

Random Power Function of Bent-cable Growth Models for Longitudinal Data: Application to AIDS Studies

Getachew A. Dagne¹

Abstract

This paper presents a power function within a bent-cable model for identifying a transition period for the development of drug resistance to antiretroviral treatment (ART) in HIV/AIDS studies or other related fields. Such a transition period usually occurs between incoming and outgoing phases of trajectories of longitudinal data. The transition curve is modeled using a power function with unknown width, and we assess this clinically important feature of the longitudinal HIV/AIDS data using the bent-cable framework with random subject-specific parameters, including parameters of the transition curve. We develop a fully Bayesian approach for fitting a power function of bent-cable Tobit model to left-censored data and illustrate the proposed methods on real data from an AIDS clinical study.

Keywords: Bayesian inference; change-points; longitudinal study; mixed-effects models; skew distribution

1. Introduction

Very often longitudinal studies produce outcome variables whose trajectories may exhibit several phasic changes over time. Characterizing and identifying those phases of trajectories is important in clinical practice. For example, in longitudinal clinical studies of the efficacy of antiretroviral (ARV) drugs, typical trajectories of viral load (number of HIV RNA copies in plasma) of patients, after receiving ARV, initially start declining (phase 1) then followed by a gradual transition period (phase 2), and finally a third phase of upward (rebound), predicting progression to AIDS (Paterson et al., 2000; Dagne, 2018). These different phases may occur because of a biological process of developing resistance to ARV drugs and non-adherence over time (Paterson et al., 2000). To describe the phasic patterns, one can use a bent-cable model (Chiu et al., 2006). The transition phase in the regular bent-cable model is modeled using a quadratic function which may be too restrictive for adequately describe it. In this paper, we develop a bent-cable model by jointly (i) relaxing the assumption of the quadratic function for the transition period, (ii) allowing subject-to-subject variability in the parameters of the proposed model, (iii) incorporating skewness and detection limits in a distribution of viral loads (Wu and Zhang, 2002), and (iv) adjusting for measurement errors in covariates.

In assessing subject-specific growth curves (trajectories) of a longitudinal response, the features of the data determine the types of models to be adopted. In cases where there are abrupt changes in trajectories, piecewise or "broken-stick" models are usually effective (Kiuchi et al., 2001; Slate et al., 2001; Hall et al., 2003; Zhao et al., 2015). In other cases where there are smooth and gradual transitions between incoming and outgoing phases, a power function of bent-cable model is preferred (Khan and Kar, 2018) instead of a piecewise model. For instance, in the case of HIV/AIDS studies, growth curves of longitudinal data may show gradual transition periods between an initial decline in viral load after treatment and rebound at later time in the study. These growth curves also exhibit subject-to-subject variability, and assuming fixed parameters instead of random may not reflect the biological reality. To properly analyze such data, we use a random power function of bent-cable model (see Section 2 for details) which also incorporates statistical adjustments for measurement error in covariates and left-censoring due to lower detection limit (LOD).

¹ Department of Epidemiology & Biostatistics, College of Public Health, MDC 56, University of South Florida, Tampa, FL 33612; Email: gdayne@health.usf.edu, phone: 813-974-6680.

