

Modeling a Mixture of Linear and Changepoint Trajectories for Longitudinal Time-Series Data

Shahedul A. Khan¹

¹ Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon, SK S7N 5E6, Canada.

Correspondence: Shahedul A. Khan, Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon, SK S7N 5E6, Canada. E-mail: s.khan@usask.ca

Received: March 6, 2015 Accepted: May 7, 2015 Online Published: July 1, 2015

doi:10.5539/ijsp.v4n3p1 URL: <http://dx.doi.org/10.5539/ijsp.v4n3p1>

Abstract

Longitudinal changepoint data naturally arise in many applications. Examples include transition of core body temperature following the hypothermia therapy and prostate-specific antigen levels following treatment. Note that the trend change occurs due to a shock (e.g., treatment) to the system. Thus, an individual exhibiting a linear trend could be an indication of insignificant effects of the shock. One of the goals of this type of study is to investigate whether the shock is significantly associated in changing the trend of a trajectory. The bent-cable model characterizes the shock-through data using three phases: (a) an incoming phase characterizing the trend before the shock comes into effect, (b) a transition due to the shock, and (c) an outgoing phase due to the after-shock effects. In this article, we develop bent-cable methodology accounting for trajectories exhibiting either a linear trend or a trend change characterized by gradual or abrupt transition.

Keywords: abrupt transition, Bayesian inference, bent-cable model, changepoint data, gradual transition, linear trend, longitudinal data, Markov Chain Monte Carlo

1. Introduction

Longitudinal changepoint trajectories with three phases – an incoming phase followed by a transition and an outgoing phase – are common in many applications including medical, biological and environmental sciences. Examples include the atmospheric concentrations of chlorofluorocarbons (CFCs) monitored from different stations all over the globe following the Montréal Protocol's ban on CFC products (Khan et al., 2009; 2012), transition of core body temperature following the hypothermia therapy to prevent brain damage of patients who have survived a cardiac arrest (Khan et al., 2013; Reynolds & Chiu, 2010), and prostate-specific antigen levels of the prostate cancer patients following treatment (Bellera et al., 2008). This type of data may be referred to as shock-through data, as the trend change occurs due to a shock (e.g., treatment) to the system. Note that an individual exhibiting a linear trend could be an indication of insignificant effects of the shock. Thus, among other related questions, one of the goals of this type of study is to investigate whether the shock has significant effects in effectively changing the trend of a trajectory. Therefore, proper characterization of the trajectories is of paramount importance to the researchers.

To illustrate the changepoint problem discussed above, we reintroduce here the example presented by Khan et al. (2013) to understand the effects of the hypothermia therapy on core body temperature (T_c). Hypothermia is a fatal condition which can occur when T_c falls below 35°C . By reducing a patient's T_c , systematic metabolism, and therefore tissue oxygen demand, can be lowered. This is believed to attenuate the effects of ischemia and the subsequent reperfusion at the cellular level (Gordon, 2001; Rincon & Mayer, 2006). Therefore, hypothermia is frequently used by the clinicians as a therapeutic tool for cardiac arrest, stroke and brain injury. In contrast, since T_c is typically associated with a critical threshold associated with a breakdown in the compensatory homeostatic mechanisms following severe hemorrhage (Rincon & Mayer, 2006; Connett et al., 1986), quantification of the transition of T_c to early hypothermia is of great clinical interest. With this objective in mind, an experiment was conducted on 38 rats to collect information about the state of hypothermia and resuscitation strategy (Reynolds & Chiu, 2010; Reynolds et al., 2007). The rats were hemorrhaged from the carotid catheter with a constraint of a mean arterial pressure threshold of 40 mm Hg. The experiment was continued until the target shed blood volume (60% of the total blood volume) was achieved. In this process, T_c was logged by automated remote data collection every 15

seconds for the duration of the trial. Figure 1 shows three T_c profiles. We see that some rats may exhibit a seemingly linear trend (e.g., Figure 1a). In addition to roughly linear incoming and outgoing phases at either end of a profile, we also see that some rats may exhibit an abrupt transition in T_c (e.g., Figure 1b), while others, a gradual transition (e.g., Figure 1c). That is, we have samples potentially coming from three different populations, say, L (linear), A (abrupt) and G (gradual), respectively, according to the type of transition for the underlying T_c trend.

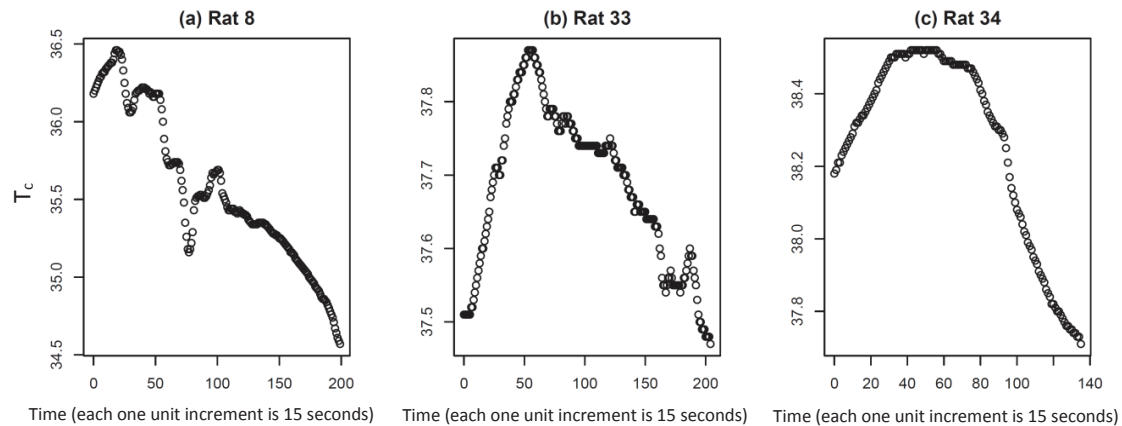


Figure 1. Core body temperature of three rat profiles out of 38, recorded during hemorrhage in a study to collect information about the state of hypothermia and resuscitation strategy.

The so-called broken-stick is a popular piece-wise linear model, and is a natural candidate to describe a continuous trend with an abrupt change (Hall et al., 2003; Lange et al., 1992; Kiuchi et al., 1995; Slate & Turnbull, 2000). Broken-stick models, however, may not be realistic in many applications where the longitudinal trajectories exhibit gradual changes (e.g., atmospheric concentration of CFCs (Khan et al., 2012)), or simply linear trends. So, it is desirable to relax the assumption of abruptness beforehand, and to formulate a model that is flexible enough to handle either type of changes, gradual or abrupt, as well as linear trends over time.

The challenge in longitudinal changepoint methodology lies not only in deriving a model that allows useful information about the population of interest, but also in developing method for statistical inference in the absence of standard regularity conditions (Chiu et al., 2006). In addition, existence of different types of transition (e.g., gradual and abrupt) across the individuals makes it an even more challenging problem to characterize the individual profiles. This is because it necessitates a flexible formulation of the model and inference technique to identify the type of transition exhibiting by a particular profile and to estimate the parameters associated with it.

Developed by Chiu et al. (2006) and Chiu & Lockhart (2010), bent-cable is an appealing changepoint model to characterize shock-through data with greatly interpretable regression coefficients that provide commonly sought information about a changepoint problem. It models a trajectory using three phases: a linear incoming phase characterizing the trend before the shock comes into effect, followed by a transition (gradual or abrupt) due to the shock, and a linear outgoing phase due to the after-shock effects (see Section 2 for details). Because of its appealing features in terms of flexibility, parsimoniousness and great interpretability, Khan et al. (2013) extended the bent-cable regression for longitudinal data by taking into account gradual and abrupt transitions as components of a mixture model; the outcome is the *flexible mixture bent-cable* methodology for samples that potentially come from two populations defined by the two types of transition: gradual and abrupt. They developed Bayesian inference implemented via *Markov chain Monte Carlo* (MCMC) techniques (Smith & Roberts, 1993). In analyzing the rat data, Khan et al. (2013) recognized that a few rats may exhibit neither an obviously gradual nor abrupt transition, but rather a linear decreasing trend. They reported that the rats exhibiting linear decreasing trends throughout were estimated to have arisen from Population A. Even if a few profiles exhibit linear decreasing trend in T_c , improper characterization of Population L by A could be misleading to the clinicians in taking appropriate resuscitation strategy. Therefore, In addition to the identification and characterization of the populations A and G, characterization of the profiles which exhibit simply linear trends would be valuable to the clinicians. Another assumption of Khan et al. (2013)'s longitudinal bent-cable methodology is that the populations share common intercept and slope parameters to describe the linear incoming and outgoing phases. Because of this assumption, the methodology is not expected to properly characterize the populations, especially in describing the linear phases. In fact, our preliminary simulation study reveals that serious estimation bias can occur under the assumption of sharing common

linear parameters when, in reality, those differ across the populations.

In this article, we show that the simple linear regression model is a special case of the bent-cable model. We then develop our longitudinal bent-cable methodology to simultaneously account for trajectories exhibiting either a linear trend or a trend change characterized by gradual or abrupt transition. In particular, we take into account

- (a) samples potentially coming from three different populations: Populations G, A and L;
- (b) population-specific linear parameters to allow proper inference at the population levels; and
- (c) hierarchical Wishart priors to model the between-individual covariance structure.

Extensions (a)-(c) to the Khan et al. (2013)'s model constitute our flexible mixture bent-cable model. Note that Khan et al. (2013) assumed Wishart priors to model the between-individual covariance structure. As pointed out by Daniels & Kass (1999), our simulation study (Section 4) reveals that Wishart priors can be sensitive to the choices of the hyperparameters, while implementing (c) with sensible choices of the hyperparameters provides robustness in estimating the between-individual covariance structure.

The rest of the paper is organized as follows. We describe the formulation of our longitudinal bent-cable model in Section 2. Technical details of the MCMC algorithm are described in Section 3. In Section 4, we present extensive simulation studies to demonstrate that (i) the proposed methodology performs well to characterize the three populations, (ii) the method based on population-specific linear parameters performs quite well even when the populations share common intercept and slope parameters in reality, and (iii) the Wishart priors can be sensitive to the choices of the hyperparameters, whereas the hierarchical Wishart priors provide robust estimates of the covariance matrices under sensible choices of the hyperparameters. We then illustrate our methodology with an application to the aforementioned rat data (Section 5). We conclude this article with a discussion in Section 6.

