Bayesian Hierarchical Varying-coefficient Mixed Effect Model

Beomjo Park [1] Taeryon Choi [1]

[1]Department of Statistics, Korea University

Functional Mixed Effect Model

- Objective : Regressing (Multi/Univariate) clustered function responses on a set of scalar predictors
- Functional Mixed Effect Model (Guo, 2002)

$$y_i(t_{ij}) = \mathrm{x}_{ij}^\top \boldsymbol{\mu}_{g(i)}(t_{ij}) + \mathrm{z}_{ij}^\top \boldsymbol{\eta}_i(t_{ij}) + e_i(t_{ij}), \quad \mathrm{e}_i \sim (\mathbf{0}, \Sigma_e)$$

$$\mathrm{E}\left(y_{ij}\right) = \underbrace{\mathrm{x}_{ij}^\top \left[I_P \otimes \boldsymbol{\varphi}_K^\top(t_{ij})\right]}_{\boldsymbol{\chi}_{ij}} \mathrm{vec}(\boldsymbol{\Theta}_{g(i)}) + \underbrace{\mathrm{z}_{ij}^\top \left[I_V \otimes \boldsymbol{\varphi}_J^\top(t_{ij})\right]}_{Z_{ij}} \mathrm{vec}(\boldsymbol{\Xi}_i)$$

- We consider
  - ▷ Hierarchical Spectral Analysis Prior
  - ▷ Shape-restriction on Fixed effects
  - ▷ DP Mixture of Ornstein-Uhlenbeck process
  - ▷ Multivariate response observed on bivariate grid

Hierarchical Spectral Analysis Prior

- The hierarchical prior specification captures similarities across groups while each having distinct profiles.

$$\boldsymbol{\beta}_{\Theta,g} \mid \tilde{\boldsymbol{\beta}} \sim \mathrm{N}_P\left(\tilde{\boldsymbol{\beta}}, V_\beta\right), \quad \tilde{\boldsymbol{\beta}} \sim \mathrm{N}_P\left(0, V_{\tilde{\beta}}\right),$$

$$\theta_{gpk} \mid \tilde{\theta}_{pk}, \tau_{gp}^2, \gamma_{gp} \sim \mathrm{N}\left(\tilde{\theta}_{pk}, \tau_{gp}^2 \exp(-k\gamma_{gp})\right), \; k = 1, \ldots, K.$$

$$\tilde{\theta}_{pk} \mid \tilde{\tau}_p^2, \tilde{\gamma}_p \sim \mathrm{N}\left(0, \tilde{\tau}_p^2 \exp(-k\tilde{\gamma}_p)\right), \; k = 1, \ldots, K.$$

  - ▷ Exponentially decaying variance known as the geometric smoother (Lenk, 1999)
  - ▷ Estimation on groupwise profiles borrows strength from the information on overall effect profile.
- The model captures the 3 layers of hierarchy :

$$\begin{aligned} \text{Overall profile} \quad & \tilde{y}(t) = \boldsymbol{\chi}\, \mathrm{vec}(\tilde{\boldsymbol{\Theta}}) \\ \text{Groupwise profile} \quad & y_g(t) = \boldsymbol{\chi}_g\, \mathrm{vec}(\boldsymbol{\Theta}_g) \\ \text{Individual profile} \quad & y_i(t) = \boldsymbol{\chi}_i\, \mathrm{vec}(\boldsymbol{\Theta}_{g(i)}) + Z_i\, \mathrm{vec}(\boldsymbol{\Xi}_i) \end{aligned}$$

Shape-restricted Fixed effects

- The shape restriction yields the reliable inference that matches with a priori domain knowledge or theory and improves the model fit by regularization.
- Monotone shape constraint (Lenk and Choi, 2017) :

$$\frac{d}{dt}\mu_{g(i)}(t) = \delta \mu_{g(i)}^2(t), \quad \delta \in \{-1, 1\}$$

$$\Longrightarrow \mathrm{x}_{ij}^\top \boldsymbol{\mu}_{g(i)}(t_{ij}) \approx \sum_p x_{ij}^{(p)} \left[\beta_{\Theta,g(i),p} + \delta_p \boldsymbol{\theta}_{g(i),p}^\top \boldsymbol{\Phi}_K^a \boldsymbol{\theta}_{g(i),p}\right]$$

- Extra constrained prior to resolve sign indeterminacy.