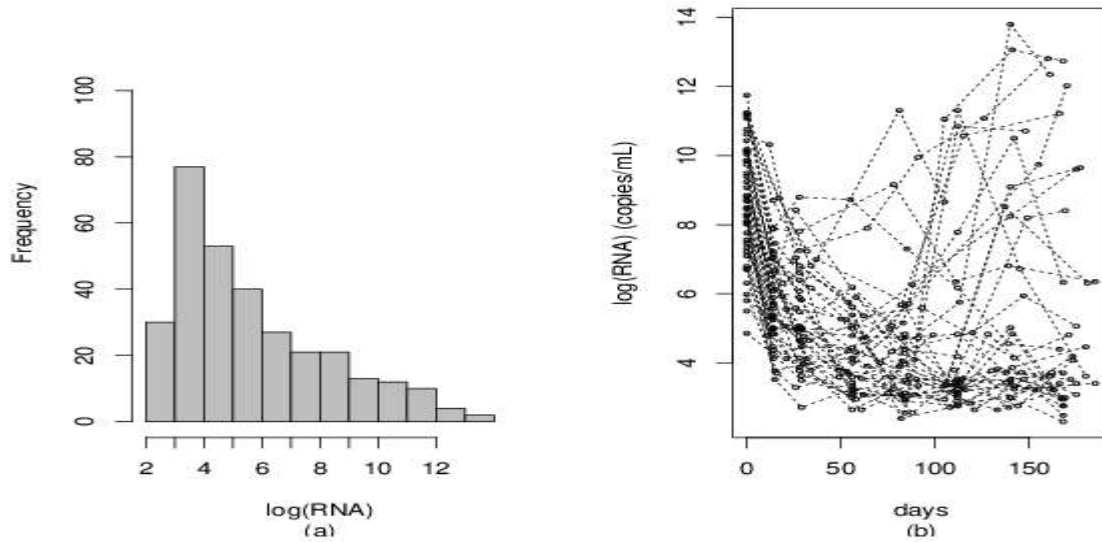


Fig. 1: (a) Histogram of log viral load; (b) spaghetti plot of log viral load

Often viral load data suffer from left-censoring and skewness. To account for left-censoring within the power function of bent-cable model, we jointly use the Tobit model (Tobin, 1958) which provides consistent estimates of parameters. The commonly used Tobit model assumes normality which may not be suitable for a highly skewed distribution of viral load measurements (Dagne, 2018). For example, Figure 1(a) shows the histogram of viral loads for patients enrolled in the AIDS clinical trial study--A5055 (Acosta et al., 2004). From the histogram one can observe that the distribution of log-transformed viral load data exhibit right skewness. In this paper, we use skew-elliptical distributions (Sahu et al., 2003; Genton, 2004), where multivariate skew-normal (SN) and multivariate skew-t (ST) distributions are special cases, for modeling left-censored HIV-RNA data with phasic patterns. In Section 2, we develop a power function of bent-cable Tobit models with multivariate ST distributions in full generality. In Section 3, we present the Bayesian inferential procedure. The proposed methodologies are illustrated using the AIDS data set in Section 4. Finally, the paper concludes with discussions in Section 5.

2. Power Function of Bent-cable Growth Models

2.1. Motivating data

The motivating data for our paper were obtained from an AIDS clinical trial study (A5055) (Acosta et al., 2004). Forty four HIV-1 infected patients were treated with a potent ARV regimen with a combined use of a protease inhibitor, nelfinavir (2,250 mg/d), and transcriptase inhibitors, zidovudine (600 mg/d) and lamivudine (300 mg/d). The outcome variable was RNA viral load (copies/mL), recorded at 0, 7, 14, 28, 56, 84, 112, 140 and 168 days of follow-up. CD4 cell counts were also measured. The number of observations varies from patient to patient with a median of 8 measurements and a standard deviation of 1.49. Figure 1(b) depicts the trajectories of log viral load over time, showing some phasic patterns. These patterns comprise an initial decreasing trend, a gradual transition part and an increasing upward trend for a typical patient. Modeling these patterns has clinical importance for assessing the effectiveness of ARV drugs and AIDS progression. In the modeling process, we also take into account left-censoring of viral load (detectable lower limit is 50 copies/mL) by using a Tobit model which is discussed in the next section.

2.2. Model specification

We now develop a power function of bent-cable Tobit model (PBTM) which simultaneously incorporates left-censoring, skewness and measurement errors in covariates. The outcome variable for our study is denoted by y_{ij}^* , and z_{ij}^* be a true covariate for the individual i at time t_{ij} ($i=1,2,\dots,n; j=1,2,\dots,n_i$). Let $c_{ij} = 0$ if y_{ij}^* is observed (above the lower detection limit (LOD)) and $c_{ij} = 1$ if y_{ij}^* is left-censored (below LOD). Some of the observed values of the outcome variable may be recorded as being equal to or below LOD due to left-censoring, making them unreliable. To distinguish these measurements, we define an indicator $c_{ij} = 0$ if y_{ij}^* is observed and $c_{ij} = 1$ if y_{ij}^* is left-censored. A Tobit model (Tobin, 1958) is adopted for modeling such left-censored longitudinal data as follows. $y_{ij} = y_{ij}^*$ if $y_{ij}^* > \rho$ or $y_{ij} = \text{missing}$ if $y_{ij}^* > \rho$ where ρ is a non-stochastic LOD, which in our example is equivalent to $\log(50)$ (1).

The main focus of PBTM is on modeling longitudinal data with multiple phasic patterns based on an ST distribution for y_{ij} as follows

$$y_{ij} = g(t_{ij}, \mu_{ij}, z_{ij}^*) + e_{ij} \quad e_i \sim ST_{n_i, \nu_e}(0, \sigma_e^2 I_{n_i}, \Delta(\delta_{e_i})) \quad (2)$$

Where μ_{ij} is the mean structure, and the error $e_i = (e_{i1}, \dots, e_{in_i})'$ follows a multivariate ST distribution with degrees of freedom ν_e , variance σ_e^2 and $n_i \times n_i$ skewness diagonal matrix $\Delta(\delta_{e_i}) = \delta_e I_{n_i}$.

The mean structure μ_{ij} in (2) is a power function of bent-cable model to assess multiple phases of trajectories such as: (i) incoming trend, (ii) outgoing trend, and (iii) transition power function that connects the incoming and outgoing trends. The power function, where a quadratic function is a special case (Chiu et al., 2006; Khan et al., 2009), represents the gradual transition period between two unknown time points T_{1i} and T_{2i} . Following Khan and Kar (2018), suppose $T_{1i} = \phi_{1i} - \phi_{2i}(\phi_{3i} - 1)$ and $T_{2i} = \phi_{1i} + \phi_{2i}$, such that $T_{1i} < \phi_{1i} < T_{2i}$ and $\phi_{1i}, \phi_{2i} > 0$, and $\phi_{3i} > 1$. These parameters can also be considered as random coefficients in the mean structure of PBTM (Tishler et al., 1981; Chiu et al., 2006) as follows.