2. The Flexible Mixture Bent-Cable Model

Suppose we have m individuals. For the j^{th} individual ($j = 1, 2, \dots, m$), let there be n_j measurements, and let t_{jk} denote the k^{th} measurement occasion, $k = 1, 2, \dots, n_j$. Without loss of generality, we assume that all individuals enter the study at time $t_{j1} = 0$. We model the corresponding response at time t_{jk} , denoted by y_{jk} , by the relationship

$$y_{jk} = f(t_{jk}, \boldsymbol{\theta}_j) + \epsilon_{jk} \quad (1)$$

where $\boldsymbol{\theta}_j$ is the vector of regression coefficients for the j th individual, $f(\cdot)$ is a function of t_{jk} and $\boldsymbol{\theta}_j$ which characterizes the trend in the response over time, and ϵ_{jk} is the random error component. The formulation of the model is completed using the bent-cable function (Chiu et al., 2006; Chiu & Lockhart, 2010), defined by

$$f(t_{jk}, \boldsymbol{\theta}_j) = \beta_{0j} + \beta_{1j}t_{jk} + \beta_{2j}q(t_{jk}, \boldsymbol{\alpha}_j) \quad (2)$$

where $\boldsymbol{\beta}_j = (\beta_{0j}, \beta_{1j}, \beta_{2j})'$, $\boldsymbol{\alpha}_j = (\gamma_j, \tau_j)'$ and $\boldsymbol{\theta}_j = (\boldsymbol{\beta}'_j, \boldsymbol{\alpha}'_j)'$. Here, $\boldsymbol{\beta}_j$ is the vector of the linear parameters with β_{0j} being the y-intercept in the incoming phase, and β_{1j} and $\beta_{1j} + \beta_{2j}$, the slopes in the incoming and outgoing phases, respectively. The function $q(\cdot)$ characterizes the transition via two parameters τ_j and γ_j , and can be expressed as

$$q(t_{jk}, \boldsymbol{\alpha}_j) = \begin{cases} \frac{(t_{jk} - \tau_j + \gamma_j)^2}{4\gamma_j} \mathbf{1}\{|t_{jk} - \tau_j| < \gamma_j\} + (t_{jk} - \tau_j) \mathbf{1}\{t_{jk} - \tau_j \geq \gamma_j\} & \text{for Population G} \\ (t_{jk} - \tau_j) \mathbf{1}\{t_{jk} - \tau_j \geq 0\} & \text{for Population A} \\ t_{jk} & \text{for Population L} \end{cases} \quad (3)$$

where $\mathbf{1}(\cdot)$ is an indicator function. Note that both τ_j and γ_j are positive for gradual transition, representing the center and the half-width of the bend, respectively; $\gamma_j = 0$ reduces the bent-cable function to a piece-wise linear model for which $q(t_{jk}, \boldsymbol{\alpha}_j) = (t_{jk} - \tau_j) \mathbf{1}\{t_{jk} - \tau_j \geq 0\}$; and $\tau_j = \gamma_j = 0$ reduces the bent-cable function to a simple linear regression model for which $q(t_{jk}, \boldsymbol{\alpha}_j) = t_{jk}$ and $f(t_{jk}, \boldsymbol{\theta}_j) = \beta_{0j} + (\beta_{1j} + \beta_{2j})t_{jk}$ (see Figure 2). Chiu & Lockhart (2010) also defined the critical time point (CTP) at which the slope of the bent-cable function changes signs: the CTP is given by $\tau_j - \gamma_j - 2\beta_{1j}\gamma_j/\beta_{2j}$ for gradual transition (Figure 2a), and τ_j for abrupt transition (Figure 2c). For gradual transition, the CTP does not represent anything meaningful when the slope of the bent-cable function does not change signs, and therefore is not marked in Figure 2b. For abrupt transition, τ_j is the time at which the slope changes signs (Figure 2c) or the slope changes magnitudes but not signs (Figure 2d).

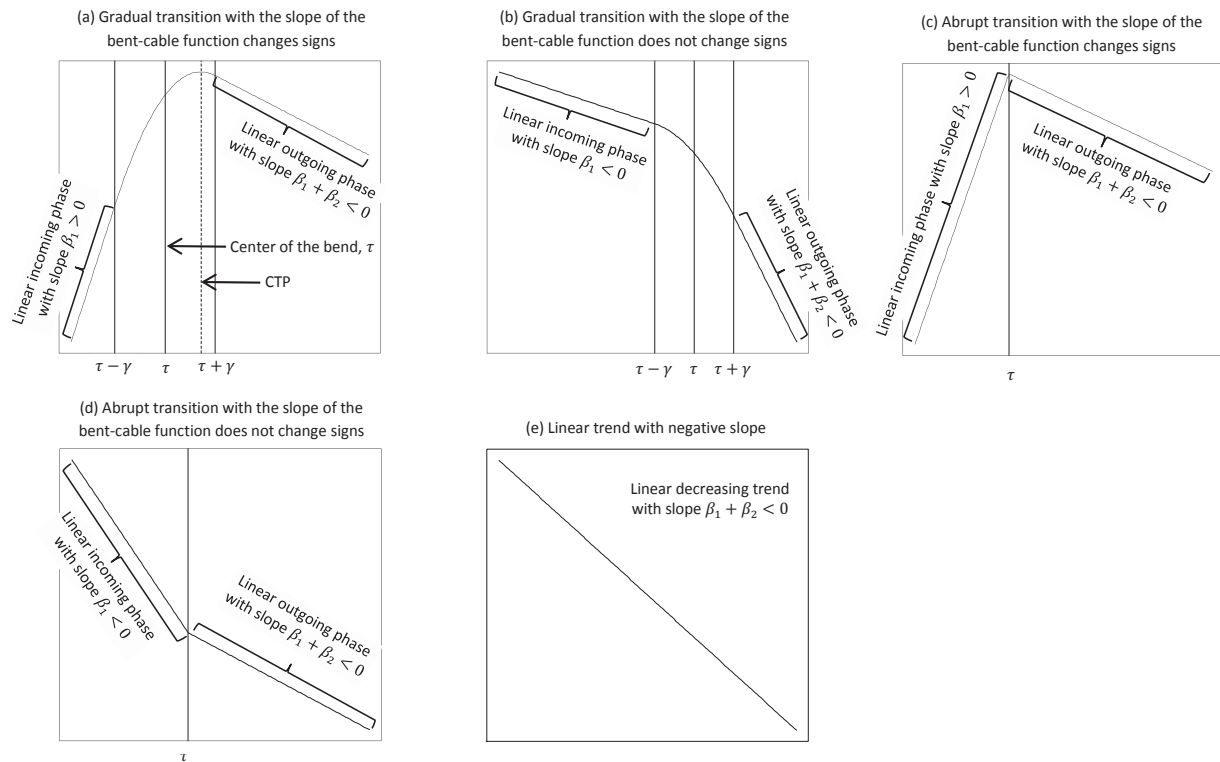


Figure 2. The bent-cable function. (a) A gradual transition for which $\tau, \gamma > 0$; the two linear segments (incoming and outgoing) are joined by a quadratic bend, ranging from $\tau - \gamma$ to $\tau + \gamma$; the slope changes signs at the critical time point (CTP), and is marked by the dotted line. (b) A gradual transition for which the slope of the bent cable does not change signs. (c) An abrupt transition with $\tau > 0$ and $\gamma = 0$ yields a piecewise linear trend; the transition takes place at time τ , which is the CTP for an abrupt transition. (d) An abrupt transition for which the slope of the bent cable does not change signs. (e) A linear decreasing trend for which $\tau_j = \gamma_j = 0$; the slope of the equation is given by $\beta_{1j} + \beta_{2j}$.

Bayesian modeling of longitudinal data involves three levels of hierarchy (Davidian & Giltinan, 1995; Wakefield et al., 1994): (a) Level 1 that characterizes the longitudinal trajectories using the regression equations (1)-(3), (b) Level 2 that characterizes the between-individual variability via the models assigned to θ_j 's, and (c) Level 3 that quantifies prior information about the random quantities of Levels 1 and 2. Note that the random error components, ϵ_{jk} 's in (1) account for within-individual variability in the measurements around the regression $f(t_{jk}, \theta_j)$. We assume at Level 1 that $\epsilon_{jk} \sim N(0, \sigma_j^2)$. Under this model, the correlation of the repeated measurements from the same individual is induced by the individual-specific random coefficients θ_j 's, and is called the conditionally independent specification of the within-individual variation (Laird & Ware, 1982). At Level 2, we assume that an individual potentially comes from one of three populations: G, A and L. Let each individual has probability p_i to have come from Population i , where $i = 1, 2, 3$ refer to G, A, and L, respectively. Under this setting, the individuals form a sample from a mixture model. We can express the mixture model under a hierarchical structure, using the unobserved data

$$z_{ij} = \begin{cases} 1, & \text{if the } j^{\text{th}} \text{ individual belongs to the } i^{\text{th}} \text{ component of the mixture} \\ 0, & \text{otherwise} \end{cases}$$

where $P(z_{ij} = 1) = p_i$. Note that the artificial random variables z_{ij} 's are introduced to facilitate implementing the Gibbs sampling for Bayesian inference (Robert, 1996).

Based on our choices of distributions for the relevant quantities at Levels 1-3, we write the hierarchical model as follows:

$$[y_j | \theta_j, \sigma_j^2] = N_{n_j}(\mathbf{f}(t_j, \theta_j), \sigma_j^2 \mathbb{I}_j) \quad \forall j = 1, 2, \dots, m, \tag{4}$$

$$\left. \begin{aligned}
[\alpha_j | \mathbf{z}_j, \xi, \eta, \Psi, \sigma_\eta^2] &= [LN_2(\alpha_j | \xi, \Psi)]^{\tau_j} [\mathbf{1}(\gamma_j = 0) LN(\tau_j | \eta, \sigma_\eta^2)]^{z_j} [\delta(\alpha_j)]^{z_{3j}} \\
[\beta_j | \mathbf{z}_j, \mu_1, \mu_2, \mu_3, \Sigma] &= \prod_{i=1}^3 [N_3(\beta_j | \mu_i, \Sigma)]^{z_{ij}} \\
[\mathbf{z}_j | p_1, p_2, p_3] &= \text{Multinomial}(\mathbf{z}_j | 1, p_1, p_2, p_3)
\end{aligned} \right\}, \quad (5)$$

$$\left. \begin{aligned}
\mu_i &\sim N_3(\mathbf{h}_i, \mathbb{H}_i) \text{ for } i = 1, 2, 3, \quad \xi \sim N_2(\mathbf{h}_\xi, \mathbb{H}_\xi), \quad \eta \sim N(a_0, a_1), \\
\Sigma^{-1} &\sim W_3(\nu_1, (\nu_1 \mathbb{A}_1)^{-1}), \quad (\mathbb{A}_1)_{ll} \sim U(b_{0l}, b_{1l}) \text{ for } l = 1, 2, 3, \\
\Psi^{-1} &\sim W_2(\nu_2, (\nu_2 \mathbb{A}_2)^{-1}), \quad (\mathbb{A}_2)_{ll} \sim U(c_{0l}, c_{1l}) \text{ for } l = 1, 2, \\
\sigma_j^{-2} &\sim G\left(\frac{d_0}{2}, \frac{d_1}{2}\right), \quad \sigma_\eta^{-2} \sim G\left(\frac{e_0}{2}, \frac{e_1}{2}\right), \quad (p_1, p_2, p_3) \sim \text{Dirichlet}(\omega_1, \omega_2, \omega_3)
\end{aligned} \right\}, \quad (6)$$

where $\mathbf{y}_j = (y_{j1}, y_{j2}, \dots, y_{jn_j})'$, $\mathbf{t}_j = (t_{j1}, t_{j2}, \dots, t_{jn_j})'$, $\mathbf{f}(\mathbf{t}_j, \boldsymbol{\theta}_j) = (f(t_{j1}, \boldsymbol{\theta}_j), f(t_{j2}, \boldsymbol{\theta}_j), \dots, f(t_{jn_j}, \boldsymbol{\theta}_j))'$ and $\mathbf{z}_j = (z_{1j}, z_{2j}, z_{3j})'$; $\mu_i = (\mu_{i0}, \mu_{i1}, \mu_{i2})'$ and Σ are, respectively, the i^{th} population mean and the common covariance matrix (shared by all three populations) of β_j ; $\xi = (\xi_1, \xi_2)'$ and Ψ are the mean and covariance of $\log(\alpha_j)$ (for Population 1); η and σ_η^2 are the mean and variance of $\log(\tau_j)$ (for population 2); N_r , LN_r , W_r , G and U stand for r -variate normal, r -variate lognormal, r -dimensional Wishart, gamma and uniform distributions, respectively; and δ stands for the Dirac delta measure (Previato, 2002), defined as

$$\delta(\alpha_j) = \begin{cases} 1, & \text{if } \gamma_j = \tau_j = 0 \\ 0, & \text{otherwise} \end{cases}$$

In Equation (6), the vectors \mathbf{h} are Gaussian means and matrices \mathbb{H} are Gaussian covariances; $\frac{d_0}{2}$ and $\frac{d_1}{2}$ are, respectively, the shape and the rate parameters of the gamma distribution $G(\frac{d_0}{2}, \frac{d_1}{2})$; and ν and \mathbb{A}^{-1} are, respectively, the degrees of freedom and the mean of the Wishart distribution $W_p(\nu, (\nu \mathbb{A})^{-1})$. Note that we prefer the Wishart distribution in the formulation of our model, because it is the conjugate prior of the inverse of a Gaussian covariance matrix. We must have $\nu \geq r$ for a proper Wishart distribution, and setting $\nu = r$ is common in Bayesian inference which makes the Wishart prior nearly flat (Wakefield et al., 1994). However, Daniels & Kass (1999) pointed out that the specification of \mathbb{A} can be quite influential; since \mathbb{A}^{-1} is the mean of the Wishart distribution, \mathbb{A} should be a close approximation of the covariance matrix. Our simulation study also reveals that the estimation of the covariance matrix is highly sensitive to the choices of \mathbb{A} . In order to overcome this problem, we propose a hierarchical Wishart prior in (6), where \mathbb{A} is a diagonal matrix with each of the diagonal elements modeled using a uniform distribution.

