Baysian modeling of hurricane trajectories

P8160 Group Project 3 Baysian modeling of hurricane trajectories

Jingchen Chai, Yi Huang, Zining Qi, Ziyi Wang, Ruihan Zhang

Columbia University

2023-04-30

Content

1. Introduction
2. Dataset
3. EDA
4. Bayesian Model
5. MCMC Results
6. Conclusion

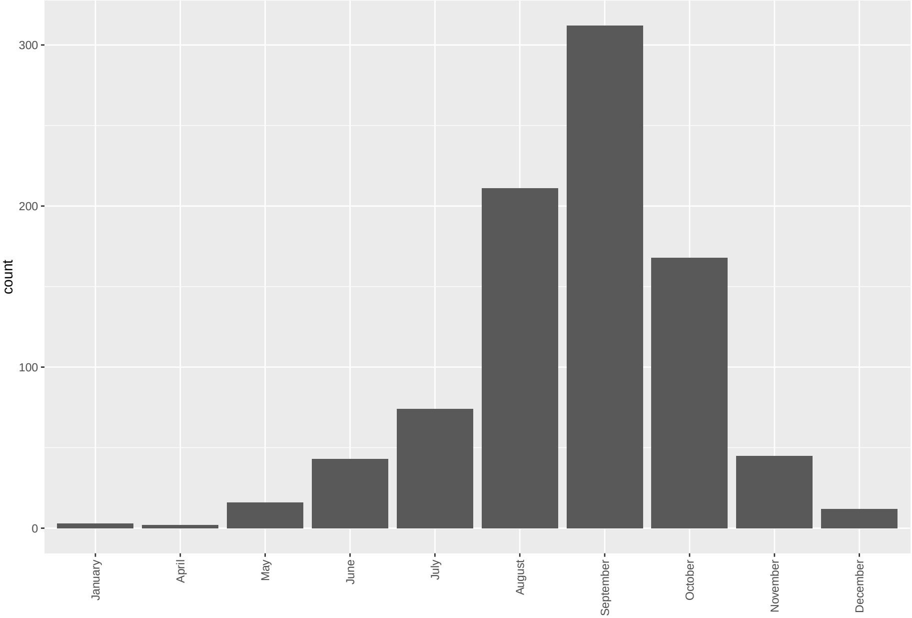
Introduction

- Hurricanes cause fatalities and property damage
- there is a growing need to accurately predict hurricane behavior, including location and speed
- This project aims to forecast wind speeds by modeling hurricane trajectories using a Hierarchical Bayesian Model.

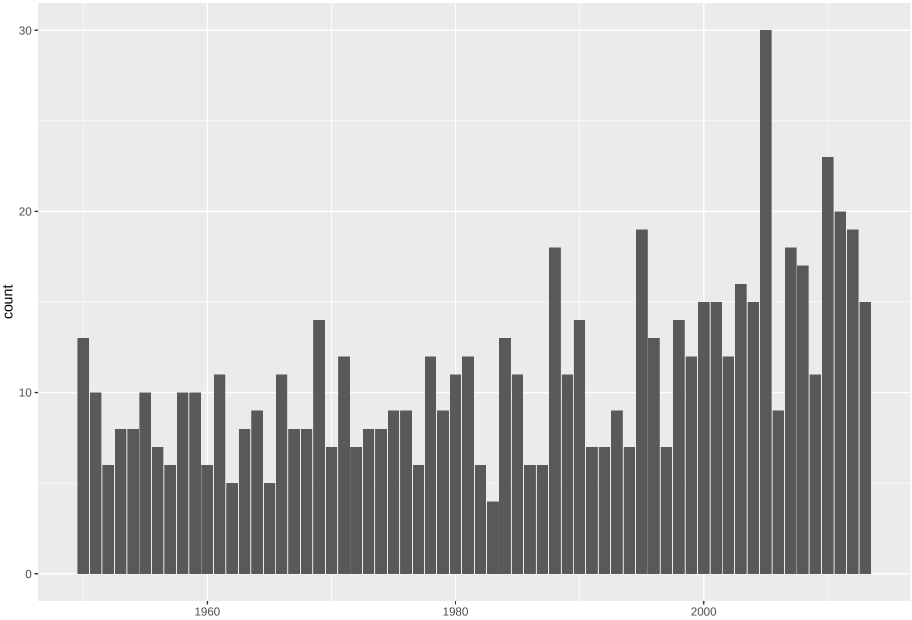
Dataset

- Hurrican703 dataset: 22038 observations $\times$ 8 variables
  - 702 hurricanes in the North Atlantic area since 1950

