. The woman-specific basis coefficients are given by α_i, α gives the global mean basis coefficients, and Ω and Ω_1 are the precision matrices for the within and between woman variability respectively. So to complete the model, we need to specify the basis functions $\{b_h(t)\}_{h=1}^{k}$ and establish the components of $_{ij}$. We'll do this in the next section when we consider a parametric hierarchical structure.

Parametric Hierarchical Model

As we discussed in Section 1, bbt levels of a healthy cycle follow a biphasic pattern where the bbt levels start at the low plateau in the follicular phase then quickly rise during ovulation to the high plateau in the luteal phase. Thus we can model η_{ij} (t) as follows [3]:

$$\eta_{ij} = \begin{cases} \theta_{ij1} & 1 \le t \le k_{ij} \\ \theta_{ij1} + \theta_{ij2}\left(\dfrac{t - k_{ij}}{r_{ij}}\right) & k_{ij} \le t \le k_{ij} + r_{ij} \\ \theta_{ij1} + \theta_{ij2} & k_{ij} + r_{ij} \le t \le n_{ij} \end{cases}$$

where $\boldsymbol{\theta}_{ij} = \left(\theta_{ij1}, \theta_{ij2}\right)'$ where θ_{ij1} and θ_{ij2} indicate the temperature during the follicular phase and the increase in bbt during ovulation respectively for cycle j of woman i. We can further define to be the last day of the follicular phase and r_{ij} to be the number of days the bbt rises during ovulation **(Figure 2).** Thus, if we consider y_{ij} (t) as Y_{ij} a vector of n_{ij} bbt values, we can represent our hierarchical model in matrix form [3]:

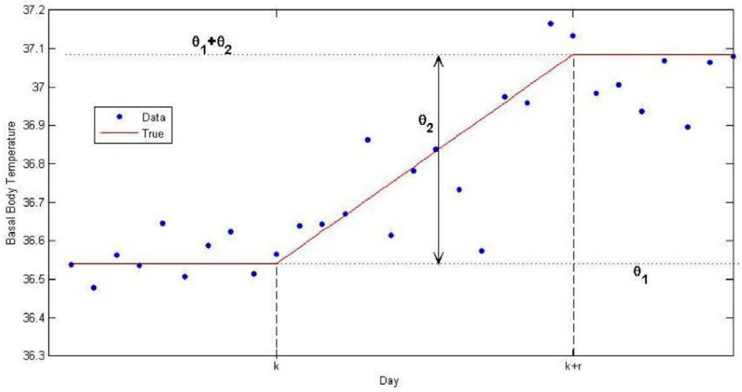

Figure 2. Graphical representation of parametric model.

$$\mathbf{Y}_{ij} = \mathbf{X}_{ij}\boldsymbol{\theta}_{ij} + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim N_{n_{ij}}\left(0, I_{n_{ij}}\sigma^2\right)$$

$$\mathbf{X}_{ij} = \left[\begin{array}{ccc|cccc|ccc} 1 & \cdots & 1 & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 0 & \cdots & 0 & \frac{1}{r} & \frac{2}{r} & \cdots & \frac{r}{r} & 1 & \cdots & 1 \end{array}\right]^T$$

Thus X_{ij} (t), the t[th] row of X, corresponds to t[th] day of cycle j of woman i. So we have defined our pre-specified basis function $\left(b_1(t) b_2(t)\right) = \mathbf{X}_{ij}(t)$ and our cycle-specific basis coefficients $\boldsymbol{\theta}_{ij} = \left(\theta_{ij1}, \theta_{ij2}\right)$.