In our formulation of the model, Levels 1 and 2 are (4) and (5), and Level 3 is (6) with all the hyperparameters at Level 3 assumed known. Note that μ_1 and μ_2 characterize the linear parameters for Populations 1 and 2, respectively. For example, μ_{11} and $\mu_{11} + \mu_{12}$ are, respectively, the slopes of the incoming and outgoing phases for Population 1. Our choices of distributions also lead to the fact that $(\beta_{0j}, \beta_{1j} + \beta_{2j})'$ for Population 3 has a normal distribution with mean $(\mu_{30}, \mu_{31} + \mu_{32})'$; therefore, we use μ_{30} and $\mu_{31} + \mu_{32}$ to characterize the linear trend for Population 3. Since our assumption for the α_j 's involves lognormal distributions, we can use Level 2 medians, namely $\exp\{\xi_1\}$ and $\exp\{\xi_2\}$ for Population 1 and $\exp\{\eta\}$ for Population 2, to describe the transition locations (Khan et al., 2013).

3. Bayesian Inference

Bayesian inference is made based on the posterior distribution of a parameter; the posterior mean or median can be considered a point estimate of the parameter, whereas the posterior standard deviation measures the uncertainty of the parameter a posteriori. The Bayesian analogue of a frequentist confidence interval is usually referred to as a credible interval. A $100(1 - 2a)\%$ credible interval is (r_1, r_2) , where r_1 and r_2 are the a^{th} and $(1 - a)^{\text{th}}$ quantiles of the posterior density, respectively.

Statistical inference is carried out by MCMC, where we sample from the posterior distribution by the Metropolis within Gibbs algorithm (Smith & Roberts, 1993). We denote all the model parameters collectively by Θ , with $\Theta^{(0)}$ being an arbitrary starting point. We generate an instance from the full conditional distributions of each of the components of the model parameters conditional on the current values of the remaining components. This completes a transition from $\Theta^{(0)}$ to $\Theta^{(1)}$. Iterating through this cycle of generating data from the full conditional distributions gives a Markov chain $\{\Theta^{(s)}, s = 1, 2, \dots, T\}$ (Davidian & Giltinan, 1995). Monte Carlo integration

(Lemieux, 2009) is then applied to the Markov chain for Bayesian inference. For example, the marginal posterior means of the intercept and the slope parameters for Population 3 can be approximated by

$$\hat{\mu}_{30} = \sum_{s=S}^T \frac{\mu_{30}^{(s)}}{T-S+1} \quad \text{and} \quad \widehat{\mu_{31} + \mu_{32}} = \sum_{s=S}^T \frac{\mu_{31}^{(s)} + \mu_{32}^{(s)}}{T-S+1},$$

where the initial $S - 1$ iterations are considered as burn-in samples.

The implementation of the Metropolis within Gibbs algorithm for our flexible mixture bent-cable model is described as follows. Let $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2, \dots, \boldsymbol{\beta}'_m)'$, $\boldsymbol{\alpha} = (\alpha'_1, \alpha'_2, \dots, \alpha'_m)'$, $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_m)'$, $\boldsymbol{\mu} = (\mu'_1, \mu'_2, \mu'_3)'$ and $\mathbf{z} = (\mathbf{z}'_1, \mathbf{z}'_2, \dots, \mathbf{z}'_m)'$. Suppose that $\{\Theta^{(0)}, \mathbf{z}^{(0)}\}$ is an arbitrary starting point, and $\{\Theta^{(s)}, \mathbf{z}^{(s)}\}$ is its update at iteration s . Given $\{\Theta^{(0)}, \mathbf{z}^{(0)}\}$, our choice of order to update the components leads to the following sequence of drawing samples from the full conditional distributions $\pi(\cdot)$ (the full conditional distributions are presented in Appendix).

1) For $j = 1$,

- (a) $(\alpha_j^{(s+1)}, \mathbf{z}_j^{(s+1)}) \sim \pi(\alpha_j, \mathbf{z}_j | \Theta^{(s)}, \mathbf{z}_j^{(s)}, \mathbf{y}_j)$, where $\pi(\alpha_j, \mathbf{z}_j | \cdot)$ can be expressed up to a proportionality constant;
- (b) $\boldsymbol{\beta}_j^{(s+1)} \sim \pi(\boldsymbol{\beta}_j | \alpha_j^{(s+1)}, \mathbf{z}_j^{(s+1)}, \sigma_j^{(s)}, \Sigma^{(s)}, \boldsymbol{\mu}^{(s)}, \mathbf{y}_j)$, where $\pi(\boldsymbol{\beta}_j | \cdot)$ is normal;
- (c) $\sigma_j^{-2(s+1)} \sim \pi(\sigma_j^{-2} | \boldsymbol{\beta}_j^{(s+1)}, \alpha_j^{(s+1)}, \mathbf{z}_j^{(s+1)}, \mathbf{y}_j)$, where $\pi(\sigma_j^{-2} | \cdot)$ is gamma.

2) Repeat step 1 for $j = 2, 3, \dots, m$, which lead to $\boldsymbol{\beta}^{(s+1)}$, $\boldsymbol{\alpha}^{(s+1)}$, $\mathbf{z}^{(s+1)}$ and $\boldsymbol{\sigma}^{(s+1)}$;

3) $(\mathbb{A}_1)_{ll}^{(s+1)} \sim \pi((\mathbb{A}_1)_{ll} | \Sigma^{(s)})$, where $\pi((\mathbb{A}_1)_{ll} | \cdot)$ is gamma truncated to (b_{0l}, b_{1l}) .

4) $(\mathbb{A}_2)_{ll}^{(s+1)} \sim \pi((\mathbb{A}_2)_{ll} | \Psi^{(s)})$, where $\pi((\mathbb{A}_2)_{ll} | \cdot)$ is gamma truncated to (c_{0l}, c_{1l}) .

5) $\Sigma^{-1(s+1)} \sim \pi(\Sigma^{-1} | \boldsymbol{\beta}^{(s+1)}, \mathbf{z}^{(s+1)}, \mathbb{A}_1^{(s+1)}, \boldsymbol{\mu}^{(s)})$, where $\pi(\Sigma^{-1} | \cdot)$ is Wishart.

6) $\Psi^{-1(s+1)} \sim \pi(\Psi^{-1} | \boldsymbol{\alpha}^{(s+1)}, \mathbf{z}^{(s+1)}, \mathbb{A}_2^{(s+1)}, \boldsymbol{\xi}^{(s)})$, where $\pi(\Psi^{-1} | \cdot)$ is Wishart.

7) $\sigma_\eta^{-2(s+1)} \sim \pi(\sigma_\eta^{-2} | \boldsymbol{\alpha}^{(s+1)}, \mathbf{z}^{(s+1)}, \eta^{(s)})$, where $\pi(\sigma_\eta^{-2} | \cdot)$ is gamma.

8) $\boldsymbol{\mu}_i^{(s+1)} \sim \pi(\boldsymbol{\mu}_i | \boldsymbol{\beta}^{(s+1)}, \mathbf{z}^{(s+1)}, \Sigma^{(s+1)})$ for $i = 1, 2, 3$, where $\pi(\boldsymbol{\mu}_i | \cdot)$ is normal.

9) $\boldsymbol{\xi}^{(s+1)} \sim \pi(\boldsymbol{\xi} | \boldsymbol{\alpha}^{(s+1)}, \mathbf{z}^{(s+1)}, \Psi^{(s+1)})$, where $\pi(\boldsymbol{\xi} | \cdot)$ is normal.

9) $\eta^{(s+1)} \sim \pi(\eta | \boldsymbol{\alpha}^{(s+1)}, \mathbf{z}^{(s+1)}, \sigma_\eta^{(s+1)})$, where $\pi(\eta | \cdot)$ is normal.

10) $(p_1, p_2, p_3) \sim \pi(p_1, p_2, p_3 | \mathbf{z}^{(s+1)})$, where $\pi(p_1, p_2, p_3 | \cdot)$ is Dirichlet.

The above cycle completes a transition from $\{\Theta^{(s)}, \mathbf{z}^{(s)}\}$ to $\{\Theta^{(s+1)}, \mathbf{z}^{(s+1)}\}$. Iterating through this cycle of generating data from the full conditional distributions leads to a Markov chain. Note that the full conditional distributions for all the components except (α_j, \mathbf{z}_j) can be expressed in closed-form expressions, and therefore we employ a Metropolis step to draw sample for (α_j, \mathbf{z}_j) . We have written our own code in C to generate MCMC samples, which we subsequently analyze using the *coda* package (Plummer et al., 2012) in R (R Development Core Team, 2014). The interface between C with R is achieved using *Rtools* (R Development Core Team, 2014).

4. Simulations

We present three scenarios to evaluate the performance of the proposed methodology in analyzing longitudinal samples coming from three different populations: G, A and L.

- **Scenario 1.** We generate data using (1a) $p_1 = 0.40$, $p_2 = 0.40$ and $p_3 = 0.20$, (1b) $p_1 = 0.40$, $p_2 = 0.20$ and $p_3 = 0.40$, and (1c) $p_1 = 0.20$, $p_2 = 0.40$ and $p_3 = 0.40$. Note that L is dominated by G and A in scenario 1a, A is dominated by G and L in scenario 1b, and G is dominated by A and L in scenario 1c. We then fit the proposed model to the simulated data to investigate the performance of our methodology when one population is dominated by the other two populations.

- Scenario 2.** Recall that our proposed methodology takes into account population-specific linear parameters to allow proper inference at the population level; we present scenario 2 to demonstrate that the proposed methodology performs well even if the populations share common intercept and slope parameters, which is one of the assumptions of Khan et al. (2013)’s methodology. As of scenario 1, we consider (2a) $p_1 = 0.40$, $p_2 = 0.40$ and $p_3 = 0.20$, (2b) $p_1 = 0.40$, $p_2 = 0.20$ and $p_3 = 0.40$, and (2c) $p_1 = 0.20$, $p_2 = 0.40$ and $p_3 = 0.40$
- Scenario 3.** We present scenario 3 to illustrate that the Wishart priors can be sensitive to the choices of the scale matrices \mathbb{A} (scenario 3a), whereas the hierarchical Wishart priors lead to robust estimates of the covariance matrices under sensible choices of the hyperparameters (scenario 3b). We only consider simulated $p_1 = 0.40$, $p_2 = 0.40$ and $p_3 = 0.20$ for both scenarios 3a and 3b.