The mean μ_{ij} is modeled as

$$\mu_{ij} = \eta z_{ij}^* + \beta_{1i} + \beta_{2i} t_{ij} + \beta_{3i} q(t_{ij}, \kappa_i) \quad (3)$$

where η measures the effect of time-varying covariate; $q(t_{ij}, \kappa_i) = \phi_{2i}(t_{ij} - \phi_{1i}(\phi_{3i} - 1)) / (\phi_{2i}\phi_{3i})^{\phi_{3i}} I(\phi_{1i} - \phi_{2i}(\phi_{3i} - 1) < t_{ij} \leq \phi_{1i} + \phi_{2i}) + (t_{ij} - \phi_{1i}) I(t_{ij} > \phi_{1i} + \phi_{2i})$ $I(\cdot)$ is an indicator function; β_{1i} and β_{2i} are the incoming random intercept and slope, respectively; β_{3i} is the change in slope between the incoming and outgoing linear phases; $\kappa_i = (\phi_{1i}, \phi_{2i}, \phi_{3i})'$ are the parameters of the power function, $q(t_{ij}, \kappa_i)$, which presents the transition phase. Note that the power function of the bent-cable model becomes an ordinary bent-cable when $\phi_{3i} = 2$ and a broken-stick model when $\phi_{2i} = 0$, $\phi_{3i} > 1$ or $\phi_{2i} > 0$, $\phi_{3i} = 1$.

The random parameters in (3) are modeled as

$$\begin{aligned} \beta_{1i} &= \beta_1 + b_{1i}, \quad \beta_{2i} = \beta_2 + b_{2i}, \quad \beta_{3i} = \beta_3 + b_{3i} \\ \phi_{1i} &= \phi_1 + b_{4i}, \quad \phi_{2i} = \phi_2 + b_{5i}, \quad \phi_{3i} = \phi_3 + b_{6i} \end{aligned} \quad (4)$$

Where $\beta = (\beta_1, \beta_2, \beta_3, \phi_1, \phi_2, \phi_3)'$ are population-based parameters. The random effects $(b_{1i}, \dots, b_{6i})'$ have a multivariate normal distribution $N(0, \Sigma_b)$ where Σ_b is a variance-covariance matrix with dimension of 6. The critical time point, where the mean trajectory shifts from a decreasing trend to an upward trend, is located at $\chi = \phi_1(\phi_3 - 1)\phi_2 + \left[\left(\frac{\beta_2}{\beta_3} \right) (\phi_2\phi_3)^{\phi_2 - 1} \right]^{\phi_3 - 1}$ (Khan and Kar, 2018). In the estimation process of these parameters of the power function of bent-cable Tobit model in (3), measurement errors in covariates (e.g., CD4) (Raboud et al., 1996) are also addressed. Next, we present a covariate measurement error model.

2.3. Covariate measurement error models

For explaining some of the variations of the outcome variable of PBTM, time-varying covariates such as CD4 cell counts are used. It is often the case that CD4 cell counts are measured with errors (Higgins et al., 1997; Wu, 2002; Carroll et al., 2006). To incorporate measurement errors, we adopt a polynomial mixed-effects model as follows.

$$z_{ij} = u_{ij} \alpha + v_{ij} a_i + \epsilon_{ij}, \quad \epsilon_i \sim ST_{n_i, \nu_z}(0, \sigma_z^2 I_{n_i}, \Delta(\delta_{\epsilon_i})) \quad (5)$$

Where z_{ij} is the observed CD4 cell count with measurement errors for individual i

at time t_{ij} and let $z_{ij}^* = u_{ij} \alpha + v_{ij} a_i$ be the true CD4 cell count at time t_{ij} ,

u_{ij} and v_{ij} are $1 \times l$ design vectors, $\alpha = (\alpha_1, \dots, \alpha_l)'$ and $a_i = (a_{i1}, \dots, a_{il})'$ are fixed-effects and random-effects parameter vectors, respectively. We assume that $a_i \sim N_l(0, \Sigma_a)$, where

Σ_a is a covariance matrix. Here, we consider the model in (5) with $u_{ij} = v_{ij} = (1, t_{ij}, \dots, t_{ij}^{l-1})'$, and AIC and BIC are used to choose the best model.

3. Bayesian estimation

The model parameters of models in (2) and (5) are estimated using a Bayesian approach. To facilitate the estimation, we use the stochastic representation for the ST distribution by introducing an $n_i \times 1$ random variable vector $w_i = (w_{i1}, \dots, w_{in_i})'$ and ξ_{e_i} as a scaling weight (Sahu et al., 2003; Dagne, 2017). Such a representation makes the use of Bayesian algorithms simpler as given below.

$$y_{ij} \sim N\left(g(t_{ij}, \mu_{ij}, z_{ij}^*) + \delta_{e_{ij}} w_{e_{ij}}, \xi_{e_i}^{-1} \sigma_e^2\right), \quad w_{e_{ij}} \sim N(0,1)I(w_{e_{ij}} > 0), \quad \xi_{e_{ij}} \sim G\left(\frac{\nu_e}{2}, \frac{\nu_e}{2}\right)$$

$$z_{ij} \sim N\left(z_{ij}^* + \delta_{e_{ij}} w_{e_{ij}}, \sigma_z^2\right), \quad a_i \sim N_l(0, \Sigma_a), \quad b_i \sim N_6(0, \Sigma_b) \tag{6}$$

where $G(\cdot)$ is a gamma distribution, $I(w_{e_{ij}} > 0)$ is an indicator with $w_{e_{ij}} \sim N(0,1)$ and $w_{e_{ij}} > 0$. z_{ij}^* is the true but unobservable covariate value at time t_{ij} . The ST distribution (6) can be reduced to the following three special cases: (i) a skew-normal (SN) distribution as $\nu_e \rightarrow \infty$ and $\xi_{e_i} \rightarrow 1$ with probability 1, (ii) a standard t-distribution as $\delta_{e_{ij}} = 0$, or (iii) a standard normal distribution as $\nu_e \rightarrow \infty$ and $\delta_{e_{ij}} = 0$.