Bayesian Specification

Now we have within (Ω) and between (Ω_1) woman variability, the last day of the follicular phase (k_{ij}), the number of days during ovulation (r_{ij}), the global mean (α), and measurement error variance (σ^2) left to model. We'll use the following equations to model the within and between woman variability:

$$\Omega = \begin{bmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{bmatrix} \qquad \Omega_1 = \begin{bmatrix} \omega_{11} & 0 \\ 0 & \omega_{12} \end{bmatrix}$$

We complete the Bayesian specification with the following priors [3]:

$$k_{ij} \sim \mathcal{U}\left(m_{ij}, m_{ij} + 20\right)$$

$$r_{ij} \sim \mathcal{U}(1, 15)$$

$$\boldsymbol{\alpha} \sim N_k\left(\boldsymbol{\alpha}_0, \boldsymbol{\Sigma}_\alpha\right)$$

$$\sigma^{-2} \sim \mathcal{G}(c, d)$$

$$\omega_h \sim \mathcal{G}\left(a_h, b_h\right) for\ h = 1,2$$

$$\omega_{1h} \sim \mathcal{G}\left(a_{1h}, b_{1h}\right) for\ h = 1,2$$

where $\mathcal{G}(a,b)$ is the gamma distribution with shape parameter a and inverse scale parameter b, $\mathcal{U}(a,b)$ is the discrete uniform distribution between a and b, and m_{ij} is the first day after menstruation. Thus, our pre-specified hyperparameters are α_0, Σ_α, c, d, a_h, b_h, a_{h1}, and b_{h2} (for h = 1, 2). A summary of the parameters, hyperparameters, their descriptions, and distributions is given in the appendix.

GIBBS SAMPLING

Our goal is to find the parameters of our model given a set of observations. Since finding the joint distribution is a complicated process, we need to use a different method. A common method to do this is to use a Monte Carlo Markov Chain as an iterative procedure. The specific type of Monte Carlo Markov Chain that we'll use is called a Gibbs sampler.

Definition

Suppose we have random variables $X_1, X_2, \ldots, X_n$.

1. Initialization: Let $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}$ be given some initial value.

2. Iteration: For $1 \le I \le K$, sample $x_j^{(i)}$ from the conditional distribution of X_j given $X_1 = x_1^{(i)}, \ldots, X_{j-1} = x_{j-1}^{(i)}$ and $X_{j+1} = x_{j+1}^{(i-1)}, \ldots, n = x_n^{(i-1)}$ for $1 \le j \le n$

This Gibbs sampling procedure is a Markov chain, where the stationary distribution is joint distribution of $X_1, X_2, \ldots X_n$. Thus, for large k, the sample $x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}$ approximates the joint distribution of $X_1, X_2, \ldots, X_n$. A blocked Gibbs sampler uses the same steps, except it groups some of the random variables together so that the variables in a block are sampled from their joint distribution conditioned on all other variabes.

Gibbs Sampling Algorithm

Here is the blocked Gibbs sampling algorithm for our model after we've given the parameters some initial value [2]:

1. Cycle-specific coefficients: $\boldsymbol{\theta}_{ij} \mid \mathbf{y}_{ij}, \ldots, \sim N_2(b_1, V_1)$

$$V_1 = \left(\Omega + \sigma^{-2}\mathbf{X}_{ij}'\mathbf{X}_{ij}\right)^{-1} \qquad b_1 = V_1\left(\Omega + \sigma^{-2}\mathbf{X}_{ij}'\mathbf{y}_{ij}\right)$$

2. Women-specific means: $\boldsymbol{\alpha}_i \mid \mathbf{y}_{ij}, \ldots, \sim N_2(b_2, V_2)$

$$V_2 = \left(\Omega_1 + n_i\Omega\right)^{-1} \qquad b_2 = V_2\left(\Omega_1 + \Omega\sum_{j=1}^{n_i}\boldsymbol{\theta}_{ij}\right)$$

3. Global mean: $\boldsymbol{\alpha}_i \mid \mathbf{y}_{ij}, \ldots, \sim N_2(b_3, V_3)$

$$V_3 = \left(\Sigma_\alpha^{-1} + n\Omega_1\right)^{-1} \qquad b_3 = V_3\left(\Sigma_\alpha^{-1}\boldsymbol{\alpha}_0 + \Omega_1\sum_{i=1}^{n}\boldsymbol{\alpha}_i\right)$$

