

P8160 - Project 3

Baysian modeling of hurrican trajectories

Amy Pitts, Jiacheng Wu, Jimmy Kelliher,
Ruiqi Yan & Tianchuan Gao

2022-04-09

Motivation

Climate researchers are interested in modeling the hurricane trajectories to forecast the windspeed.

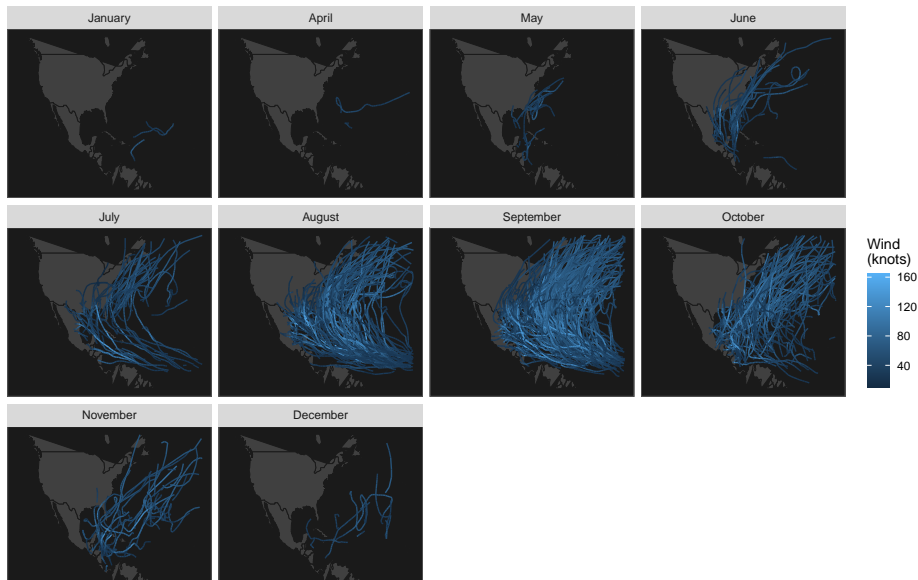
Data

- **ID:** ID of hurricanes
- **Year:** In which the hurricane occurred
- **Month:** In which the hurricane occurred
- **Nature:** Nature of the hurricane
 - ET: Extra Tropical
 - DS: Disturbance
 - NR: Not Rated
 - SS: Sub Tropical
 - TS: Tropical Storm
- **Time:** dates and time of the record
- **Latitude** and **Longitude:** The location of a hurricane check point
- **Wind.kt:** Maximum wind speed (in Knot) at each check point

- ➊ Exploration into the Data
- ➋ Bayesian modeling of hurricane wind speed
 - Model Equation
 - Posterior Derivation
 - MCMC Algorithm
- ➌ How Month, Year, and the Nature of the hurricane affect wind speed
 - Explore seasonal differences
 - Explore if wind speeds is increasing over years
- ➍ Exploring wind speeds impact on death and damages
 - As well as the characteristic of a hurricane associated with damages and deaths

Data

Atlantic named Windstorm Trajectories by Month (1950 – 2013)



- We are only concerned about observations that occurred on 6 hour intervals. hour 0, 6, 12, and 18.
- In addition we will exclude all hurricane IDs that have less than 7 observations.
- We used a lag (t to $t - 6$) for the latitude, longitude and wind speed.

Through this process we remove 460 observations so we are left with 21578 observations and 681 unique hurricanes.

Bayesian Model for Hurricane Trajectories

To model the wind speed of the i^{th} hurricane at time t we will use

$$Y_i(t+6) = \beta_{0,i} + \beta_{1,i}Y_i(t) + \beta_{2,i}\Delta_{i,1}(t) + \beta_{3,i}\Delta_{i,2}(t) + \beta_{4,i}\Delta_{i,3}(t) + \epsilon_i(t)$$

Where

- $\Delta_{i,1}(t)$, $\Delta_{i,2}(t)$ and $\Delta_{i,3}(t)$ are changes in latitude longitude and wind speed respectively between $t-6$ and t
- $\epsilon_i(t) \sim N(0, \sigma^2)$ independent across t
- Let $\beta_i = (\beta_{0,i}, \beta_{1,i}, \beta_{2,i}, \beta_{3,i}, \beta_{4,i}) \sim \mathcal{N}(\mu, \Sigma)$ be multivariate normal distribution where $\mu \in \mathbb{R}^d$ and $\Sigma \in \mathbb{R}^{d \times d}$.