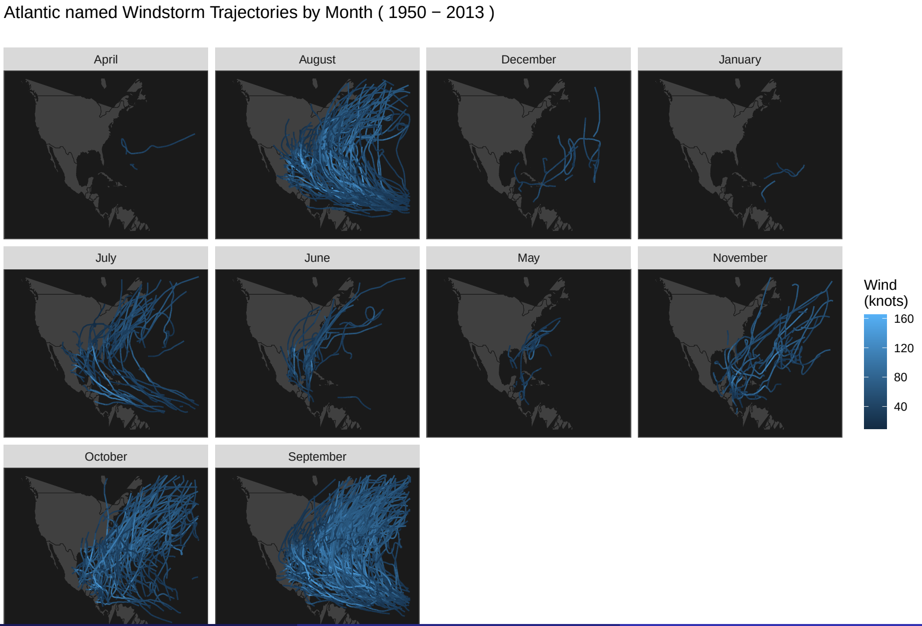
EDA-Count of Hurricanes in each Month



EDA-Count of Hurricanes in each Year



Show hurricance tracks by month



Bayesian Model

The suggested Bayesian model is
$Y_i(t+6) = \beta_{0,i} + \beta_{1,i} Y_i(t) + \beta_{2,i}\Delta_{i,1}(t) + \beta_{3,i}\Delta_{i,2}(t) + \beta_{4,i}\Delta_{i,3}(t) + \boldsymbol{X_i}\gamma + \epsilon_i(t)$

- where $Y_i(t)$ the wind speed at time t (i.e. 6 hours earlier), $\Delta_{i,1}(t)$, $\Delta_{i,2}(t)$ and $\Delta_{i,3}(t)$ are the changes of latitude, longitude and wind speed between t and $t-6$, and $\epsilon_{i,t}$ follows a normal distributions with mean zero and variance σ^2, independent across t.
- $\boldsymbol{\beta}_i = (\beta_{0,i}, \beta_{1,i}, ..., \beta_{5,i})$, we assume that $\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where d is dimension of $\boldsymbol{\beta}_i$.

Priors

$$P(\boldsymbol{\mu}) = \frac{1}{\sqrt{2\pi}|\boldsymbol{V}|^{\frac{1}{2}}} \exp\{-\frac{1}{2}\boldsymbol{\mu}^\top \boldsymbol{V}^{-1}\boldsymbol{\mu}\} \propto |\boldsymbol{V}|^{-\frac{1}{2}} \exp\{-\frac{1}{2}\boldsymbol{\mu}^\top \boldsymbol{V}^{-1}\boldsymbol{\mu}\} \quad (2)$$

$$P(\Sigma^{-1}) \propto |\Sigma|^{-\frac{(\nu+d+1)}{2}} \exp(-\frac{1}{2}tr(S\Sigma^{-1})) \quad (3)$$

$$P(\gamma) \propto exp(-\frac{\gamma^2}{2*(0.05)^2}) = e^{-200\gamma^2} \quad (4)$$

$$P(\sigma) = \frac{2\alpha}{\pi+\alpha^2} \propto \frac{1}{\sigma^2+\alpha^2} \quad (5)$$

Posterior

Let $\mathbf{B} = (\boldsymbol{\beta}_1^\top, ..., \boldsymbol{\beta}_n^\top)^\top$, derive the posterior distribution of the parameters $\Theta = (\mathbf{B}^\top, \boldsymbol{\mu}^\top, \sigma^2, \Sigma, \gamma)$.