4. Components of Ω (h = 1,2): $\omega_h \mid \mathbf{y}_{ij}, \ldots, \sim \mathcal{G}(p_1, q_1)$

$$p_1 = a_h + \frac{1}{2}\sum_{i=1}^{n} n_i \qquad q_1 = b_h + \frac{1}{2}\sum_{i=1}^{n} n_i \sum_{i=j}^{n_i}\left(\theta_{ijh} - \alpha_{ih}\right)^2$$

5. Components of Ω_1 (h = 1,2): $\omega_{1h} \mid \mathbf{y}_{ij}, \ldots, \sim \mathcal{G}(p_2, q_2)$

$$p_2 = a_{1h} + \frac{n}{2} \qquad q_2 = b_{1h} + \frac{1}{2}\sum_{i=1}^{n}\left(\alpha_{ih} - \alpha_h\right)^2$$

6. Error variance: $\sigma^{-2} \mid \mathbf{y}_{ij}, \ldots, \sim \mathcal{G}(p_3, q_3)$

$$p_3 = c + \frac{n}{2}\sum_{i=1}^{n}\sum_{j=1}^{n_i} n_{ij} \qquad q_3 = d + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n_i}\sum_{t=1}^{n_{ij}}\left(y_{ij}(t) -_{ij}(t)\boldsymbol{\theta}_{ij}\right)^2$$

7. Last day of follicular phase: Use the change-point stopping rule for k_{ij} (see Section 4).

8. Ovulation period: Use the change-point stopping rule for r_{ij} (see Section 4).

So if we use the blocked Gibbs sampling steps to update the unknown parameters from their conditional distributions, then after a large number of iterations, we'll be sampling from the joint distribution of all the unknown parameters.

CHANGE-POINT PROBLEM

Now we need to consider k_{ij} and r_{ij} since these are specific points (called change-points) in the bbt curve where the shape of the curve changes. The process for finding a change-point P is intuitive. Given a set of observations $\mathbf{y}_m = (y_1, y_2, \ldots, y_m)$ we'll find the probability that $p \le m$ given y_m. We want this probability to be higher than some high threshold Q. Thus we have a "stopping rule" [4]:

Terminate at m if and only if $Pr(P \le m \mid \mathbf{y}_m) > Q$.

Starting with m = 1, the first time that the probability is greater than Q is our change-point P.

Finding k_{ij}

First we wish to find k_{ij} assuming that the other parameters are known. Let X_k be the first m rows of the matrix X_{ij} established above but with $k_{ij=k}$. Thus $\mathbf{y}_{ij,m} = \mathbf{X}_k \boldsymbol{\theta}_{ij} + \varepsilon_{ij,m}$. Let $\{\pi_k\}$ be the prior distribution for k_{ij},$\pi(.)$ be the density function of θ_{ij}, and f(.) be the density function of the data $y_{ij,m.}$ We now need to find the posterior distribution of k_{ij} in order to find $Pr(k \le m | \mathbf{y}_m)$ for our stopping rule. Let $L_k = f\left(_{ij,m} | k_{ij} = k\right)$ be the change-point likelihood. Then implementing Bayes' theorem, we get [4]

$$L_k = \frac{f\left(\mathbf{y}_{ij,m} | \boldsymbol{\theta}_{ij} = k\right)\pi\left(\boldsymbol{\theta}_{ij}\right)}{\pi\left(\boldsymbol{\theta}_{ij} | \mathbf{y}_{ij,m}k\right)} d$$