Let $\theta = (\alpha, \beta, \eta, \sigma_z^2, \sigma_e^2, \Sigma_a, \Sigma_b, \nu_e, \Delta_{e_i}, i = 1, \dots, n)$ be the unknown parameters in models (2, 5) with the joint prior distribution $g(\theta)$ where assessments of specific prior distributions are given in Section 4.

Statistical inference for the parameters of the proposed models is carried out using their posterior distributions, which are obtained by combining the likelihood of the observed data and the prior distributions via the Bayes' theorem. Let $f(\cdot|\cdot)$ and $F(\cdot|\cdot)$ denote a probability distribution and cumulative distribution, respectively. For a left-censored response, a detectable observed value y_{ij} contributes $f(y_{ij})$, while a non-detectable value contributes $F(\rho)$ in the likelihood. Based on this formulation of the likelihood, the joint posterior density of θ based on the observed data $D = (y_i, z_i, i = 1, \dots, n)$ can be given by

$$f(\theta|D) \propto g(\theta) \prod_i^n \prod_j^{n_i} \int \int [1 - \Pr(S_{ij}=1)] f(y_{ij} - c_{ij} | 1 - \Pr(S_{ij}=1) + \Pr(S_{ij}=1) F(\rho_{cij} | z_{ij} w_{e_{ij}} w_{e_{ij}} > 0) f(w_{e_{ij}} w_{e_{ij}} > 0) f(a_i) f(b_i) f(\Delta_{e_i}) \tag{7}$$

where $\Pr(S_{ij} = 1)$ is the probability that a patient's outcome is left-censored. The posterior function in (7) involves integrals of high dimensions and getting closed form solutions is thus difficult. Alternatively, we can use the Monte Carlo Markov chain (MCMC) algorithm to estimate the parameters in (7).

4. Application to HIV/AIDS data

4.1. Specification of models

We now demonstrate the proposed models by fitting them to a real data set from an AIDS clinical study (Acosta et al., 2004). We fit PBTM given in (2-4) to the the natural logarithm of HIV viral load for the i th subject at time t_{ij} ($i=1, \dots, n=44, j=1, \dots, n_i$), and here z_{ij}^* represents the true value of CD4. To estimate the true values of CD4 cell count, a linear mixed-effects model is adopted as given in (5). In particular, we set $u_{ij} = v_{ij} = (1, t_{ij}, \dots, t_{ij}^{l-1})'$, taking linear ($l=2$), quadratic ($l=3$) and cubic ($l=4$) polynomials with corresponding AIC (BIC) values of 799.03 (821.74), 703.56 (744.42) and 766.18 (782.08), respectively. Based on these model criteria values, the best model is a quadratic mixed-effects model for the CD4 process (Dagne, 2019).

$$z_{ij} = (\alpha_1 + a_{i1}) + (\alpha_2 + a_{i2})t_{ij} + (\alpha_3 + a_{i3})t_{ij}^2 + \epsilon_{ij}, \tag{8}$$

Where $z_{ij}^* = (\alpha_1 + a_{i1}) + (\alpha_2 + a_{i2})t_{ij} + (\alpha_3 + a_{i3})t_{ij}^2$, $\alpha = (\alpha_1, \dots, \alpha_3)'$ is a vector of population (fixed-effects) parameters, $\epsilon_i = (\epsilon_{i1}, \dots, \epsilon_{in_i})' \sim ST_{n_i, \nu_z}(0, \sigma_z^2 I_{n_i}, \delta_\epsilon I_{n_i})$ and individual-specific random-effects $a_i = (a_{i1}, \dots, a_{i3})' \sim N_3(0, \Sigma_a)$. For numerical stability, we standardized the time-varying covariate CD4 cell counts.

To describe the response process with left-censoring and multiple phases of trajectory of viral load, we provide three competing models. These are (i) a power function of bent-cable Tobit model with normal distributions of random errors (Model I); (ii) a power function of bent-cable Tobit model with skew-normal distributions of random errors (Model II); and

(iii) a power function of bent-cable Tobit model with skew-t distributions of random errors (Model III). In all these models, we assess how the true time-varying covariates, CD4 cell count, would affect the multiple phases of trajectories.