Let

$$\boldsymbol{Z}_i(t)\boldsymbol{\beta}_i^\top = \beta_{0,i} + \beta_{1,i} Y_i(t) + \beta_{2,i}\Delta_{i,1}(t) + \beta_{3,i}\Delta_{i,2}(t) + \beta_{4,i}\Delta_{i,3}(t)$$

We can find that

$$\boldsymbol{Y_i} \sim MVN(\boldsymbol{Z}_i \beta_i, \sigma^2 I) \tag{6}$$

The likelihood for our data is

$$\begin{aligned} L(\boldsymbol{Y}_i \mid \boldsymbol{B}_i, \sigma^2) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2}(\boldsymbol{y}_i - \boldsymbol{Z}_i\beta_i - \boldsymbol{X_i}\gamma_i)^\top (\sigma^2 I)^{-1}(\boldsymbol{y}_i - \boldsymbol{Z}_i\beta_i - \\ &\propto (\sigma^{-\frac{N}{2}}) \prod_{i=1}^{n} \exp\{-\frac{1}{2}(\boldsymbol{Y}_i - \boldsymbol{Z}_i\beta_i - \boldsymbol{X_i}\gamma_i)^\top (\sigma^2 I)^{-1}(\boldsymbol{Y}_i - \boldsymbol{Z}_i\beta \end{aligned} \tag{7}$$

Joint Posrerior

$$
\begin{aligned}
P(\Theta|Y) =& P(B, \mu, \sigma^2, A, \gamma|Y) \\
& \propto L(Y|B, \sigma^2) L(B|\mu, \Sigma) p(\mu) p(\sigma) p(A) p(\gamma) \\
& \propto \frac{1}{\sigma^N(\sigma^2 + 10^2)} \prod_{i=1}^{n} \exp\{ -\frac{1}{2}(\boldsymbol{Y}_i - \boldsymbol{Z}_i \boldsymbol{\beta}_i - \boldsymbol{X_i} \gamma_i)^\top (\sigma^2 I)^{-1} (\boldsymbol{Y}_i - \\
& \times \exp\{-\frac{1}{2} \sum_{i}^{n} (\beta_i - \mu)^\top A (\beta_i - \mu)\} |A|^{\frac{N+d+6}{2}} \exp\{-\frac{1}{2} tr(SA)\} |\boldsymbol{V}|^{-\frac{1}{2}}
\end{aligned} \tag{9}
$$

Conditional Distributions

Generate B_t from $f(B_t, \sigma_{t-1}, \mu_{t-1}, \Sigma_{t-1}^{-1})$

$$f(B^\top|\sigma^2, \mu, \Sigma) \propto \prod_{i=1}^{n} \exp\{-\frac{(Y_j - Z_j\beta_j - X_j\gamma)^\top (Y_j Z_j\beta_j - X_j\gamma)}{2\sigma^2} - \frac{(\beta_j - \mu)^\top \Sigma}{}$$

$$\propto \prod_{i=1}^{n} \exp(-\frac{1}{2}(\beta_j^\top(\frac{Z_j^\top Z_j}{\sigma^2} + A)\beta_j - 2(\frac{Y_j^\top Z_j - X_j^\top Z_j\gamma}{\sigma^2}$$

$$\propto \prod_{i=1}^{n} \exp\{-\frac{1}{2}(\beta_j - (\frac{Z_j^\top Z_j}{\sigma^2} + A)(\beta_j - (\frac{Z_j^\top Z_j}{\sigma^2} + A)^{-1}(\frac{Y_j^\top Z_j - X_j^\top Z_j\gamma}{\sigma^2}$$

$$MVN_d((\frac{Z_j^\top Z_j}{\sigma^2} + A)^{-1}(\frac{Y_j^\top Z_j^\top - X_j^\top Z_j\gamma}{\sigma^2} + \mu^\top A)^\top, \frac{Z_j^\top Z_j}{\sigma^2} + A)^{-1})$$

MCMC Algorithm

$$f(\mathbf{B}|\boldsymbol{\mu}^\top, \sigma^2, A, \gamma, Y^\top) \sim MVN_d((\frac{Z_j^\top Z_j}{\sigma^2} + A)^{-1}(\frac{Y_j^\top Z_j^\top - X_j^\top Z_j\gamma}{\sigma^2} + \mu^\top A)^\top, \frac{Z_j}{\sigma}$$

$$f(\mu_t|B^\top, \sigma^2, A, \gamma, Y^\top) \sim MVN_d((NA + \frac{1}{v})^{-1}(\sum_i^n \beta_j A)^\top, (NA + \frac{1}{v})^{-1})$$

$$f(\sigma|B^\top, \mu^\top, A, \gamma, Y^\top) \sim MVN_d((X_i^\top \sigma^{-2} I X_i + 400I)^{-1}(X_i^\top \sigma^{-2} I Y_i - X_i^\top \sigma^{-2} I Z_i$$