Since L_k is not a function θ_{ij}, we may replace θ_{ij} with 0. Also $f\left(\mathbf{y}_{ij,m} | 0, k\right)$ does not depend on k, so L_k is proportional in k to

$$W_k = \left(\pi\left(0 | \mathbf{y}_{ij,m}, k\right)\right)^{-1}$$

The posterior distribution of θ_{ij} is $N_2(\mu_k,V_k)$ where

$$\mathbf{V}_k = \left(\sigma^{-2}\mathbf{X}_k'\mathbf{X}_k + \Omega\right)^{-1}$$

$$\boldsymbol{\mu}_k = \mathbf{V}_k\left(\sigma^{-2}\mathbf{X}_k'\mathbf{y}_{ij,m} + \Omega\boldsymbol{\alpha}_i\right)$$

Note that this is the same posterior distribution found for θ_{ij} found in Section 3, but with the altered matrix X_k. Using this posterior distribution for θ_{ij}, we find that

$$W_k = \left(det_k\right)^{-\frac{1}{2}} exp\left(\frac{1}{2}\boldsymbol{\mu}_k'\mathbf{V}_k\boldsymbol{\mu}_k\right).$$

Since we need to find $Pr\left(k \le m | \mathbf{y}_{ij,m}\right)$, we'll first find $Pr\left(k_{ij} = k | \mathbf{y}_{ij,m}\right)$. From applying Bayes' theorem to $Pr\left(k_{ij} = k | \mathbf{y}_{ij,m}\right)$, we get

$$Pr\left(k_{ij} = k | \mathbf{y}_{ij,m}\right) = L_k\pi_k\left(\sum_{s=1}^{n_{ij}} L_s\pi_s\right).$$

Since L_k is proportional in k to W_k, then we can substitute W_k in for L_k , resulting in

$$Pr\left(k_{ij} = k | \mathbf{y}_{ij,m}\right) = W_k\pi_k\left(\sum_{s=1}^{n_{ij}} W_s\pi_s\right)^{-1}.$$

These probabilities are easy to compute as the data becomes available. When $k \ge m$,

$$\mathbf{X}_k = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix}^T$$

So X_k does not depend on k (and hence $W_k = W_m$). Thus, our stopping rule is to terminate at m if and only if

$$(1 = Q)\sum_{k=1}^{m} W_k\pi_k > QW_m \sum_{k=m+1}^{n_{ij}} \pi_k$$

Once we hit our stopping rule [4], let k_{ij} =m.

Finding r_{ij}

We can apply the same approach to finding r_{ij} by using a transformation to examine the data "backwards", i.e. consider the last day first, and proceed through the cycle in reverse order. So our new reversed data is

$$\mathbf{y}_{ij}^* = \left(y_{ij}\left(n_{ij}\right), y_{ij}\left(n_{ij}-1\right), \ldots, y_{ij}(1)\right).$$

Let $= \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ be our transformation matrix. Then

$$\boldsymbol{\theta}_{ij}^* = \mathbf{A}\boldsymbol{\theta}_{ij} = \begin{pmatrix} \theta_{ij1} + \theta_{ij2} \\ -\theta_{ij2} \end{pmatrix}$$

$$y_{ij}^*(t) = \begin{cases} \theta_{ij1} + \theta_{ij2} & 1 \le t \le n_{ij} - r \\ \theta_{ij1} + \theta_{ij2} - \theta_{ij2}\left(\dfrac{t - n_{ij} + r}{r - k_{ij}}\right) & n_{ij} - r \le t \le n_{ij} - k_{ij} \\ \theta_{ij1} & n_{ij} - k_{ij} \le t \le n_{ij} \end{cases}$$

Thus if we let X_r* correspond to the first m rows of the above system, then

$$y^*_{ij,m} = \mathbf{X}^*_r \boldsymbol{\theta}^*_{ij} + \varepsilon^*_{ij,m}$$