Table 1. Simulation scenario 1a results with simulated $p_1 = 0.40$, $p_2 = 0.40$ and $p_3 = 0.20$, and $m = 40$: average of 500 posterior medians, and coverage of 95% credible intervals.

Population	Parameter	True value	$n = n_j = 100$ ($j = 1, 2, \dots, 40$)		$n = n_j = 24$ ($j = 1, 2, \dots, 40$)	
			Mean	Coverage	Mean	Coverage
G	μ_{10}	244.00	243.872	0.944	243.920	0.954
	μ_{11}	0.50	0.501	0.968	0.502	0.964
	μ_{12}	-0.75	-0.752	0.972	-0.748	0.958
	ξ_1	1.50	1.499	0.990	1.500	0.970
	ξ_2	2.50	2.497	0.970	2.496	0.980
	$(\Psi)_{11}$	0.020	0.021	0.976	0.022	0.986
	$(\Psi)_{22}$	0.030	0.033	0.972	0.032	0.966
	$(\Psi)_{12}$	0.005	0.004	0.992	0.005	0.996
A	μ_{20}	230.00	229.947	0.948	230.115	0.958
	μ_{21}	1.00	1.001	0.956	1.001	0.950
	μ_{22}	-1.25	-1.248	0.978	-1.251	0.972
	η	2.00	2.000	0.952	1.999	0.952
	σ_η^2	0.050	0.054	0.948	0.052	0.958
L	μ_{30}	250.00	250.028	0.970	250.103	0.950
	$\mu_{31} + \mu_{32}$	-0.25	-0.248	0.966	-0.254	0.958
Common parameters	$(\Sigma)_{11}$	125.000	126.067	0.960	128.319	0.966
	$(\Sigma)_{22}$	0.030	0.031	0.970	0.031	0.964
	$(\Sigma)_{33}$	0.030	0.030	0.980	0.031	0.974
	$(\Sigma)_{12}$	-1.000	-1.012	0.974	-1.045	0.960
	$(\Sigma)_{13}$	0.500	0.483	0.962	0.518	0.962
	$(\Sigma)_{23}$	-0.010	-0.010	0.984	-0.011	0.970
Mixing proportion	p_1	0.40	0.392	0.938	0.396	0.948
	p_2	0.40	0.400	0.952	0.393	0.946
	p_3	0.20	0.208	0.956	0.211	0.948

We investigate the performance of our methodology with fairly vague, minimally informative priors. For multivariate normal, the hyperparameters – mean vector and covariance matrix – are commonly chosen to be a zero vector and a diagonal matrix, say, \mathbb{H} , respectively, such that $\mathbb{H}^{-1} \approx \mathbb{O}$, where \mathbb{O} is a matrix with all its elements zero. We follow this convention which leads to a flat prior (Davidian & Giltinan, 1995). Similarly, we consider flat prior for the univariate normal distribution: $a_0 = 0$ and $a_1 = 10^5$. For the hierarchical Wishart priors in scenarios 1 and 2, we choose $(\mathbb{A}_1)_{11} \sim U(100, 150)$, $(\mathbb{A}_1)_{22} \sim U(0, 10)$, $(\mathbb{A}_1)_{33} \sim U(0, 10)$, $(\mathbb{A}_2)_{11} \sim U(0, 10)$ and $(\mathbb{A}_2)_{22} \sim U(0, 10)$. For a gamma prior, small values of the hyperprior parameters lead to a flat prior; we consider $d_0 = e_0 = 0.01$ and $d_1 = e_1 = 0.01$ in all simulation scenarios. Finally, for the mixture probabilities (p_1, p_2, p_3) , we choose the dirichlet prior $Dirichlet(\omega_1, \omega_2, \omega_3)$ with $\omega_1 = \omega_2 = \omega_3 = 1$.

In all the scenarios, we take $m = 40$ and define time by $t_{jk} = k - 1$ for $k = 1, 2, \dots, n_j$ and $j = 1, 2, \dots, m$. Simulations for each scenario are performed using both a large number of repeated measurements ($n = n_j = 100$) and a moderate number of repeated measurements ($n = n_j = 24$). Model parameter values are chosen to allow reasonable generalization and are given in Tables 1-4. Note that σ_j^2 's ($j = 1, 2, \dots, 40$) were randomly drawn from

the $U(0, 0.01)$ distribution. For each simulation, 500 data sets were generated, and 100,000 MCMC iterations after burn-in were used to approximate posterior distributions per set. Posterior summaries were averaged over the 500 sets for each parameter, and the coverage probability of 95% credible intervals (proportion of such credible intervals out of 500 that capture the truth) was calculated.

4.1 Results for Scenario 1

Numerical results for scenario 1a are presented in Table 1, with the top three panels depict the results for the parameters relevant to only Populations G, A and L, respectively, the second last panel presents the results for the covariance matrix Σ which is shared by all the three populations, and the last panel shows results for the mixing proportions (p_1, p_2, p_3) . We see that our methodology performs well with respect to all the parameters for both large and moderate number of repeated measurements (i.e., $n = n_j = 100$ and $n = n_j = 24$): averages of posterior medians are all close to the true parameter values, and coverage probabilities are all reasonably close to the nominal 0.95. In particular, coverage probabilities range from 0.938 to 0.992 for $n = 100$ and from 0.946 to 0.996 for $n = 24$. We observed similar results for scenarios 1b and 1c (numerical results for scenarios 1b and 1c are not presented in this article).

Table 2. Simulation scenario 2a results with simulated $p_1 = 0.40$, $p_2 = 0.40$ and $p_3 = 0.20$, and $m = 40$: average of 500 posterior medians, and coverage of 95% credible intervals.

Population	Parameter	True value	$n = n_j = 100$ ($j = 1, 2, \dots, 40$)		$n = n_j = 24$ ($j = 1, 2, \dots, 40$)	
			Mean	Coverage	Mean	Coverage
G	μ_{10}	244.00	244.020	0.958	244.062	0.938
	μ_{11}	0.50	0.499	0.978	0.499	0.964
	μ_{12}	-0.75	-0.746	0.940	-0.747	0.970
	ξ_1	1.50	1.497	0.970	1.497	0.982
	ξ_2	2.50	2.499	0.974	2.498	0.978
	$(\Psi)_{11}$	0.020	0.022	0.970	0.020	0.978
	$(\Psi)_{22}$	0.030	0.032	0.976	0.032	0.962
	$(\Psi)_{12}$	0.005	0.004	0.982	0.004	0.996
A	μ_{20}	244.00	244.085	0.958	244.098	0.956
	μ_{21}	0.50	0.499	0.980	0.499	0.976
	μ_{22}	-0.75	-0.750	0.962	-0.751	0.974
	η	2.00	2.002	0.956	2.002	0.964
	σ_η^2	0.050	0.053	0.948	0.053	0.954
L	μ_{30}	244.00	243.875	0.944	243.929	0.952
	$\mu_{31} + \mu_{32}$	-0.25	-0.251	0.982	-0.256	0.976
Common parameters	$(\Sigma)_{11}$	125.000	125.409	0.978	125.528	0.958
	$(\Sigma)_{22}$	0.030	0.030	0.972	0.031	0.956
	$(\Sigma)_{33}$	0.030	0.031	0.972	0.031	0.966
	$(\Sigma)_{12}$	-1.000	-1.003	0.960	-1.029	0.952
	$(\Sigma)_{13}$	0.500	0.515	0.952	0.519	0.978
	$(\Sigma)_{23}$	-0.010	-0.011	0.982	-0.011	0.978
Mixing proportion	p_1	0.40	0.392	0.952	0.401	0.960
	p_2	0.40	0.395	0.958	0.389	0.948
	p_3	0.20	0.213	0.954	0.210	0.956

4.2 Results for Scenario 2

As mentioned in Section 1, the assumption of the Khan et al. (2013)'s methodology that the populations share common intercept and slope parameters is restrictive, because the bent-cable methodology under this assumption cannot properly characterize the populations, especially in describing the linear phases. Our preliminary simulation study supports this fact: serious estimation bias can occur under this assumption when, in reality, those differ across the populations. Simulation scenario 2 is introduced to evaluate the performance of our methodology under the assumption that the linear parameters do not differ across the populations, that is, we analyze the simulated data using our methodology which are generated assuming that the populations share common intercept and slope parameters. Numerical results for scenario 2a are summarized in Table 2. We see that our methodology performs

well with respect to all the parameters for scenario 2a: averages of posterior medians are all close to the true parameter values, and coverage probabilities are all reasonably close to the nominal 0.95, ranging from 0.940 to 0.982 for $n = n_j = 100$ and 0.938 to 0.996 for $n = n_j = 24$. We observed similar results for scenarios 2b and 2c (numerical results for scenarios 2b and 2c are not presented in this article).

Table 3. Simulation scenario 3a (Wishart priors) results with simulated $p_1 = 0.40$, $p_2 = 0.40$ and $p_3 = 0.20$, and $n_j = 24$ for all j and $m = 40$: average of 500 posterior medians, and coverage of 95% credible intervals.

		Wishart prior						
		Prior ₁ 1			Prior ₁ 2		Prior ₁ 3	
		$\mathbb{A}_1 = \text{diag}(125, 0.05, 0.05)$ $\mathbb{A}_2 = \text{diag}(0.05, 0.05)$			$\mathbb{A}_1 = \text{diag}(100, 0.15, 0.15)$ $\mathbb{A}_2 = \text{diag}(0.15, 0.15)$		$\mathbb{A}_1 = \text{diag}(100, 0.25, 0.25)$ $\mathbb{A}_2 = \text{diag}(0.25, 0.25)$	
Population	Parameter	True value	Mean	Coverage	Mean	Coverage	Mean	Coverage
G	μ_{10}	244.00	243.951	0.964	244.005	0.956	243.898	0.954
	μ_{11}	0.50	0.505	0.974	0.499	0.986	0.499	0.998
	μ_{12}	-0.75	-0.754	0.966	-0.755	0.988	-0.752	0.994
	ξ_1	1.50	1.499	0.982	1.501	0.999	1.502	0.999
	ξ_2	2.50	2.496	0.976	2.500	0.986	2.501	0.994
	$(\Psi)_{11}$	0.020	0.024	0.996	0.040	0.752	0.055	0.052
	$(\Psi)_{22}$	0.030	0.032	0.988	0.043	0.938	0.056	0.718
	$(\Psi)_{12}$	0.005	0.005	0.999	0.006	0.999	0.006	0.999
A	μ_{20}	230.00	230.070	0.968	230.088	0.952	229.924	0.964
	μ_{21}	1.00	0.999	0.980	0.999	0.986	1.000	0.996
	μ_{22}	-1.25	-1.251	0.962	-1.251	0.984	-1.248	0.999
	η	2.00	2.004	0.956	1.997	0.964	1.998	0.952
	σ_η^2	0.050	0.054	0.946	0.053	0.960	0.052	0.954
	μ_{30}	250.00	250.113	0.944	249.840	0.944	249.983	0.946
L	$\mu_{31} + \mu_{32}$	-0.25	-0.242	0.974	-0.248	0.990	-0.250	0.996
Common parameters	$(\Sigma)_{11}$	125.000	125.968	0.980	123.212	0.962	123.435	0.962
	$(\Sigma)_{22}$	0.030	0.032	0.990	0.045	0.876	0.049	0.494
	$(\Sigma)_{33}$	0.030	0.032	0.970	0.043	0.892	0.049	0.532
	$(\Sigma)_{12}$	-1.000	-1.001	0.968	-1.004	0.968	-1.001	0.986
	$(\Sigma)_{13}$	0.500	0.508	0.980	0.441	0.974	0.506	0.990
	$(\Sigma)_{23}$	-0.010	-0.011	0.990	-0.010	0.996	-0.013	0.999
	Mixing proportion	p_1	0.40	0.394	0.938	0.397	0.934	0.393
	p_2	0.40	0.395	0.950	0.397	0.948	0.399	0.944
	p_3	0.20	0.211	0.944	0.206	0.966	0.208	0.962