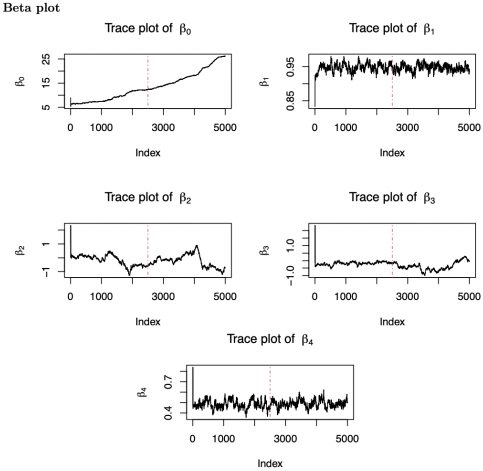
$$f(\boldsymbol{A}|\boldsymbol{B}^\top \sigma^2, \gamma, \boldsymbol{Y}^\top) \sim w^{-1}(\boldsymbol{S} + \sum_i^n (\beta_i - \boldsymbol{\mu})(\beta_i - \boldsymbol{\mu})^\top, \boldsymbol{n} + \boldsymbol{v})$$

$$f(\gamma|B^\top, \mu^t op, \sigma^2, A, Y^\top) \sim MVN((\frac{X_j^\top X_j}{\sigma^2} + 400I)^{-1}(\sum_i^n Y_j^\top X_j - \sum_i^n X_j Z_j \beta_j)^\top,$$

MCMC Results - Details

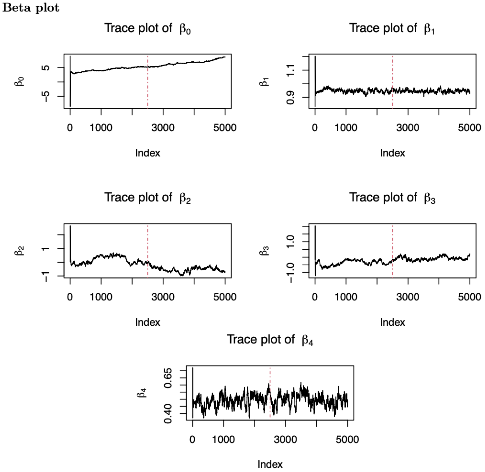
- 5000 iterations
- Estimates and inferences based on 5000 MCMC samples

MCMC Results - Beta Plots 1



-Trace plots of variance parameters, based on 5000 MCMC sample.

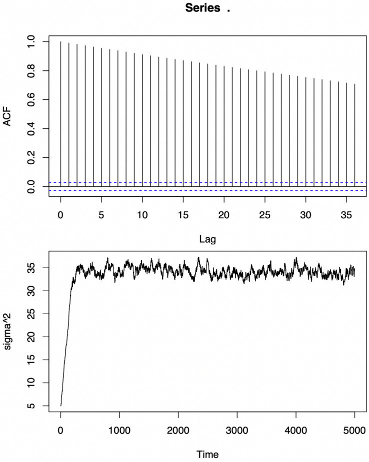
MCMC Results - Beta Plots 2

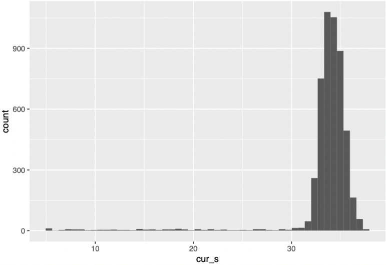


-Trace plots of variance parameters, based on 5000 MCMC sample.