$$\theta_{gp0}^{(1)} \mid \tilde{\theta}_{p0}^{(1)} \sim \mathrm{N}\left(\tilde{\theta}_{p0}^{(1)}, v_{\theta_0}^2\right) \mathrm{I}(\theta_{gp0}^{(1)} \geq 0)$$

$$\tilde{\theta}_{p0}^{(1)} \sim \mathrm{N}\left(0, v_{\tilde{\theta}_0}^2\right) \mathrm{I}(\tilde{\theta}_{p0}^{(1)} \geq 0)$$

DP Mixture of Ornstein-Uhlenbeck process

- For temporal data, we consider Ornstein-Uhlenbeck (OU) process to capture serial correlation.
- OU error process in SDE form :

$$d\boldsymbol{e}_i(t_{ij}) = -A\boldsymbol{e}_i(t_{ij}) + B d\boldsymbol{W}(t_{ij})$$

- Motivated from Quintana et al. (2017), we consider Dirichlet process (DP) mixture of OU process for specifying flexible serial error process.
- Denote $\zeta_i := [A_i, C_i]$ for every subject i. $(C := B^2)$

$$\zeta_1, \cdots, \zeta_I \mid G \overset{ind}{\sim} G, \quad G \sim \mathrm{DP}(\alpha, G_0)$$

$$G_0 = \mathrm{logNormal}\left(A; 0, \sigma_A^2\right) \times \mathrm{logNormal}\left(C; 0, \sigma_C^2\right)$$

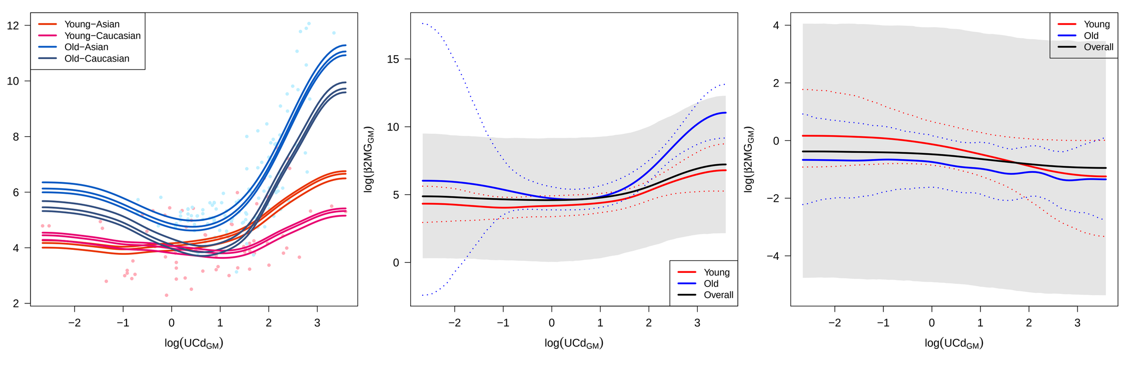
where α be a dispersion parameter and G_0 be a base measure.

Cadmium Toxicity Data

- Purpose : Obtain dose-response relationship between cadmium exposure and renal damage measured by urinary cadmium concentration (Ucd) and β2-microglobulin (β_2M) from multiple studies.
- Partially linear varying coefficient mixed model

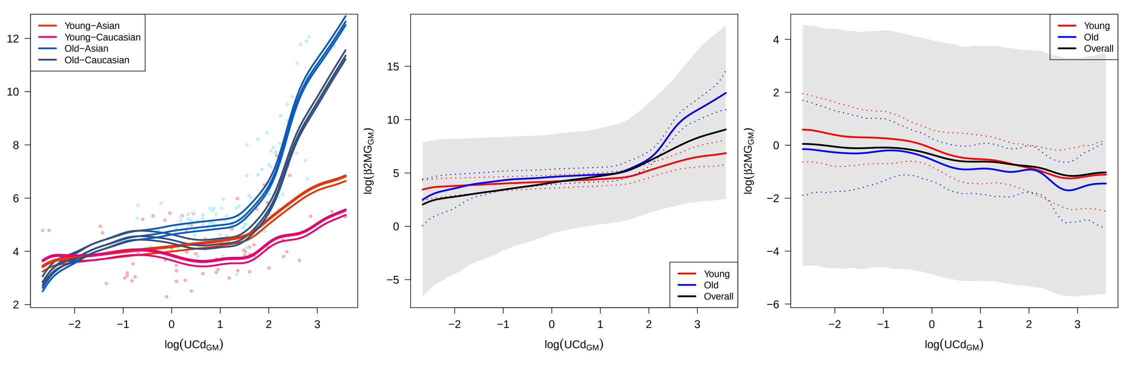
$$\log(\beta 2\mathrm{M}_{ij}) = W_{ij}\boldsymbol{\alpha}_{g(i,j)} + X_{ij}^* \boldsymbol{\mu}_{g(i,j)}(\log(\mathrm{Ucd}_{ij})) + \mathrm{Study}_i + \epsilon_{ij},$$