4.3 Results for Scenario 3

First, we present simulation results when the covariance matrices are modeled using the Wishart distributions (scenario 3a). To make the Wishart priors for Σ^{-1} and Ψ^{-1} nearly flat, we choose $\nu_1 = 3$ and $\nu_2 = 2$. The scale matrices \mathbb{A}_1 and \mathbb{A}_2 are considered diagonals in each simulation. Note that the diagonal elements of true Σ and Ψ are (125.00, 0.03, 0.03) and (0.02, 0.03), respectively. To investigate the sensitivity of the specification of Wishart distributions in modeling the covariance matrices, we consider three sets of priors: the diagonal elements of \mathbb{A}_1 and \mathbb{A}_2 are all relatively close to those of the true Σ and Ψ , respectively, with the closest one specified as Prior₁1 : $\{\mathbb{A}_1 = \text{diag}(125, 0.05, 0.05), \mathbb{A}_2 = \text{diag}(0.05, 0.05)\}$, the farthest one as Prior₁3 : $\{\mathbb{A}_1 = \text{diag}(100, 0.25, 0.25), \mathbb{A}_2 = \text{diag}(0.25, 0.25)\}$, and an intermediate to these two specifications as Prior₁2 : $\{\mathbb{A}_1 = \text{diag}(100, 0.15, 0.15), \mathbb{A}_2 = \text{diag}(0.15, 0.15)\}$. Numerical results under these three sets of priors for $n = n_j = 24$ are summarized in Table 3. We see that our methodology with Prior₁1 performs well with respect to all the parameters: averages of posterior medians are all close to the true parameter values, and coverage probabilities are all reasonably close to the nominal 0.95. However, as the priors for \mathbb{A}_1 and \mathbb{A}_2 deviate slightly from those of Prior₁1 but still reasonably close to the truth, under-coverage and large bias result for at least one of the diagonal elements (variance parameters) of Σ and Ψ . For example, under Prior₁1 and Prior₁2, the coverage probabilities for $(\Psi)_{11}$ are 0.996 and 0.752, respectively; thus, the coverage under Prior₁1 is 77% closer (Note 1) to the nominal 0.95 when compared with Prior₁2. We see the worst scenario under Prior₁3: the coverage probabilities for $(\Psi)_{11}$, $(\Psi)_{22}$, $(\Sigma)_{22}$ and $(\Sigma)_{33}$ are 0.052, 0.718, 0.494 and 0.532, respectively, each of which is substantially lower than the nominal 0.95. The above results support the fact that the bent-cable methodology can be sensitive to the choices of the hyperparameters for the Wishart priors: the model performs well when \mathbb{A}_1 and \mathbb{A}_2 are very close to the truth, whereas slight deviations from the truth can result in drastically lower coverage probabilities for the variance parameters in Σ and Ψ . Scenario 3a with $n = n_j = 100$ also leads to the same conclusion (numerical results for $n = n_j = 100$ are not presented in this article).

Now, we consider the case when the covariance matrices are modeled using the hierarchical Wishart priors (scenario 3b). We consider three sets of priors as presented in Table 4, with Prior₂1 is relatively the most informative, Prior₂3 is the least informative, and Prior₂2 is in between Prior₂2 and Prior₂3. Numerical results for $n = n_j = 24$ are summarized in Table 4. Under all the three sets of priors, our methodology performs well with respect to all the parameters: coverage probabilities range from 0.946 to 0.996 for both Prior₂1 and Prior₂2, and from 0.948 to 0.998 for Prior₂3. Averages of the posterior medians are all very close to the true parameter values as well. Scenario 3b with $n = n_j = 100$ also leads to similar results (numerical results for $n = n_j = 100$ are not presented here).

Table 4. Simulation scenario 3b (hierarchical Wishart) results with simulated $p_1 = 0.40$, $p_2 = 0.40$ and $p_3 = 0.20$, and $n_j = 24$ for all j and $m = 40$: average of 500 posterior medians, and coverage of 95% credible intervals.

		Hierarchical Wishart prior							
		Prior ₂ 1		Prior ₂ 2		Prior ₂ 3			
		$(\mathbb{A}_1)_{11} \sim U(100, 150)$		$(\mathbb{A}_1)_{11} \sim U(50, 150)$		$(\mathbb{A}_1)_{11} \sim U(50, 150)$			
		$(\mathbb{A}_1)_{22} \sim U(0, 10)$		$(\mathbb{A}_1)_{22} \sim U(0, 20)$		$(\mathbb{A}_1)_{22} \sim U(0, 30)$			
		$(\mathbb{A}_1)_{33} \sim U(0, 10)$		$(\mathbb{A}_1)_{33} \sim U(0, 20)$		$(\mathbb{A}_1)_{33} \sim U(0, 30)$			
		$(\mathbb{A}_2)_{11} \sim U(0, 10)$		$(\mathbb{A}_2)_{11} \sim U(0, 20)$		$(\mathbb{A}_2)_{11} \sim U(0, 30)$			
		$(\mathbb{A}_2)_{22} \sim U(0, 10)$		$(\mathbb{A}_2)_{22} \sim U(0, 20)$		$(\mathbb{A}_2)_{22} \sim U(0, 30)$			
Population	Parameter	True value	Mean	Coverage	Mean	Coverage	Mean	Coverage	
G	μ_{10}	244.00	243.920	0.954	243.983	0.946	244.176	0.948	
	μ_{11}	0.50	0.502	0.964	0.502	0.956	0.497	0.962	
	μ_{12}	-0.75	-0.748	0.958	-0.753	0.972	-0.749	0.966	
	ξ_1	1.50	1.500	0.970	1.501	0.982	1.503	0.970	
	ξ_2	2.50	2.496	0.980	2.501	0.970	2.501	0.972	
	$(\Psi)_{11}$	0.020	0.022	0.986	0.022	0.972	0.021	0.980	
	$(\Psi)_{22}$	0.030	0.032	0.966	0.033	0.982	0.032	0.982	
	$(\Psi)_{12}$	0.005	0.005	0.996	0.004	0.996	0.005	0.998	
A	μ_{20}	230.00	230.115	0.958	229.821	0.952	229.992	0.968	
	μ_{21}	1.00	1.001	0.950	1.002	0.968	1.000	0.972	
	μ_{22}	-1.25	-1.251	0.972	-1.252	0.970	-1.249	0.964	
	η	2.00	1.999	0.952	2.003	0.954	2.000	0.968	
	σ_η^2	0.050	0.052	0.958	0.052	0.962	0.056	0.968	
L	μ_{30}	250.00	250.103	0.950	250.232	0.960	250.209	0.952	
	$\mu_{31} + \mu_{32}$	-0.25	-0.254	0.958	-0.245	0.974	-0.247	0.972	
Common parameters	$(\Sigma)_{11}$	125.000	128.319	0.966	122.996	0.960	123.627	0.970	
	$(\Sigma)_{22}$	0.030	0.031	0.964	0.030	0.958	0.030	0.968	
	$(\Sigma)_{33}$	0.030	0.031	0.974	0.031	0.976	0.031	0.962	
	$(\Sigma)_{12}$	-1.000	-1.045	0.960	-0.987	0.956	-1.020	0.960	
	$(\Sigma)_{13}$	0.500	0.518	0.962	0.488	0.974	0.495	0.972	
	$(\Sigma)_{23}$	-0.010	-0.011	0.970	-0.011	0.974	-0.011	0.976	
Mixing proportion	p_1	0.40	0.396	0.948	0.396	0.954	0.395	0.950	
	p_2	0.40	0.393	0.946	0.396	0.946	0.398	0.954	
	p_3	0.20	0.211	0.948	0.208	0.966	0.207	0.948	

In summary, the proposed methodology based on the Wishart priors for the covariance matrices performs well as long as \mathbb{A}_1 and \mathbb{A}_2 almost accurately approximate the covariance matrices Σ and Ψ , respectively, (e.g., Prior₁1 in Table 3). Note that such accurate approximation can rarely be achieved in practical situations. Slight departures of \mathbb{A}_1 and \mathbb{A}_2 from the truth may result in severe under-coverage for the variance parameters in Σ and Ψ . In contrast, the proposed methodology based on the hierarchical Wishart priors performs well under sensible choices of the priors for \mathbb{A}_1 and \mathbb{A}_2 .

Numerical results for the σ_j^2 's are not presented in this article. For each simulation scenario considered in this study, the average of the posterior medians for each σ_j^2 were very close to the truth, and the coverage probabilities were all close to the nominal 0.95.

5. Example: Rat Data Analysis

In Section 1, we introduced an experiment on 38 rats to understand the effects of the hypothermia therapy on core body temperature, T_c . This dataset was analyzed by Khan et al. (2013) using their longitudinal bent-cable methodology which accounts for only gradual and abrupt transitions. In this section, we re-analyze this dataset using our proposed methodology which simultaneously accounts for transitional (gradual and abrupt) as well as linear trends over time. We also compare the fits of the proposed model with that of Khan et al. (2013)'s using the deviance information criterion (DIC) (Spiegelhalter et al., 2002).

As in Khan et al. (2013), we define time by t_{jk} , $k = 0, 1, \dots, n_j$ where $t_{j1} = 0$ refers to the starting point of the

study for rat j ($j = 1, 2, \dots, 38$), and each subsequent time increment is 15 seconds. Since any function $h(\Theta)$ of the parameters has its own posterior distribution, we consider the posterior median of $h(\Theta)$ as its estimate. For example, the CTP for Population G is $h(\Theta) = \exp\{\xi_2\} - \exp\{\xi_1\} - 2\mu_{11} \exp\{\xi_1\}/\mu_{12}$, and therefore the posterior median of $\{\exp\{\xi_2^{(l)}\} - \exp\{\xi_1^{(l)}\} - 2\mu_{11}^{(l)} \exp\{\xi_1^{(l)}\}/\mu_{12}^{(l)}, l = S, S + 1, \dots, T\}$ is considered as its estimate. Similarly, the MCMC sample median of $f(t_{jk}, \theta_j)$ is considered as the fitted value \hat{y}_{jk} at time t_{jk} for individual j . A fitted population curve is produced based on the estimates of the theoretical medians for β_j and α_j from Level 2.

5.1 Results

The DIC for the proposed model and the Khan et al. (2013)'s model fit to the rat data are approximately $-16,739$ and $-14,729$, respectively, suggesting that our methodology provides substantial improvement in fit, with DIC considerably less than that reported by Khan et al. (2013).