MCMC Results - sigma^2 Plots 1

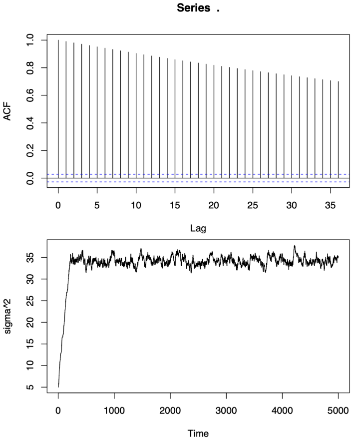
sigma^2 plot

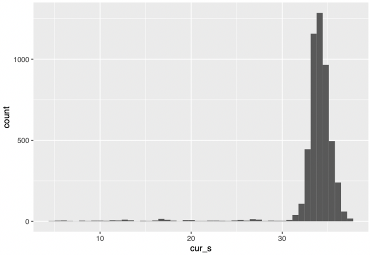




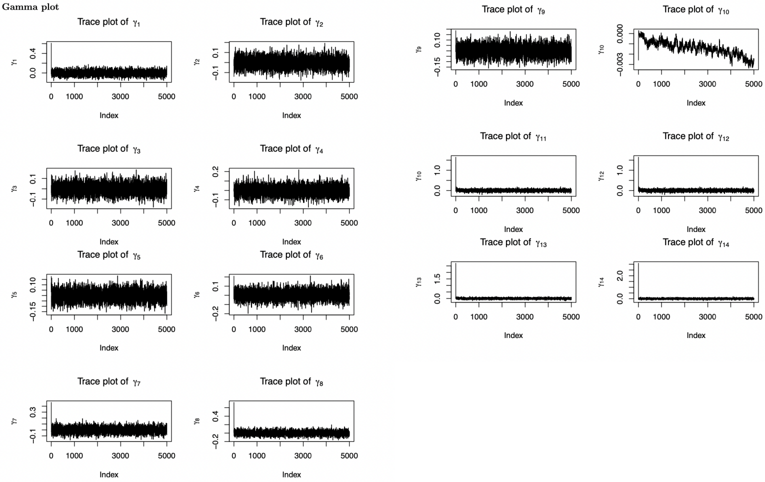
MCMC Results - sigma^2 Plots 2

sigma^2 plot

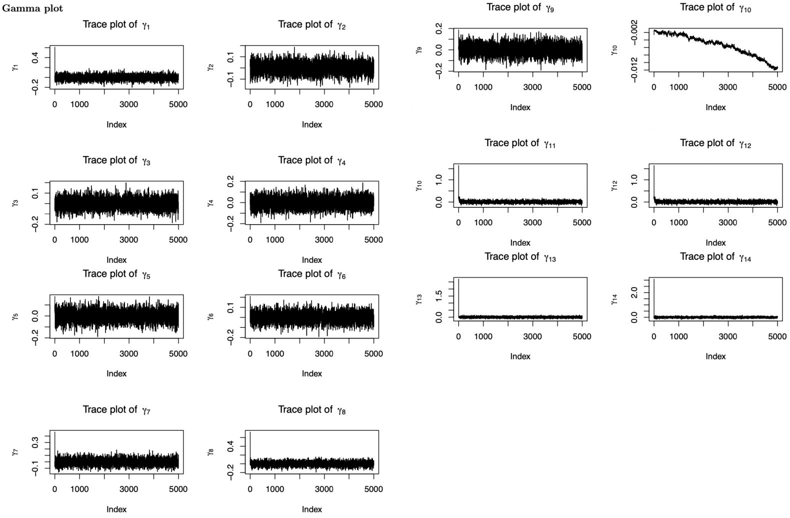




MCMC Results - Gamma Plots 1

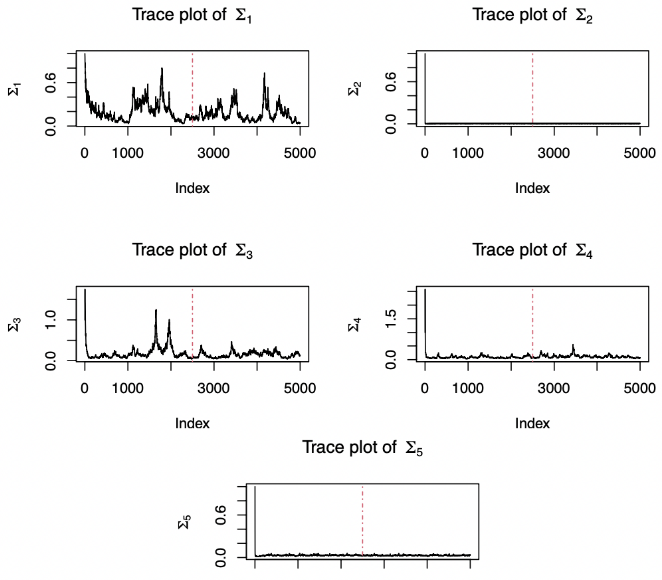


MCMC Results - Gamma Plots 2



MCMC Results - Sigma Inverse Plots 1

Sigma_inverse plot



MCMC Results - Sigma Inverse Plots 2

Sigma_inverse plot