where $W_{ij} \in \{0, 1\}^2$ be the male and female indicator of the observation, and $X_{ij}^* \in \{1\} \times \{0, 1\}$ be the baseline and ethnicity indicator.



(a) Fixed effects (b) Baseline (c) Caucasian effect

Figure: Fixed effects profile without restriction ($\mathcal{M}1$)



(a) Fixed effects (b) Baseline (c) Caucasian effect

Figure: Fixed effects profile with restriction on baseline ($\mathcal{M}2$)

- Shape-restricted model is more preferred model.

	Baseline	$\mathcal{M}1$	$\mathcal{M}2$
RMSE	0.70	0.72	0.69
LPML	-273	-268	-248
WAIC	533	534	490

Table: Model selection criteria comparison

Study of Women's health Across the Nation

- Purpose : Characterize the menopausal transition with temporal trend of Follicle-stimulating hormone (FSH) of women from different 2 age and 4 ethnic groups ($\mathrm{x}_i \in \{0, 1\}^8$) who went through Menopause.
- $\overline{\mathtt{FSH}} = \tilde{\boldsymbol{\mu}}(t), \; \mathtt{FSH}_{ij} = \mathrm{x}_i^\top \boldsymbol{\mu}(t_{ij}) + \eta_i(t_{ij}) + e_i(t_{ij})$