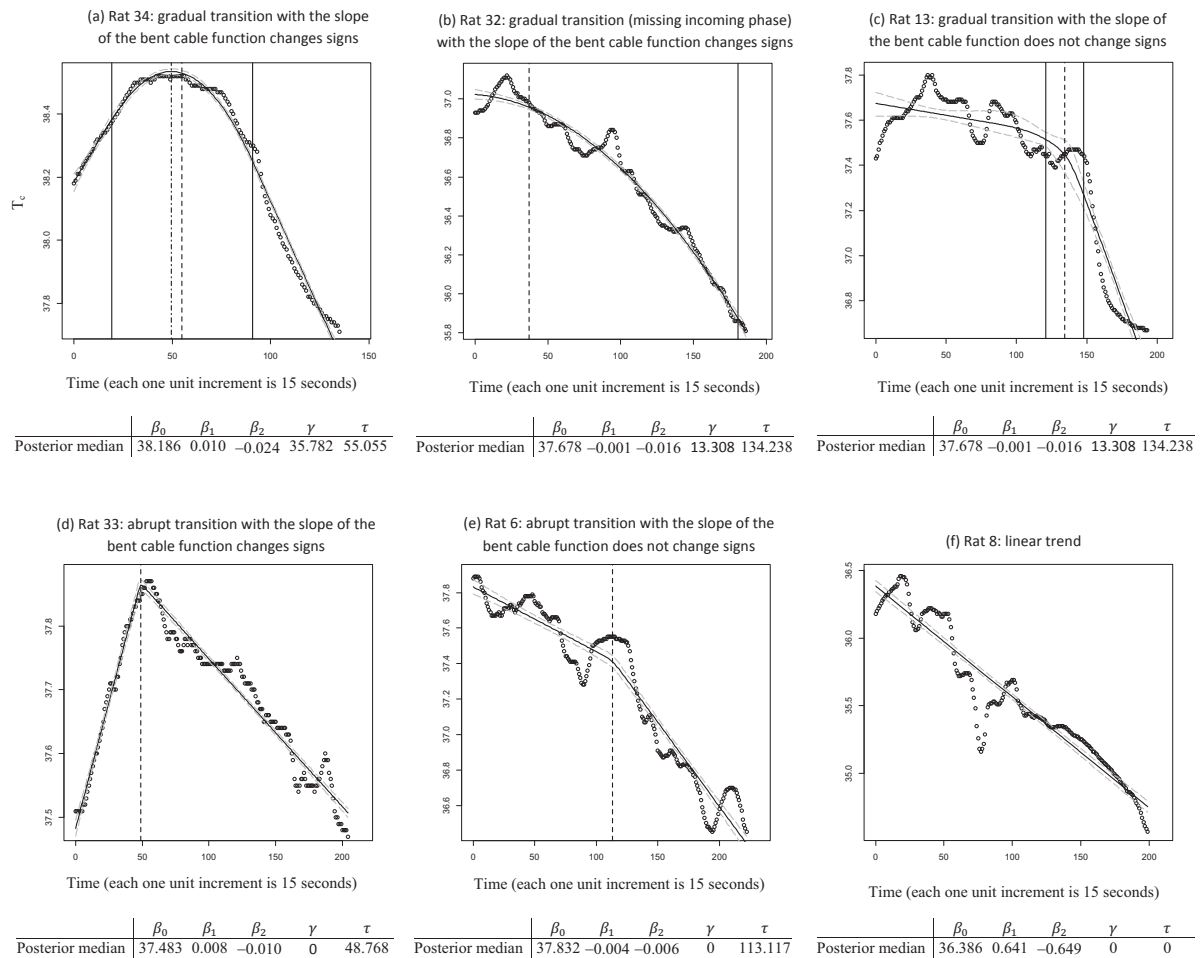


Figure 3. Fits of six representative rat profiles out of 38. The dotted curves indicate the observed profiles. Superimposed on the plots are the fitted curves/lines along with 95% pointwise credible intervals indicated by the dashed curves. Rats 34, 32 and 13 are estimated to have arisen from Population G (gradual transition) for which the transition periods are marked by the solid vertical lines, the centers of the bends by the dashed vertical lines, and the CTP (for rat 34) by the dot-dashed vertical line. Rats 33 and 6 are estimated to have arisen from Population A (abrupt transition) for which the changes in slopes occur at the dashed vertical lines. Rat 8 exhibits simply a linear decreasing trend with estimated slope -0.008 (i.e., $\widehat{\beta_1 + \beta_2}$).

Figure 3 displays the fits of six representative rat profiles along with the posterior medians for the regression coefficients. Rats 34, 32 and 13 are estimated to have arisen from Population G for which both $\hat{\gamma}$ and $\hat{\tau}$ are positive. Rat 34 exhibits a linear increasing trend in T_c , followed by a gradual transition and a linear decreasing trend thereafter, whereas rat 13 exhibits linear decreasing trends at either end of the profile, with a continuous transition between the two phases (i.e., the fitted slope of the bent cable changes only its magnitudes but not signs). In

total, sixteen rats are characterized by gradual transition of which twelve exhibit sign changes in the fitted slope of the bent cable function (i.e., an incoming phase with linear increasing trend and an outgoing phase with linear decreasing trend are joined by a quadratic bend; see rat 34 in Figure 3), and only four show no sign change in the fitted slope of the bent cable function (e.g., rat 13). Note that rat 32 exhibits a gradual drop in T_c from the beginning of the study period. The missing incoming phase for this profile suggests that it does not obviously follow the shape of the bent cable. Our methodology picks up this anomaly adequately – a gradual transition is in progress for this rat from the beginning of the study period. Note that one of the appealing features of the multilevel model is that pooling information from many individuals leads to shrinkage towards the population, so that mild deviations of observed profiles from the bent cable structure do not hinder model fitting (Khan et al., 2013). This feature leads the estimates of the parameters characterizing a missing phase as in rat 32 to shrink towards their population counterparts. Rats 33 and 6 exhibit abrupt transition for which $\hat{\gamma} = 0$ and $\hat{\tau} > 0$. In total, eighteen rats are characterized by abrupt transition of which sixteen exhibit sign changes in the fitted slope of the bent cable function (i.e., an incoming phase with linear increasing trend is joined by an outgoing phase with linear decreasing trend; see rat 33 in Figure 3), and only two exhibit no sign change in the fitted slope of the bent cable function (e.g., rat 6). Finally, only four rat profiles show linear decreasing trend throughout (e.g., rat 8). Note that about 40%, 50% and 10% of the rats belong to Populations G, A and L, respectively, based on the posterior medians of p_1 , p_2 and p_3 ($\hat{p}_1 = 0.40$, $\hat{p}_2 = 0.50$ and $\hat{p}_3 = 0.10$; see Table 5). In summary, the fits appear very reasonable as the observed data agree closely with the respective fitted curves, and the estimated τ and γ demonstrate that our methodology picks up the three types of trend adequately.

Table 5 summarizes the posterior characteristics of the population quantities, and Figure 4 displays the fitted curves for Populations G, A and L overlaid with the rat profiles which exhibit gradual and abrupt transitions and linear decreasing trend, respectively.

Characterization of Population G. Population G is characterized by three phases: (a) a linear incoming phase with estimated slope 0.005 (posterior median for μ_{11} is 0.005 with 95% credible interval (0.001, 0.010)), (b) a transitional phase ranging from 29.732 to 113.503 (posterior medians for $\exp\{\xi_2\} \pm \exp\{\xi_1\}$ are 29.732 and 113.503), and (c) a linear outgoing phase with estimated slope -0.019 (the posterior median for $\mu_{11} + \mu_{12}$ is -0.019 with 95% credible interval $(-0.026, -0.012)$). Thus, on average, T_c for Population G rats significantly increases at the rate of $0.005^\circ C$ per 15 seconds in the incoming phase, followed by a gradual transition which begins at approximately 7.433 minutes ($(29.732 \times 15)/60 = 7.433$) from the time of hemorrhage and lasts for about 28.376 minutes ($(113.503 \times 15)/60 = 28.376$), and a significant linear decrease at the rate of $-0.019^\circ C$ per 15 seconds in the outgoing phase. The estimate of the CTP for Population G is 48.017, suggesting that the drop in T_c begins at approximately 12 minutes after hemorrhage. Note that since a few rat profiles do not obviously follow the shape of the bent cable because of missing incoming and/or outgoing phases, the estimation of the transition parameters $\exp\{\xi_1\}$, $\exp\{\xi_2\}$ and CTP for Population G involves considerable uncertainty. This phenomenon is reflected in the credible intervals and standard deviations of the transition parameters: we see wider credible intervals for $\exp\{\xi_1\}$, $\exp\{\xi_2\}$ and CTP, as well as for the posterior standard deviations of γ_j and τ_j .

Characterization of Population A. Population A is characterized by two phases: (a) a linear incoming phase with estimated slope 0.004 (posterior median for μ_{21} is 0.004 with 95% credible interval $(-0.0002, 0.0085)$), and (b) a linear outgoing phase with estimated slope -0.009 (posterior median for $\mu_{22} + \mu_{23}$ is -0.009 with 95% credible interval $(-0.015, -0.004)$). Note that the incoming slope is not statistically significant, as the 95% credible interval for μ_{21} includes zero. Since the posterior median for the incoming intercept, μ_{20} , is 37.375, T_c for Population A rats remains more or less constant at around $37.375^\circ C$ until approximately 14.11 minutes elapsed (the posterior median for the Population A CTP is $(56.430 \times 15)/60 = 14.11$ minutes) before a significant drop occurs at the rate of $-0.009^\circ C$ per 15 seconds. Note that the relatively narrower credible intervals for the CTP of Population A (i.e., $\exp\{\eta\}$) and the standard deviation of τ_j indicate less uncertainty in estimating the transition parameter for Population A.

Characterization of Population L. The posterior median for $\mu_{31} + \mu_{32}$ is -0.005 , suggesting that the rate of decrease in T_c for Population L rats is around $-0.005^\circ C$ per 15 seconds. However, the 95% credible interval for $\mu_{31} + \mu_{32}$ is $(-0.015, 0.006)$, which includes zero. Thus, the rate of decrease in T_c for Population L is not statistically significant. Since the the posterior median for the intercept parameter, μ_{30} , is 37.147, T_c for Population L rats remains approximately constant at around $37.147^\circ C$ until the end of the study period.

Table 5. Posterior summaries for the three populations of rats (the time scale is defined by nonnegative integers with each time increment is 15 seconds).