For fitting these models using a Bayesian approach, the hyperparameters of the prior distributions are assessed as follows. (i) Parameters for fixed-effects are assumed to have a normal distribution $N(0, 100)$ for each element of α , β , and η . (ii) For σ_e^2 and σ_z^2 we assume inverse gamma prior distributions $IG(0.01, 0.01)$. (iii) For the variance matrices of the random-effects Σ_a and Σ_b , we assume inverse Wishart distributions $IW(\Omega_1, \rho_1)$ and $IW(\Omega_2, \rho_2)$ with covariance matrices $\Omega_1 = \text{diag}(.01)$, $\Omega_2 = \text{diag}(.01)$ and $\rho_1 = 3$ and $\rho_2 = 6$, respectively. (iv) The parameter for the degrees of freedom ν_e , it follows a gamma distribution $G(1.0, .01)$. (v) For the skewness parameter δ_e , a normal distribution $N(0, 100)$ is considered.

After applying the Bayes' theorem to the likelihood function and the prior distributions specified above, Bayesian estimation of parameters using the MCMC algorithm is straightforward. Checking the convergence of the MCMC algorithm was carried out using the tools within the WinBUGS software (Lunn et al., 2000). First, to see how well the MCMC generated chains were mixing (showing convergence), we inspected trace plots of the iteration number against the values of the draw of parameters at each iteration. Upon inspection and taking in consideration of the complexity of the power function of bent-cable models, some generated values for some parameters took longer iterations to converge. Convergence was getting better after 100,000 iterations, and thus the first 100,000 iterations were ignored as burn-in. Second, the computed autocorrelations were small after using a thinning of 50, suggesting a good mixing. Third, the Markov chain (MC) errors were less than 5% of posterior standard deviation values for the parameters, showing good precision and convergence of MCMC (Ntzoufras, 2009). Fourth, the Gelman-Rubin potential scale reduction factors were close to 1 with a range of 1-1.2, confirming convergence. Finally, we retained 10,000 MCMC samples for the purpose of making posterior inference of the unknown parameters of interest. The computational burden to run MCMC for Model II was close to 6 hours on Latitude E5540, Intel(R) Core(TM) i7-4600U @2.10GHz.

4.2. Results of model fit

4.2.1. Model comparison

For choosing the best model out of the three models mentioned above, the values of the Bayesian model selection criteria are presented in Table 1. Accordingly, Model I has the biggest values of deviance (1254.0), RSS (230.649), EPD (2.577), and DIC (1110.820) the second largest values of deviance (988.8), RSS (4.308), EPD (.327), and DIC (822.071) belong to Model III. Model II has the least values of deviance (919.5), RSS (2.108), EPD (.326) and of DIC (631.709), indicating that the skew-normal gives the best fit.

Table 1. Model comparison using Bayesian model selection criteria (RSS=residual sum of squares, EPD= expected predictive deviance, DIC= deviance information criterion)

Criterion	Model I	Model II	Model III
Deviance	1254.0	919.5	988.8
RSS	230.649	2.108	4.308
EPD	2.647	0.326	0.327
DIC	1110.820	631.709	822.071

Results of indicators of goodness-of-fit are depicted in Figure 2 to supplement the findings given above. When we look at the plots of the observed values versus the fitted values in the second column of Figure 2, we see that Model II provides better fit to the observed data as compared to either Model I or Model III. Similarly, in the right panel, one can see that the Q-Q plot for Model II gives a better fit to the data as compared to either Model I or Model III. All the results suggest that Model II is the best model for fitting the HIV/AIDS viral load data which has multiple phasic patterns, skewness and left-censoring.