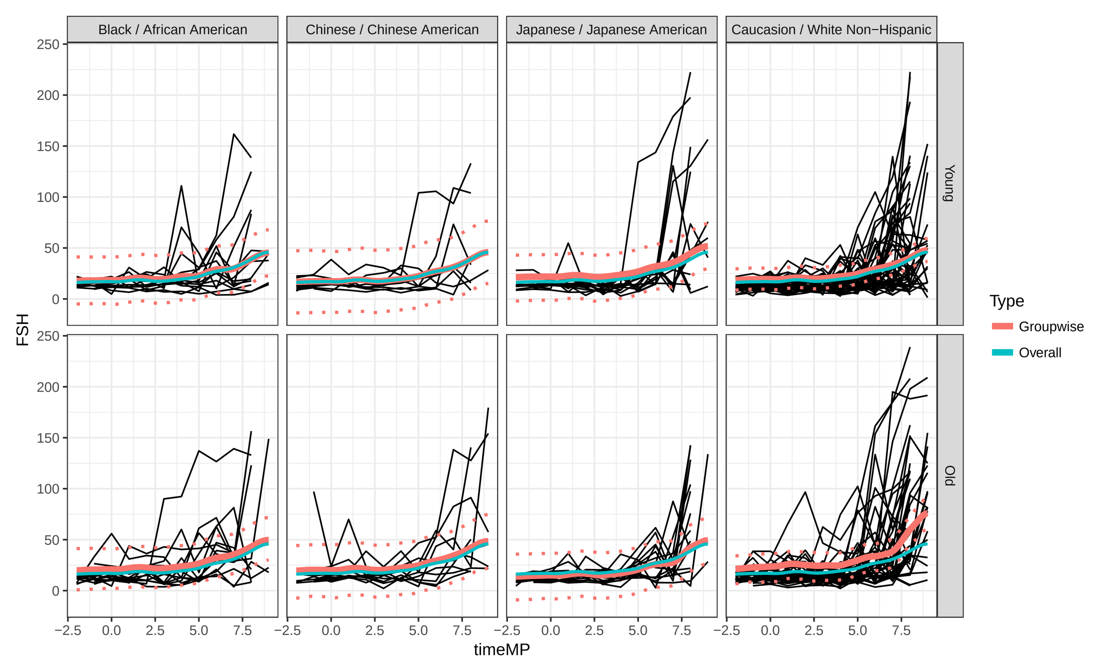


Figure: Fitted result without shape restriction

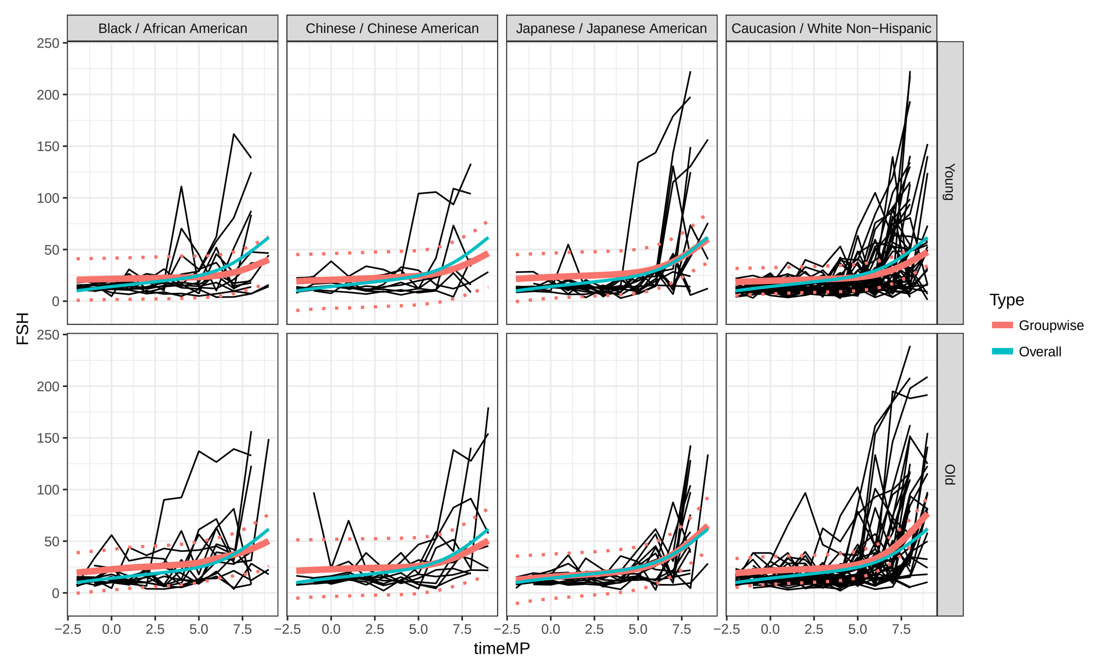
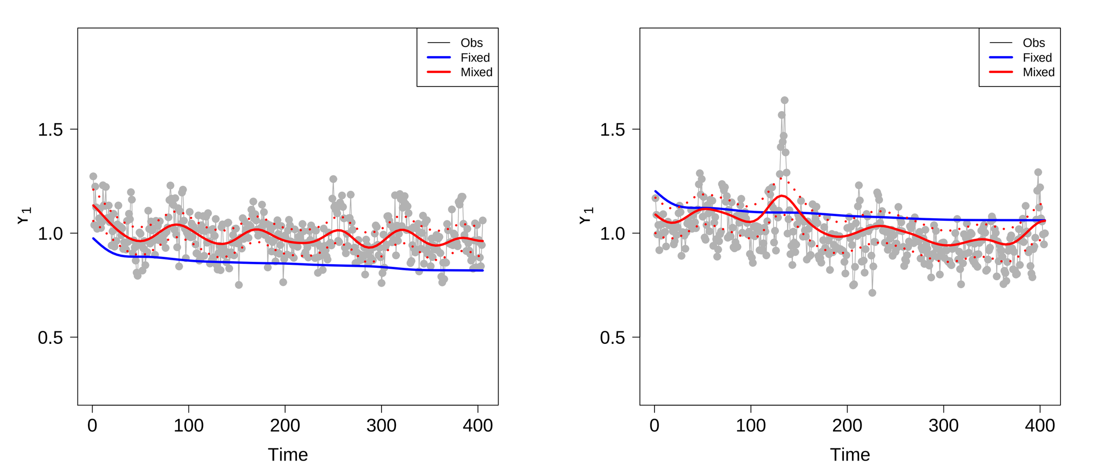


Figure: Fitted result with Increasing shape restriction

Sleeping Energy Expenditure Data

- Purpose : Identify the sleeping energy expenditure (SEE) profiles of each people grouped by obesity.
- Monotone decreasing shape restricted model

$$\mathrm{SEE}(t_{ij}) = -\boldsymbol{\theta}_{g(i)}^\top \boldsymbol{\Phi}_i^a(t_{ij}) \boldsymbol{\theta}_{g(i)} + \boldsymbol{\varphi}_i^\top(t_{ij}) \boldsymbol{\xi}_i + e_i(t_{ij}), \quad g = 1, 2$$