This transformation lets us use the same process that we used to find k_{ij}. Let $\{\pi_r\}$ be the prior distribution for r (the second change-point). The new posterior distribution for θ^*_{ij} is $N_2(\mu_r, V_r)$, where

$$\mathbf{V}_r = \left(\sigma^{-2}\left(\mathbf{X}^*_r\right)' \mathbf{X}^*_r + \mathbf{A}\Omega\mathbf{A}'\right)^{-1}$$

$$\boldsymbol{\mu}_r = \mathbf{V}_r\left(\sigma^{-2}\left(\mathbf{X}^*_r\right)' \mathbf{y}^*_{ij,m} + \mathbf{A}\Omega\mathbf{A}'\mathbf{A}\boldsymbol{\alpha}_i\right).$$

Note that since $\boldsymbol{\theta}^*_{ij} = \mathbf{A}\boldsymbol{\theta}_{ij}, \boldsymbol{\theta}_{ij} \sim N_2\left(\mathbf{A}\boldsymbol{\alpha}_i, \mathbf{A}\Omega\mathbf{A}^T\right)$ thus giving a similar posterior distribution found in Section 3 for θ_{ij} . Using this posterior distribution for $*_{ij}$ we find that

$$W_r = \left(det\,\mathbf{V}_r\right)^{-\frac{1}{2}} exp\left(\frac{1}{2}\boldsymbol{\mu}'_r \mathbf{V}_r \boldsymbol{\mu}_r\right).$$

Thus, our stopping rule is to terminate at m if and only if

$$(1 = Q)\sum_{r=1}^{m} W_r \pi_r > QW_m \sum_{r=m+1}^{n_{ij}} \pi_r$$

Once we hit our stopping rule, $r = n_{ij} - m$ (since we transformed our data). Since r_{ij} is the number of days between k_{ij} and r, then $r_{ij} = r - k_{ij}$

SIMULATION

Data simulation

Now that the method has been described, we'll introduce a simulation in order to illustrate the method. The simulated data consisted of 20 women, each with 10 cycles, so we had a total of 200 cycles. Each cycle had a length of 30 days [5]. We arbitrarily assigned the global mean and covariance matrices to be

$$\boldsymbol{\alpha} = \left(36.5, 0.4\right)' \qquad \Omega^{-1} = \Omega_1^{-1} = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}.$$

Each woman specific mean was generated as per the model discussed in Section 2, so $\boldsymbol{\alpha}_i \sim N_2\left(\boldsymbol{\alpha}, \Omega_1\right)$. We then generated k_{ij} and r_{ij} for each cycle according to $k_{ij} \sim \mathcal{U}\left(1,14\right)$ and $r_{ij} \sim \mathcal{U}\left(1,15\right)$. Finally, we generated the cycle specific coefficients according to $\boldsymbol{\theta}_{ij} \sim N_2\left(\boldsymbol{\alpha}_i, \Omega\right)$ and the measurement error terms according to $\varepsilon_{ij} \sim N_{30}\left(0, \sigma^2\mathbf{I}\right)$ where σ^2 = 0.01. Now that we generated all the parameters needed for our model, we generated the bbt values according to

$$\mathbf{y}_{ij} = \mathbf{X}_{ij}\boldsymbol{\theta}_{ij} + \varepsilon_{ij}$$

Initialization

In order to run the algorithm, we need to specify the hyperparameters (our initial guesses). In order to account for error, we intentionally perturb the hyperparameters away from the actual parameters established in Section Data simulation So we set our hyperparameters to be

$$\boldsymbol{\alpha}_0 = \begin{bmatrix} 36 \\ 1 \end{bmatrix} \qquad \Sigma_\alpha = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.01 \end{bmatrix}$$

$$a_1 = a_2 = 1 \qquad b_1 = b_2 = 1$$

$$a_{11} = a_{12} = 1 \qquad b_{11} = b_{12} = 1$$

$$c = 1 \qquad d = 1$$

These hyperparamters are used in the Gibbs sampling steps outlined in Section 4. But before we can implement the Gibbs sampling algorithm, we need to initialize our unknowns from their prior distributions using the hyperparameters, as seen below:

$$\boldsymbol{\alpha} \sim N_2\left(\begin{bmatrix} 36 \\ 1 \end{bmatrix}, \begin{bmatrix} 0.25 & 0 \\ 0 & 0.01 \end{bmatrix}\right)$$

$$\omega_1, \omega_2 \sim \mathcal{G}\left(1,1\right)$$

$$\omega_{11}, \omega_{12} \sim \mathcal{G}\left(1,1\right)$$

$$\sigma^{-2} \sim \mathcal{G}\left(1,1\right)$$

We also need to initialize the α_is from their prior distributions since the first Gibbs sampling step is finding the posterior distribution of θ_{ij} conditional on α_i. So we'll generate α_i values from the prior distribution

$$\boldsymbol{\alpha}_i = N_2(\boldsymbol{\alpha},\Omega)$$

The final step for initialization was to set the k_{ij} and r_{ij} values all to a fixed value of 7.

Running the algorithm

Now that initialization is completed, we can run the algorithm. We ran the algorithm for a total of 6,000 iterations with a 1,000 iteration burn-in. To make sure that the parameters were converging, we performed traceplots of a number of the variables. **Figures 3 and 4** show the traceplots for α_{11} and α_{12} respectively.

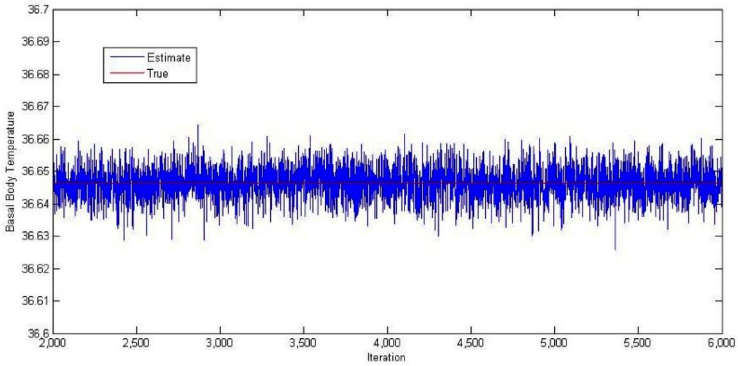

Figure 3. Trace plot of α11.

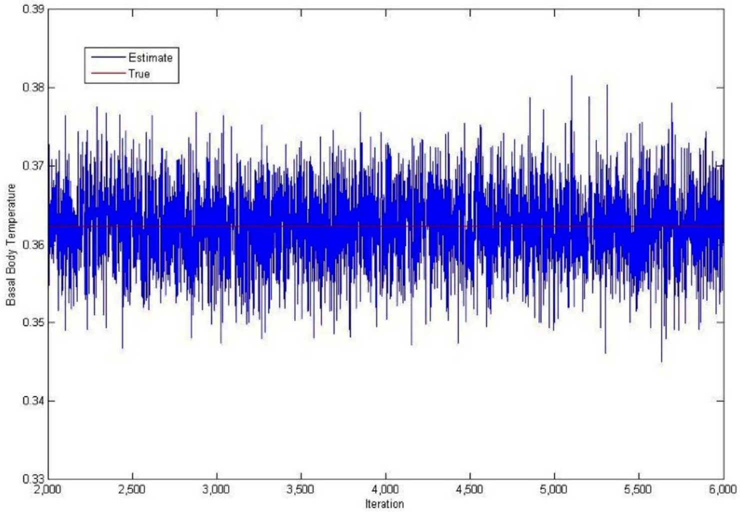

Figure 4. Trace plot of α12.

Tables 1 and 2 summarize the results from estimating the parameters in our model. The cycle specific parameters (k_{ij}, r_{ij}, θ_{ij1} and θ_{ij2}) are summaries over all cycles, while the women specific parameters (α_{i1} and α_{i2}) are summaries over all women. As evidenced by the tables, the distribution of these parameters was well estimated. Figures 5 and 6 show graphs of the true and estimated distributions for two example cycles.

Table 1. Posterior summaries of the true cycle and woman specific parameters.