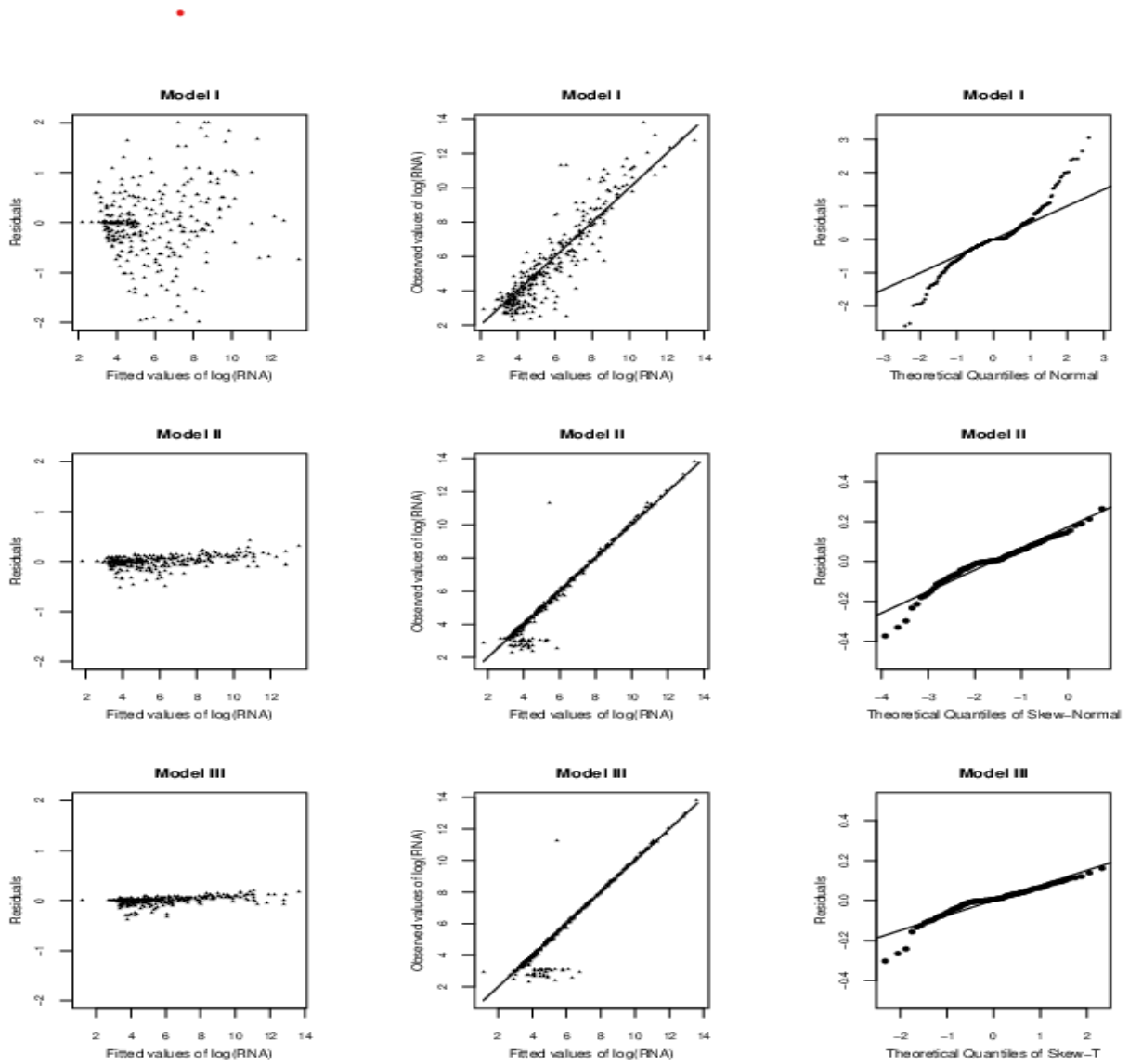


Fig. 2. Plots of goodness-of-fit statistics for three Models: (i) The first column has the residual plots against the fitted values, (ii) the second has plots of the fitted against observed $\log(\text{RNA})$ values, and (iii) the last column has the Q-Q plots of residuals.

4.2.2. Bayesian inference under Model II

The best model, Model II, describes well the phasic patterns of viral trajectories by using a power function of the bent-cable with a skew-normal distribution for the error terms. The posterior values of the parameters of Model II are reported in Table 2, specifically in the middle row. The posterior means for the intercept and slope parameters of the incoming trend are $\widehat{\beta}_1 = 4.102$ and $\widehat{\beta}_2 = -1.27$, respectively; for β_3 , which represents the change in slope between the incoming trend and outgoing trend, the posterior mean is $\widehat{\beta}_3 = 1.621$. For the parameters of the power function which represents the transition phase, the posterior means are $\widehat{\phi}_1 = 6.886$, $\widehat{\phi}_2 = 8.958$, $\widehat{\phi}_3 = 2.207$. The fitted power function of the bent-cable curve is displayed in Figure 3.

From the graph we observe that patients, after receiving the ARV treatment, encounter a decrease in viral load followed by a gradual transition period up until $\widehat{\phi}_1 + \widehat{\phi}_2 = 15.84$ weeks. This means that after receiving a treatment the decrease in viral load production is stronger at the start of the study but does not last long.

Right after approximately 16 weeks in the study (the rightdotted line in Figure 3, patients encountered a linear increase at the rate of 1.621 per week. Also, the time point at which the mean trajectory of viral load turns to an upward direction from a decreasing trend is estimated to be 12.225 weeks after the initiation of treatment. In summary, after receiving a treatment, the estimated mean growth trajectory of viral load suggests that the effect of the ARV treatment declines over time by not providing protection against the production of HIV virus. This explains why we see the upward trend towards the end of the study. For Model II, the posterior mean of the variance parameter σ_e^2 , .164, is relatively small implying that taking into account skewness is beneficial. The posterior mean of the skewness parameter (δ_e) is 1.81 with a 95% credible interval of (1.257, 2.508), confirming that the use of skewed distributions is very useful. The effect of CD4 on viral load is negative and strong ($\hat{\eta} = -.893$, 95% CI: (-1.337,-.485)), suggesting that as CD4 cell count gets lower, viral load rebounds.

Table 2. A summary of the posterior means (PM) of population parameters along with the corresponding lower limit (L_{CI}) and upper limit (U_{CI}) of 95% equal-tail credible interval.