(a) Subject 37 – Non-obese group (b) Subject 86 – Obese group

Figure: Selected subjects from each obesity group

Multivariate model with Tensor product basis

- Rosen and Thompson (2009) considered multivariate extension of Guo (2002). Let $\boldsymbol{y}$ be L-dim functions.

$$\boldsymbol{y}_i(t_{ij}) = (I_L \otimes \mathrm{x}_{ij}^\top)\boldsymbol{\mu}_{g(i)}(t_{ij}) + (I_L \otimes \mathrm{z}_{ij}^\top)\boldsymbol{\eta}_i(t_{ij}) + \boldsymbol{e}_i(t_{ij})$$

$$\mathrm{E}\left(\boldsymbol{y}_{ij}\right) = [I_L \otimes \boldsymbol{\chi}_{ij}]\, \mathrm{vec}(\boldsymbol{\Theta}_{g(i)}^*) + [I_L \otimes Z_{ij}]\, \mathrm{vec}(\boldsymbol{\Xi}_i^*)$$

- Let $\varphi_0, \varphi_1, \ldots$ be an orthonormal basis for $L^2([0, 1])$. Then, a function f observed on bivariate grid (b_1, b_2) is

$$f(b_1, b_2) \approx \sum_{j_1=0}^{J_1} \sum_{j_2=0}^{J_2} \theta_{j_1,j_2} \varphi_{j_1}(b_1) \varphi_{j_2}(b_2) = \mathrm{vec}(\boldsymbol{\phi})^\top \mathrm{vec}(\overline{\boldsymbol{\Theta}})$$

- Applying KL expansion with respect to b_1, b_2,

$$\boldsymbol{\mu}_g(b_1, b_2, t) = \left[I_P \otimes \mathrm{vec}(\boldsymbol{\phi}_{K_1,K_2})^\top(b_1, b_2, t)\right] \mathrm{vec}(\overline{\boldsymbol{\Theta}}_g)$$

$$\boldsymbol{\eta}_i(b_1, b_2, t) = \left[I_V \otimes \mathrm{vec}(\boldsymbol{\phi}_{J_1,J_2})^\top(b_1, b_2, t)\right] \mathrm{vec}(\overline{\boldsymbol{\Xi}}_i)$$

- Hierarchical Spectral Analysis Prior

$$\mathrm{vec}(\overline{\boldsymbol{\Theta}}_g) \mid \mathrm{vec}(\widetilde{\boldsymbol{\Theta}}) \sim \mathrm{N}\left(\mathrm{vec}(\widetilde{\boldsymbol{\Theta}}), \boldsymbol{V}_{0,\overline{\Theta}_g}\right),$$

$$\mathrm{vec}(\widetilde{\boldsymbol{\Theta}}) \sim \mathrm{N}\left(\boldsymbol{m}_{0,\widetilde{\Theta}}, \boldsymbol{V}_{0,\widetilde{\Theta}}\right).$$

where $\boldsymbol{V}_{0,\overline{\Theta}_g} = \mathrm{bdiag}\left[\tau_{g1}^2 \Gamma_{g12} \otimes \Gamma_{g11}, \cdots, \tau_{gL}^2 \Gamma_{gL2} \otimes \Gamma_{gL1}\right]$,

$\boldsymbol{V}_{0,\tilde{\Xi}} = \mathrm{bdiag}\left[\boldsymbol{V}_{0,\tilde{\Xi}_1}, \cdots, \boldsymbol{V}_{0,\tilde{\Xi}_L}\right]$,

$\boldsymbol{V}_{0,\tilde{\Xi}_l} = \mathrm{diag}\left(\frac{w_0}{j_1+w_0}\right)\Big|_{j_1=0}^{J_1} \otimes \mathrm{diag}\left(\frac{w_0}{j_2+w_0}\right)\Big|_{j_2=0}^{J_2}$

  - ▷ Algebraic smoother (Lenk, 1999) on $\widetilde{\boldsymbol{\Theta}}$ allows the overall effect function to retain more high-frequency components compared to geometric smoothers.