True distribution					
Parameter	Mean	Standard Deviation	First Quartile	Median	Third Quartile
k_{ij}	7.07	4.00	4	7	11
r_{ij}	8.60	4.60	4	9	13
θ_{ij1}	36.49	0.13	36.40	36.50	36.58
θ_{ij2}	0.40	0.13	0.31	0.41	0.47
α_{i1}	36.49	0.08	36.43	36.51	36.55
α_{i1}	0.41	0.09	0.35	0.38	0.48

Table 2. Posterior summaries of the estimated cycle and woman specific parameters.

	Estimates				
Parameter	**Mean**	**Standard Deviation**	**First Quartile**	**Median**	**Third Quartile**
k_{ij}	7.91	4.04	4.5	8	12
r_{ij}	7.83	4.87	3	8	12.5
θ_{ij1}	36.49	0.13	36.39	36.50	36.57
θ_{ij2}	0.40	0.13	0.30	0.40	0.47
α_{i1}	36.50	0.08	36.44	36.50	36.54
α_{i2}	0.40	0.10	0.34	0.39	0.47

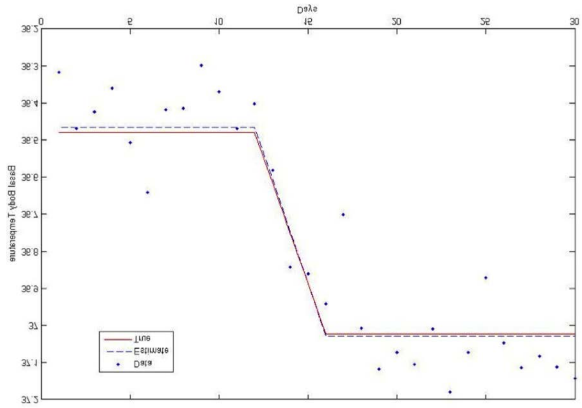

Figure 5. Woman 2, cycle 2 .

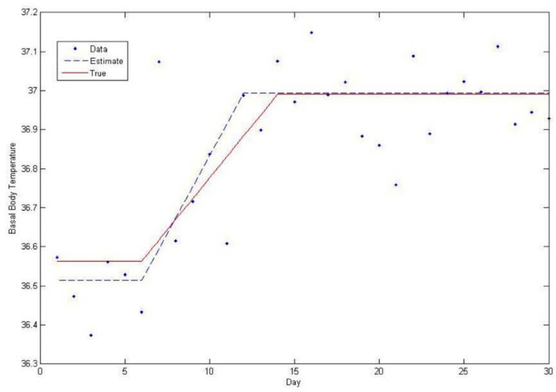

Figure 6. Woman 6, cycle 4.

CONCLUSION

So, we can see that the blocked Gibbs sampling algorithm was an accurate method for estimating the various parameters of our hierarchical model. Most importantly, the change-point analysis gave accurate estimates for the last day of the follicular stage and the number of days that the temperature rises during ovulation (Appendix). These two figures are critical for determining the most fertile period of the menstrual cycle. A natural extension of this topic would be to model unhealthy cycles using a nonparametric approach. Combining the parametric and nonparametric models would allow for broader analysis and more unusual data sets.

REFERENCES

1. McClure Browne JC. Postgraduate Obstetrics and Gynaecology 4th ed. 355. London: Butterworth. 1973.
2. Ovagraph Fertility Charting. www.ovagraph.com/chart-tags/typical-cycles.html 2015.
3. Scarpa B and Dunson DB. Bayesian hierarchical functional data analysis via contaminated information priors. Biometrics. 2009; 65: 772-780.
4. Carter RL and Blight BJN. A Bayesian change-point problem with an application to the prediction and detection of ovulation in women. Biometrics. 1981; 371: 743-751.
5. Royston JP and Abrams RM. An objective method for detecting the shift in basal body temperature in women. Biometrics. 1980; 36: 217-224