Model		β_1	β_2	β_3	η	ϕ_1	ϕ_2	ϕ_3	α_1	α_2	α_3	σ_e^2	σ_z^2	δ_e
I	PM	7.136	-	1.751	-.912	2.343	4.405	2.286	-.903	.623	-	1.328	.060	--
	L_{CI}	4.312	-	.913	-	.098	1.667	1.447	-	.110	-	1.039	.024	--
	U_{CI}	8.524	-.814	3.306	-.487	4.748	6.93	2.965	-.563	1.124	.3162	1.678	.099	--
II	PM	4.102	-1.27	1.621	-.893	6.886	8.958	2.207	-.929	.630	-.257	.164	.056	1.81
	L_{CI}	1.546	-	.604	-	.292	2.883	1.137	-	.092	-	.006	.019	1.257
	U_{CI}	5.973	-.293	2.581	-.485	31.13	25.04	2.97	-.609	1.155	.315	.676	.095	2.508
III	PM	5.036	2.239	-	-.900	.097	.544	1.07	-.699	.778	-	.153	.072	2.312
	L_{CI}	4.193	1.453	-	-	.003	.021	1.002	-	.246	-	.006	.031	1.796
	U_{CI}	5.849	2.95	-	-.400	.362	1.677	1.25	.406	1.319	.160	.744	.122	2.779

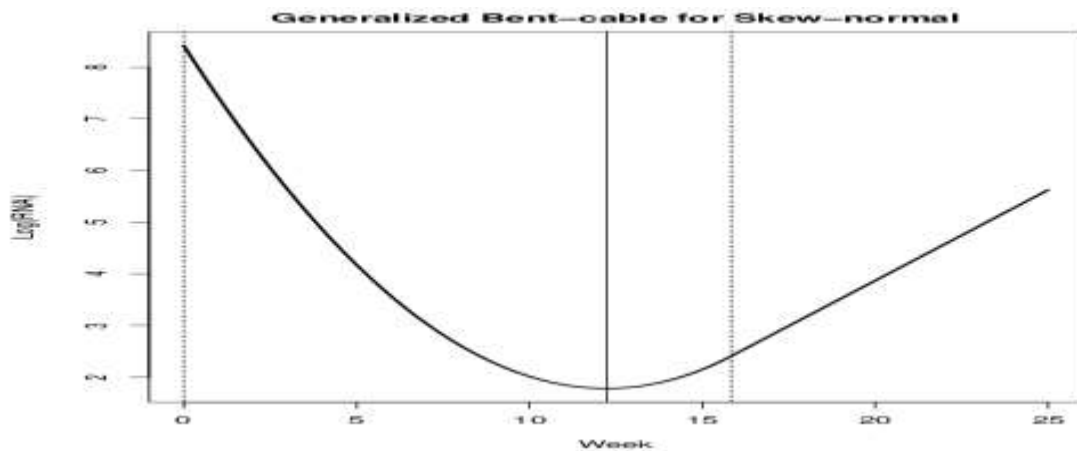


Fig.3 Fitted power function of bent-cable growth curve for log(viral load). The dotted vertical lines represent the lower ($\hat{\phi}_1 - \hat{\phi}_2(\hat{\phi}_3 - 1)$, truncated at zero) and upper ($\hat{\phi}_1 + \hat{\phi}_2 = 15.84$) limits of the gradual transition period; an estimated critical time point is 12.225 weeks (solid line).

For the CD4 covariate measurement error model (8), the posterior means are also in Table 2.

In the middle of Table 2 under Model II, the posterior means of the parameters for the linear and quadratic components are $\widehat{\alpha}_2 = .630$ and $\widehat{\alpha}_3 = -.257$, respectively. The linear slope is positive and significant since the 95% credible interval does not include zero while 95% credible interval for the quadratic coefficient contains zero, showing a weak relation. This result shows that there is a strong and positive linear relation between CD4 cell count and time.

5. Conclusion

For modeling phasic patterns of longitudinal data, this paper presents a power function of the bent-cable growth model (PBTM). Specifically, the power function models the transition period between incoming and outgoing segments of growth trajectories as a generalization of a quadratic function, providing a precise estimation of the time at which drug resistance may develop or rebound. The PBTM under skew-normal (Model II) is selected as the 'best' fit to the HIV/AIDS data, and the main findings suggest that, after the initiation of the ARV treatment, a significant decrease in viral load production is followed by a gradual transition period upward until 15.84 weeks. Furthermore, this proposed method (PBTM) also accounts for a skewness and left-censoring in the outcome and a measurement error in covariates.

In modeling left-censored viral loads, we use the Bayesian approach to predict them as missing values instead of using a substitution method such as LOD/2 or LOD (Hornung and Reed, 1990). Paxton et al. (1997) and Hughes (1999) stated that employing a substitution method for left-censored data often leads to biased estimators. In the case of right-censored data, our method can also be applied in a straightforward way. Thus, based on the Bayesian methodology, PBTM with skew distributions easily handles both censored and uncensored data as demonstrated in this paper.

In conclusion, we have demonstrated the application of flexible skew-distributions in the framework of a Bayesian PBTM for jointly analyzing left-censored and skewed longitudinal data with phasic changes. The proposed methods have a promising potential to better understand the features of trajectories of longitudinal data and measurement errors in covariates.