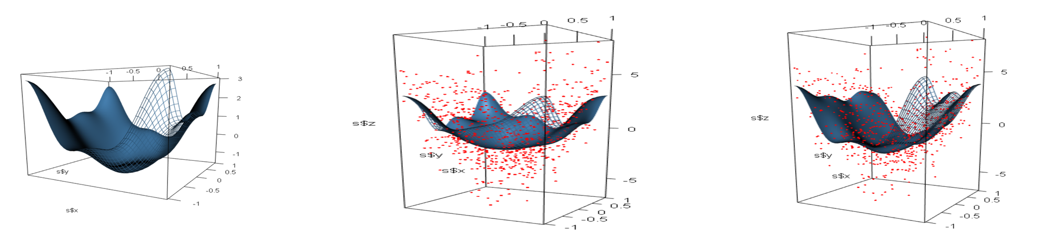
Simulation

- Bivariate temporal functions observed on bivariate grid
- 2000 obs. from 100 subjects divided into 2 groups

$$\boldsymbol{y}_{git} = \boldsymbol{\beta}_g^\top \boldsymbol{w}_{it} + f_g(x_{1it}, x_{2it}) + \boldsymbol{r}_i + \boldsymbol{e}_{it}$$

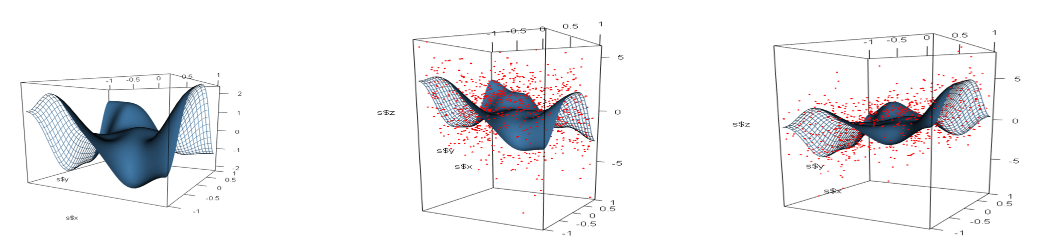
$$\boldsymbol{e}_{it} = -\begin{pmatrix} 2 & -0.6 \\ -0.6 & 2 \end{pmatrix} \boldsymbol{e}_{it} + \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \mathrm{d}\boldsymbol{W}_{it} \quad \text{or}$$

$$-\begin{pmatrix} 3 & -2 \\ -0.1 & 3 \end{pmatrix} \boldsymbol{e}_{it} + \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \mathrm{d}\boldsymbol{W}_{it}$$



(a) Overall (b) Group 1 (c) Group 2

Figure: Fitted functions of the 1st dimension



(a) Overall (b) Group 1 (c) Group 2

Figure: Fitted functions of the 2nd dimension

	β_{11}	β_{12}	β_{13}	β_{14}	β_{21}	β_{22}	β_{23}	β_{24}	σ_R
True	-2	4	-5	3	-4	2	-3	5	2
Fitted	-2.1	3.9	-5.2	3.1	-4.1	1.9	-3.0	5.1	1.66
SE	(0.12)	(0.13)	(0.11)	(0.11)	(0.12)	(0.12)	(0.13)	(0.11)	(0.20)

Table: Posterior mean and s.e. of linear components and random effect

  - ▷ DP Mixture identified the true cluster assignment with 97% accuracy with MAP estimate.

Concluding Remarks

- Unifying multi-dimensional mixed effect model for clustered functional & longitudinal data.
- Handling measurement error on smoothing variable.
  - ▷ Bridging measurement error (ME) model on FDA as an extension of ME model in classical regression.
  - ▷ Application in meta-analysis, i.e., Cadmium Toxicity, and temporal longitudinal data, i.e., SWAN data

Selected References

- Guo, W., 2002. Functional mixed effects models. Biom. 58 (1), 121–128.
- Lenk, P. J., 1999. Bayesian inference for semiparametric regression using a fourier representation. J. R. Stat. Soc. Ser.B 61 (4), 863–879.
- Lenk, P. J., Choi, T., 2017. Bayesian analysis of shape-restricted functions using Gaussian process priors. Stat. Sin. 27 (1), 43–69.
- Rosen, O., Thompson, W. K., 2009. A bayesian regression model for multivariate functional data. Comp. Stat. & Dat. Anal. 53 (11), 3773–3786.
- Quintana, F. A., Johnson, W. O., Waetjen, L. E., B. Gold, E., 2016. Bayesian non-parametric longitudinal data analysis. J. Am. Stat. Assoc. 111 (515), 1168–1181.