	Parameter	Posterior median	95% credible interval
Population G	Incoming intercept (μ_{10})	37.490	(37.200, 37.780)
	Incoming slope (μ_{11})	0.005	(0.001, 0.010)
	Outgoing slope ($\mu_{11} + \mu_{12}$)	-0.019	(-0.026, -0.012)
	Half width of the bend ($\exp\{\xi_1\}$)	41.886	(16.542, 67.703)
	Center of the bend ($\exp\{\xi_2\}$)	71.618	(40.463, 104.446)
	CTP ($\exp\{\xi_2\} - \exp\{\xi_1\} - 2\mu_{11} \exp\{\xi_1\}/\mu_{12}$)	48.017	(6.947, 86.870)
	Transition period ($\exp\{\xi_2\} \pm \exp\{\xi_1\}$)	29.732 to 113.503	
	S.D. of γ_j ($(\exp\{2\xi_1 + (\Psi)_{11}\} \times [\exp\{(\Psi)_{11}\} - 1])^{1/2}$)	65.585	(15.730, 404.003)
	S.D. of τ_j ($(\exp\{2\xi_2 + (\Psi)_{22}\} \times [\exp\{(\Psi)_{22}\} - 1])^{1/2}$)	67.421	(24.451, 202.086)
	Corr. between γ_j and τ_j	-0.203	(-0.560, 0.183)
	$([\exp\{(\Psi)_{12}\} - 1] / \sqrt{[\exp\{(\Psi)_{11}\} - 1] \times [\exp\{(\Psi)_{22}\} - 1]})$		
Population A	Incoming intercept (μ_{20})	37.375	(37.103, 37.649)
	Incoming slope (μ_{21})	0.004	(0.000, 0.009)
	Outgoing slope ($\mu_{21} + \mu_{22}$)	-0.009	(-0.015, -0.004)
	Critical time point ($\exp\{\eta\}$)	56.430	(41.824, 71.484)
	S.D. of τ_j ($(\exp\{2\eta + \sigma_\eta^2\} \times [\exp\{\sigma_\eta^2\} - 1])^{1/2}$)	29.367	(13.646, 56.456)
Population L	Intercept (μ_{30})	37.147	(36.622, 37.664)
	slope ($\mu_{31} + \mu_{32}$)	-0.005	(-0.015, 0.006)
Common parameters	S.D. of β_{0j} ($(\Sigma)_{11})^{1/2}$)	0.541	(0.417, 0.693)
	S.D. of β_{1j} ($(\Sigma)_{22})^{1/2}$)	0.008	(0.006, 0.011)
	S.D. of β_{2j} ($(\Sigma)_{33})^{1/2}$)	0.012	(0.008, 0.015)
	S.D. of $\beta_{1j} + \beta_{2j}$ ($(\Sigma)_{22} + (\Sigma)_{33} + 2 \times (\Sigma)_{23})^{1/2}$)	0.011	(0.008, 0.015)
	Corr. between β_{0j} and β_{1j} ($(\Sigma)_{12}/\{(\Sigma)_{11} \times (\Sigma)_{22}\}^{1/2}$)	-0.089	(-0.440, 0.261)
	Corr. between β_{0j} and β_{2j} ($(\Sigma)_{13}/\{(\Sigma)_{11} \times (\Sigma)_{33}\}^{1/2}$)	0.241	(-0.11, 0.564)
	Corr. between β_{1j} and β_{2j} ($(\Sigma)_{23}/\{(\Sigma)_{22} \times (\Sigma)_{33}\}^{1/2}$)	-0.425	(-0.695, -0.101)
Mixing proportion	Probability of coming from Population G (p_1)	0.40	(0.245, 0.586)
	Probability of coming from Population A (p_2)	0.50	(0.279, 0.627)
	Probability of coming from Population L (p_3)	0.10	(0.037, 0.245)

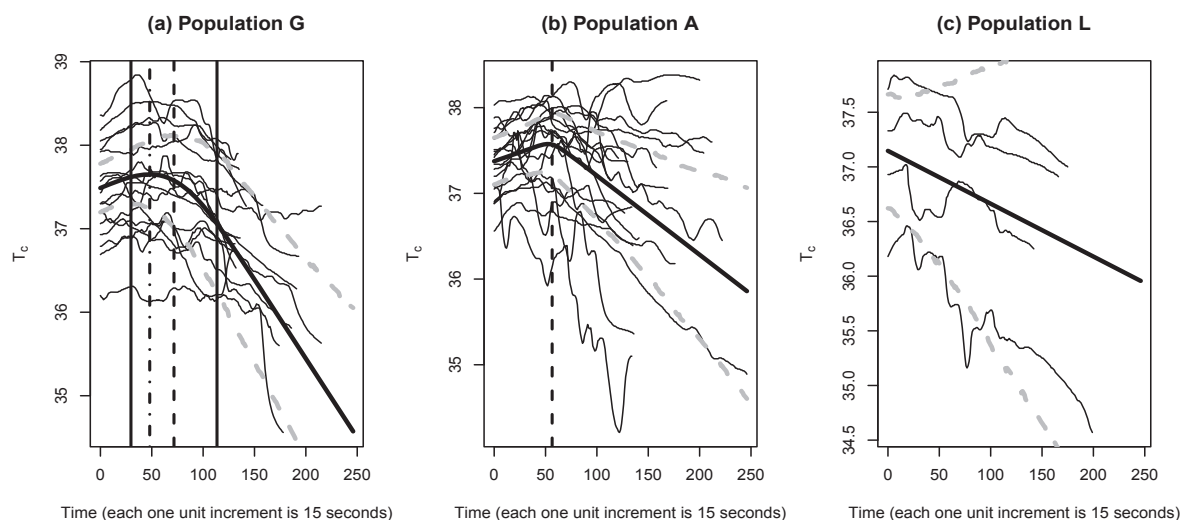


Figure 4. Population fitted curves (in bold) along with 95% pointwise credible intervals (gray curves). (a) Population G fitted curve overlaid with the sixteen rat profiles which exhibit gradual transition. The estimated transitions ($\widehat{\exp\{\xi_2\} \pm \exp\{\xi_1\}}$) are marked by the solid vertical lines, the center of the bend ($\widehat{\exp\{\xi_2\}}$) by the dashed line and the CTP by the dot-dashed line. (b) Population A fitted curve overlaid with the eighteen rat profiles which exhibit abrupt transition; the estimated transition ($\widehat{\exp\{\eta\}}$) is marked by the dashed vertical line. (c) Population L fitted curve overlaid with the four rat profiles which exhibit decreasing trend.

Our analysis reveals very similar results for Σ as reported by Khan et al. (2013). Small variation in the slope parameters (the posterior medians for the standard deviations of β_{1j} and $\beta_{1j} + \beta_{2j}$ are 0.008 and 0.011, respectively) suggests that all rats exhibit very similar rates of increase/decrease before/after the transition period. The posterior

median for the standard deviation of the intercept parameter β_{0j} is 0.541, suggesting some variation in T_c 's at the time of administering hemorrhage. Significant negative correlation between β_{1j} and β_{2j} ($\text{corr}(\widehat{\beta}_{1j}, \widehat{\beta}_{2j}) = -0.425$ with 95% credible interval $(-0.695, -0.101)$) indicates that the steeper the incoming slope going up, the bigger the drop in slope for the outgoing phase.

We summarize here a few interesting findings of our analysis.

- If a rat exhibits a significant increase in T_c at the rate of approximately $0.005^\circ C$ per 15 seconds in the incoming phase, it is highly likely that it belongs to Population G. On the other hand, a rat exhibiting approximately a constant trend in T_c in the incoming phase is more likely to have come from Population A or L. These phenomena are supported by the facts that the incoming slope for Population G is statistically significant, but the incoming slope for Population A and the slope for Population L are both statistically insignificant.
- If a rat exhibits approximately a constant trend in T_c in the incoming phase, the T_c will either remain roughly constant over time (Population L for which the rate of decrease is not significant) or will decrease at a significant rate of approximately $-0.009^\circ C$ per 15 seconds after around 14.11 minutes from hemorrhage (Population A for which the rate of decrease after around 14.11 minutes from hemorrhage is significant).
- If a rat exhibits a gradual but nonlinear drop in T_c after hemorrhage (i.e., missing incoming phase), it is more likely that it belongs to Population G.
- The hypothermia therapy appears to have no significant effects in reducing T_c for about 10% of the rats which belong to Population L.

6. Conclusion

In many longitudinal studies, an individual may exhibit simply a linear trend over time or may initially exhibit a linear increasing/decreasing trend, followed by a transition period and then a linear decreasing/increasing trend. As far as the transition period is concerned, some individuals may exhibit a gradual transition, whereas for others an abrupt transition. Statistical techniques for this context revolve around broken-stick (piecewise linear) regression, typically ignoring temporal autocorrelation. Reynolds & Chiu (2010) provided a brief critique of the restrictiveness of an abrupt transition as imposed by the broken stick, and of the shortcomings of ignoring autocorrelation. They subsequently attempted to improve the statistical inference by applying the bent-cable time series model (Chiu & Lockhart, 2010). This approach is a time-series extension of frequentist bent-cable regression of Chiu et al. (2006). For both models, asymptotic theory was developed; the focus was to quantify the nature of the transition between approximately linear phases by modelling the transition as a quadratic phase with unknown, non-negative width. However, inference is only based on a single observational unit at a time. Moreover, the authors admitted that the segmented nature of the model leads to poor asymptotic approximation in many practical settings involving finite samples. This shortcoming partially caused Reynolds & Chiu (2010) to discard various observational units from their analysis. Khan et al. (2013) extended the single-profile bent-cable model for longitudinal data which allows the transition of each individual's underlying temporal trend to be either abrupt or gradual, through inference for whether the individual belongs to Population A or Population G. Thus, inference for transition type is data driven rather than pre-assumed, the latter being required by many changepoint analysis methods. A Bayesian framework was considered for statistical inference, so that the concern about unsatisfactory performance of bent-cable asymptotics is irrelevant.

In this manuscript, we develop a substantial extension of the Khan et al. (2013)'s bent-cable methodology for longitudinal data by taking into account samples potentially coming from three different populations: Populations G, A and L. We have tailored our work for the scientific context of the rat model as described in Section 1 to understand hypothermia therapy for human populations. Note that abruptness of change modeled by either a piecewise linear model or a smooth function is unrealistic as discussed in Khan et al. (2013). Instead, it would be more meaningful to distinguish the three types of trends as being samples from three different populations. The populations may also differ with respect to the rates of linear changes in the incoming and outgoing phases. As opposed to the Khan et al. (2013)'s bent-cable methodology, the extended bent-cable methodology presented in this article by taking into account the population-specific intercept and slope parameters in addition to the three types of transitions is expected to more appropriately characterize the populations under investigation. Moreover, the proposed hierarchical Wishart prior is extremely important to achieve robust inference, even when little is reliably known about the between-individual covariance structure. The performance of our methodology is evaluated

using extensive simulation study: our simulation study reveals that the proposed methodology performs well to allow proper inference at the population level by relaxing the assumption of universal abruptness, gradualness or linearity. Note that there is also penalized spline regression (Ruppert et al., 2003) to flexibly model the shape of the individual trends, though the flexible mixture bent-cable model is more appealing in the context described in this paper due to its parsimoniousness and greatly interpretable regression coefficients.

Our methodology provides a general framework to analyze longitudinal data that exhibit either a transitional trend (gradual or abrupt) between two approximately linear phases or simply a linear trend over time. No previous work has been done under a general regression modelling framework to classify observational units in the same longitudinal study as exhibiting either an abrupt or gradual transition or simply a linear trend. Such classification provides not only inference that is more realistic, but also insights into the underlying behaviour within a population. Aside from crucial information at the population level, another aspect of our longitudinal framework which researchers would find valuable is the inference for the temporal trend exhibited by individual observational units. Although our work is motivated by a case study in therapeutic hypothermia, our model can be applied to other longitudinal data that exhibit similar statistical characteristics as the rat study.

Some caution is required in using our proposed methodology for the following reasons: (1) the proposed methodology should be used if there is strong reason to believe that the sample potentially comes from the three populations mentioned above; Khan et al. (2013)'s bent-cable methodology with extensions (b) and (c) mentioned in Section 1 is the more appropriate technique when the sample potentially comes from populations G and A only, and Khan et al. (2009)'s bent-cable methodology with extension (c) of Section 1 is the appropriate technique when all the trajectories exhibit gradual transition, and (2) the proposed methodology is intended for longitudinal time-series data, and may not be efficient to identify the three types of trends for longitudinal data with a small number of repeated measures.

Acknowledgements

This work was partially supported by the University of Saskatchewan through New Faculty Start-up Operating Fund and President's NSERC grant to S. A. Khan. We thank Dr. P. S. Reynolds, Department of Emergency Medicine and Virginia Commonwealth University Reanimation Science Center, Virginia Commonwealth University Medical Center, Richmond, VA USA, for permitting us to use the rat data in this article. We also wish to express our sincere gratitude and appreciation to Professors Joel A. Dubin and Grace S. Chiu for their valuable comments and suggestions in preparing this manuscript.