6. References

- Acosta EP, Wu H, Walawander A, Eron J, Pettinelli C, Yu S, Neath D, Ferguson E, Saah AJ, Kuritzkes DR, Gerber JG, for the Adult ACTG 5055 Protocol Team. (2004). Comparison of two indinavir/ritonavir regimens in treatment-experienced HIV-infected individuals. *Journal of Acquired Immune Deficiency Syndromes*, 37, 1358-1366.
- Arellano-Valle, R.B., Genton, M.G. (2005). On fundamental skew distributions. *Journal of Multivariate Analysis*, 96, 93-116.
- Carroll RJ, Ruppert D, Stefanski LA, et al. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*. 2nd edition. London: Chapman & Hall.
- Chiu G, Lockhart R, Routledge R. (2006). Bent-cable regression theory and applications. *Journal of the American Statistical Association*, 101, 542--553.
- Chiu GS, Lockhart RA. (2010). Bent-cable regression with autoregressive noise. *The Canadian Journal of Statistics*, 38, 386--407.
- Dagne GA, (2017). Bayesian two-part bent-cable Tobit models with skew distributions: Application to AIDS studies. *Statistical Methods in Medical Research*. *Statistical Methods in Medical Research*, 27(12) 3696--3708.
- Dagne GA. (2018). Bayesian two-part bent-cable Tobit models with skew distributions: Application to AIDS studies. *Statistical Methods in Medical Research*, 27(12,) 3696--3708.
- Dagne GA. (2019). Bayesian semiparametric growth models for measurement error and missing data in CD4/CD8 ratio: Application to AIDS Study. *Statistical Methods in Medical Research*, 0(0) 1--11.
- Gelman A, Carlin JB, Stern HS, et al. (2004). *Bayesian Data Analysis*. New York: Chapman & Hall/CRC.
- Genton MG. (2004). *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Edited Volume, Boca Raton: Chapman & Hall / CRC.
- Hall CB, Ying J, Kuo L, et al. (2003). Bayesian and profile likelihood changepoint methods for modeling cognitive function over time. *Computational statistics and data analysis*, 42, 91--109.
- Higgins KM, Davidian M, Giltinan DM. (1997). A two-step method to measurement error in time dependent covariates in nonlinear mixed-effects models, with application to IGF-I pharmacokinetics. *Journal of the American Statistical Association*, 92, 436--448.

- Hornung, R. W., and Reed, L. D. (1990). Estimation of average concentration in the presence of nondetectable values. *Applied Occupational and Environmental Hygiene*, 4: 46-51.
- Hughes, J.P. (1999). Mixed effects models with censored data with application to HIV RNA levels. *Biometrics*, 55, 625-629.
- Khan SA, Chiu G, Dubin JA. (2009). Atmospheric concentration of chlorofluorocarbons: addressing the global concern with the longitudinal bent-cable model. *Chance*, 22, 8--17.
- Kiuchi AS, Hartigan JA, Holford TR, et al. (1995). Change Points in the Series of T4 Counts Prior to AIDS. *Biometrics*, 51, 236-48.
- Lunn DJ, Thomas A, Best N, et al. (2000). WinBUGS -- a Bayesian modelling framework: concepts, structure, an extensibility. *Statistics and Computing*, 10, 325--337.
- Ntzoufras I. (2009). *Bayesian Modeling Using WinBUGS*. Wiley: New Jersey: Wiley.
- Paterson DL, Swindells S, Mohr J, et al. (2000). Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. *Ann Intern Med*, 133, 21--30.
- Paxton, W. B., Coombs, R. W., McElrath, M.J., Keefer, M.C., Hughes, J., and Corey, L. (1997). Longitudinal analysis of quantitative virologic measures in human immunodeficiency virus-infected subjects with <400 CD4 lymphocytes: Implication for applying measurements to individual patients. *Journal of Infectious Disease*, 175, 247-254.
- Pinheiro JC, Bates DM. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer.
- Raboud JM, Montaner JSG, Conway B, et al. (1996). Variation in plasma RNA levels, CD4 cell counts, and p24 antigen levels in clinically stable men with human immunodeficiency virus infection. *Journal of Infectious Diseases*, 174(1), 191--194.
- Sahu, S.K., Dey, D.K., Branco, M.D. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *The Canadian Journal of Statistics*, 31, 129--150.
- Khan, AS and Kar, SC. (2018) Generalized bent-cable methodology for changepoint data: a Bayesian approach. *Journal of Applied Statistics*, 45(10), 1799--1812, DOI:10.1080/02664763.2017.1391754
- Slate E, Lark L. (2001). Using PSA to detect prostate cancer onset: an application of Bayesian retrospective and prospective changepoint identification. *Journal of Educational and Behavioral Statistics*, 26, 443-468.
- Spiegelhalter DJ, Best NG, Carlin BP, et al. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, 64, 583--639.
- Tishler A, Zang I. (1981). A new maximum likelihood algorithm for piecewise regression. *Journal of the American Statistical Association*, 76, 980-987.
- Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26, 24-36.
- Verbeke G, Lesaffre E. (1998). A linear mixed-effects model with heterogeneity in random-effects population. *Journal of the American Statistical Association*, 91, 217--221.
- Wu H, Zhang J-T. (2002). The study of long-term HIV dynamics using semi-parametric non-linear mixed-effects models. *Statistics in Medicine*, 21, 3655--3675.
- Wu, L. (2002). A joint model for nonlinear mixed-effects models with censoring and covariates measured with error, with application to AIDS studies. *Journal of the American Statistical Association*, 97, 955--964.
- Zhao L, Feng D, Neelon B, Buyse M. (2015). Evaluation of treatment efficacy using a Bayesian mixture piecewise linear model of longitudinal biomarkers. *Statistics in Medicine*, 34(10), 1733--1746.