References

- Bellera, C. A., Hanley, J. A., Joseph, L., & Albertsen, P. C. (2008). Hierarchical changepoint models for biochemical markers illustrated by tracking postradiotherapy prostate-specific antigen series in men with prostate cancer. *Annals of Epidemiology*, *18*, 270–282. <http://dx.doi.org/10.1016/j.annepidem.2007.10.006>
- Chiu, G. S., & Lockhart, R. A. (2010). Bent-cable regression with autoregressive noise. *The Canadian Journal of Statistics*, *38*, 386–407. <http://onlinelibrary.wiley.com/doi/10.1002/cjs.10070/pdf>
- Chiu, G., Lockhart, R., & Routledge, R. (2006). Bent-cable regression: theory and applications. *Journal of the American Statistical Association*, *101*, 542–553. <http://amstat.tandfonline.com/doi/abs/10.1198/016214505000001177>
- Connett, R. J., Pearce, F. J., & Drucker, W. R. (1986). Scaling of physiological responses: A new approach for hemorrhage shock. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, *250*, R951–R959. <http://ajpregu.physiology.org/content/ajpregu/250/6/R951.full.pdf>
- Daniels, M. J., & Kass, R. E. (1999). Nonconjugate Bayesian estimation of covariance matrices and its use in hierarchical models. *Journal of the American Statistical Association*, *94*, 1254–1263. <http://www.tandfonline.com/doi/abs/10.1080/01621459.1999.10473878>
- Davidian, M., & Giltinan, D. M. (1995). *Nonlinear models for repeated measurement data*. New York, NY: Chapman and Hall.
- Gordon, C. J. (2001). The therapeutic potential of regulated hypothermia. *Emergency Medical Journal*, *18*, 81–89. <http://dx.doi.org/10.1136/emj.18.2.81>
- Hall, C. B., Ying, J., Kuo, L., & Lipton, R. B. (2003). Bayesian and profile likelihood change point methods for modeling cognitive function over time. *Computational Statistics and Data Analysis*, *42*, 91–109.

- [http://dx.doi.org/10.1016/S0167-9473\(02\)00148-2](http://dx.doi.org/10.1016/S0167-9473(02)00148-2)
- Khan, S. A., Chiu, G., & Dubin, J. A. (2009). Atmospheric concentration of chlorofluorocarbons: Addressing the global concern with the longitudinal bent-cable model. *CHANCE*, 22, 8–17. <http://www.tandfonline.com/doi/abs/10.1080/09332480.2009.10722965>
- Khan, S. A., Chiu, G., & Dubin, J. A. (2013). Therapeutic hypothermia: Quantification of the transition of core body temperature using the flexible mixture bent-cable model for longitudinal data. *Australian & New Zealand Journal of Statistics*, 55, 369–385. <http://onlinelibrary.wiley.com/doi/10.1111/anzs.12047/full>
- Khan, S. A., Rana, M., Li, L., & Dubin, J. A. (2012). A statistical investigation to monitor and understand atmospheric CFC decline with the spatial-longitudinal bent-cable model. *International Journal of Statistics and Probability*, 1, 56–68. <http://dx.doi.org/10.5539/ijsp.v1n2p56>
- Kiuchi, A. S., Hartigan, J. A., Holford, T. R., Rubinstein, P., & Stevens, C. E. (1995). Change points in the series of T4 counts prior to aids. *Biometrics*, 51, 236–248. <http://dx.doi.org/10.2307/2533329>
- Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38, 963–974. <http://dx.doi.org/10.2307%2F2529876>
- Lange, N., Carlin, B. P., & Helfand, A. E. (1992). Hierarchical Bayes models for the progression of HIV infection using longitudinal CD4 T-cell numbers. *Journal of the American Statistical Association*, 87, 615–626. <http://www.tandfonline.com/doi/abs/10.1080/01621459.1992.10475258>
- Lemieux, C. (2009). *Monte Carlo and Quasi-Monte Carlo Sampling*. New York, NY: Springer.
- Plummer, M., Best, N., Cowles, K., Vines, K., Sarkar, D., & Almond, R. (2012). *coda: Output analysis and diagnostics for MCMC*. R Package Version 0.16-1. <http://cran.r-project.org/web/packages/coda/index.html>
- Previato, E. (Eds.) (2002). *Dictionary of applied math for engineers and scientists*. New York, NY: CRC Press.
- R Development Core Team. (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- Reynolds, P. S., Barbee, R. W., Skafien, M. S., & Ward, K. R. (2007). Low-volume resuscitation cocktail extends survival after severe hemorrhagic shock. *Shock*, 28, 45–52. <http://www.readcube.com/articles/10.1097/shk.0b013e31802eb779>
- Reynolds, P. S., & Chiu, G. S. (2010). Understanding thermoregulatory transitions during haemorrhage by piecewise regression. arXiv:1006.5117v1. <http://arxiv.org/abs/1006.5117>
- Rincon, F., & Mayer, S. A. (2006). Therapeutic hypothermia for brain injury after cardiac arrest. *Seminars in Neurology*, 26, 387–395. <https://www.thieme-connect.com/DOI/DOI?10.1055/s-2006-948319>
- Robert, C. P. (1996). Mixtures of distributions: Inference and estimation. In Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (Eds.) *Markov Chain Monte Carlo in practice*. New York, NY: Chapman and Hall/CRC, 441–464.
- Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric Regression*. New York, NY: Cambridge University Press.
- Slate, E. H., & Turnbull, B. W. (2000). Statistical Models for Longitudinal Biomarkers of Disease Onset. *Statistics in Medicine*, 19, 617–637. [http://dx.doi.org/10.1002%2f\(SICI\)1097-0258\(20000229\)19%3a4%3c617%3a%3aAID-SIM360%3e3.0.CO%3b2-R](http://dx.doi.org/10.1002%2f(SICI)1097-0258(20000229)19%3a4%3c617%3a%3aAID-SIM360%3e3.0.CO%3b2-R)
- Smith, A. F. M., & Roberts, G. O. (1993). Bayesian computation via Gibbs sampler and related Markov Chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B*, 55, 3–23. <http://dx.doi.org/10.1080%2F01621459.1990.10476213>
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society: Series B*, 64, 583–616. <http://onlinelibrary.wiley.com/doi/10.1111/1467-9868.00353/full>
- Wakefield, J. C., Smith, A. F. M., Racine-Poon, A., & Gelfand, A. E. (1994). Bayesian analysis of linear and non-linear population models by using the Gibbs sampler. *Applied Statistics*, 43, 201–221. <http://dx.doi.org/10.2307/2986121>

Appendix: Full Conditional Distributions

In this section, we present the full conditional distributions of the parameters. Let

$$\mathbb{X}_j = \begin{pmatrix} 1 & t_{j1} & q(t_{j1}, \alpha_j) \\ 1 & t_{j2} & q(t_{j2}, \alpha_j) \\ \vdots & \vdots & \vdots \\ 1 & t_{jn_j} & q(t_{jn_j}, \alpha_j) \end{pmatrix}.$$

Then, the full conditional distributions of the the model parameters can be expressed as follows.

- $[\beta_j|\cdot] = N_3(\mathbb{M}_j[\sigma_j^{-2}\mathbb{X}'_j\mathbf{y}_j + \Sigma^{-1}(\sum_{i=1}^3 z_{ij}\boldsymbol{\mu}_i)], \mathbb{M}_j)$, where $\mathbb{M}_j^{-1} = \sigma_j^{-2}\mathbb{X}'_j\mathbb{X}_j + \Sigma^{-1}$.
- $\pi(\alpha_j, \mathbf{z}_j|\cdot) \propto \exp\left\{-\frac{1}{2\sigma_j^2}(\mathbf{y}_j - \mathbb{X}'_j\boldsymbol{\beta}_j)'(\mathbf{y}_j - \mathbb{X}'_j\boldsymbol{\beta}_j)\right\}$
 $\times [p_1 \times LN_2(\alpha_j|\xi, \Psi) \times N_3(\boldsymbol{\beta}_j|\boldsymbol{\mu}_1, \Sigma)]^{z_{1j}}$
 $\times [p_2 \times LN(\tau_j|\eta, \sigma_\eta^2) \times N_3(\boldsymbol{\beta}_j|\boldsymbol{\mu}_2, \Sigma)]^{z_{2j}}$
 $\times [p_3 \times N_3(\boldsymbol{\beta}_j|\boldsymbol{\mu}_3, \Sigma)]^{z_{3j}}$.
- $[\sigma_j^{-2}|\cdot] = G\left(\frac{n_j+d_0}{2}, \frac{(\mathbf{y}_j - \mathbb{X}'_j\boldsymbol{\beta}_j)'(\mathbf{y}_j - \mathbb{X}'_j\boldsymbol{\beta}_j) + d_1}{2}\right)$.
- $[\boldsymbol{\mu}_i|\cdot] = N_3(\mathbb{U}_i[\Sigma^{-1}\boldsymbol{\lambda}_i + \mathbb{H}_i^{-1}\mathbf{h}_i], \mathbb{U}_i)$ for $i = 1, 2, 3$, where $\mathbb{U}_i^{-1} = m_i\Sigma^{-1} + \mathbb{H}_i^{-1}$, $m_i = \sum_{j=1}^m z_{ij}$ and $\boldsymbol{\lambda}_i = \sum_{j=1}^m z_{ij}\boldsymbol{\beta}_j$.
- $[\xi|\cdot] = N_2(\mathbb{V}[\Psi^{-1}\zeta + \mathbb{H}_\xi^{-1}\mathbf{h}_\xi], \mathbb{V})$, where $\mathbb{V}^{-1} = m_1\Psi^{-1} + \mathbb{H}_\xi^{-1}$, $m_1 = \sum_{j=1}^m z_{1j}$ and $\zeta = \sum_{j=1}^m z_{1j} \log \alpha_j$.
- $[\eta|\cdot] = N\left(\frac{\sigma_\eta^{-2}\kappa + a_1^{-1}a_0}{m_2\sigma_\eta^{-2} + a_1^{-1}}, \frac{1}{m_2\sigma_\eta^{-2} + a_1^{-1}}\right)$, where $\kappa = \sum_{j=1}^m z_{2j} \log \tau_j$ and $m_2 = \sum_{j=1}^m z_{2j}$.
- $[\Sigma^{-1}|\cdot] = W_3\left(m + \nu_1, \left[\sum_{i=1}^3 \sum_{j=1}^m z_{ij}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_i)(\boldsymbol{\beta}_j - \boldsymbol{\mu}_i)' + \nu_1\mathbb{A}_1\right]^{-1}\right)$.
- $[\Psi^{-1}|\cdot] = W_2\left(m_1 + \nu_2, \left[\sum_{j=1}^m z_{1j}(\log \alpha_j - \xi)(\log \alpha_j - \xi)' + \nu_2\mathbb{A}_2\right]^{-1}\right)$.
- $[\sigma_\eta^{-2}|\cdot] = G\left(\frac{m_2 + e_0}{2}, \frac{\sum_{j=1}^m z_{2j}(\log \tau_j - \eta)^2 + e_1}{2}\right)$.
- $[(\mathbb{A}_1)_{ll}|\cdot] = G\left(\frac{\nu_1+2}{2}, \frac{\nu_1(\Sigma^{-1})_{ll}}{2}\right)\mathbf{1}_{\{b_{0l}, b_{1l}\}}$, for $l = 1, 2, 3$.
- $[(\mathbb{A}_2)_{ll}|\cdot] = G\left(\frac{\nu_2+2}{2}, \frac{\nu_2(\Psi^{-1})_{ll}}{2}\right)\mathbf{1}_{\{c_{0l}, c_{1l}\}}$, for $l = 1, 2$.
- $[p_1, p_2, p_3|\cdot] = Dirichlet(m_1 + \omega_1, m_2 + \omega_2, m_3 + \omega_3)$.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).