Prior Distributions Assumptions:

- For σ^2 we assume $\pi(\sigma^2) \propto \frac{1}{\sigma^2}$
- For μ we assume $\pi(\mu) \propto 1$
- For Σ we assume $\pi(\Sigma^{-1}) \propto |\Sigma|^{-(d+1)} \exp\{-\frac{1}{2}\Sigma^{-1}\}$

Goal: Estimate $\Theta = (B, \mu, \Sigma^{-1}, \sigma^2)$

Likelihood & Prior Function

Likelihood

$$L(Y | \theta) \propto \prod_{i=1}^m (\sigma^2)^{-\frac{n_i}{2}} \exp \left\{ -\frac{1}{2\sigma^2} (Y_i - X_i \beta_i)^T (Y_i - X_i \beta_i) \right\}$$

where m is the number of hurricane and n_i is the number of observations for i^{th} hurricane

Prior Let $\theta = (B, \mu, \Sigma^{-1}, \sigma^2)$

$$\pi(\theta) \propto (\sigma^2)^{-1} |\Sigma^{-1}|^{d+1} \exp \left\{ -\frac{1}{2} \text{tr}(\Sigma^{-1}) \right\} \prod_{i=1}^m |\Sigma^{-1}|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\beta_i - \mu)^T \Sigma^{-1} (\beta_i - \mu) \right\}$$

where d is the dimension of μ

Posterior Calculation

Posterior

$$\begin{aligned}\pi(\theta | Y) \propto (\sigma^2)^{-\left(1 + \frac{\sum_{i=1}^m n_i}{2}\right)} |\Sigma^{-1}|^{d+1+\frac{m}{2}} \exp\left\{-\frac{1}{2} \text{tr}(\Sigma^{-1})\right\} \\ \times \exp\left\{-\frac{1}{2} \sum_{i=1}^m (\beta_i - \mu)^T \Sigma^{-1} (\beta_i - \mu)\right\} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^m (Y_i - X_i \beta_i)^T (Y_i - X_i \beta_i)\right\}\end{aligned}$$

Conditional Posterior

$$\beta_i : \pi(\beta_i | \theta_{(-\beta_i)} Y) \propto \exp\left\{-\frac{1}{2} (\beta_i - \mu)^T \Sigma^{-1} (\beta_i - \mu) - \frac{1}{2\sigma^2} (Y_i - X_i \beta_i)^T (Y_i - X_i \beta_i)\right\}$$

$$\mu : \pi(\mu | \theta_{(-\mu)}, Y) \sim N(\bar{\beta}, \Sigma/m)$$

$$\sigma^2 : \pi(\sigma^2 | \theta_{(-\sigma^2)}, Y) \propto (\sigma^2)^{-\left(1 + \frac{\sum_{i=1}^m n_i}{2}\right)} \times \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^m (Y_i - X_i \beta_i)^T (Y_i - X_i \beta_i)\right\}$$

$$\Sigma^{-1} : \pi(\Sigma^{-1} | \theta_{(-\Sigma^{-1})}, Y) \sim \text{Wishart}\left(3d + 3 + m, \left(I + \sum_{i=1}^m (\beta_i - \mu)(\beta_i - \mu)^T\right)^{-1}\right)$$

MCMC Algorithm

We apply hybrid algorithm consisting with Metropolis-Hastings steps and Gibbs steps.

Update component wise:

- Sampling proposed β'_{ij} , $j = 0, 1 \dots 4$ for i^{th} hurricane from proposal distribution $U(\beta_{ij}^{(t)} - a_{ij}, \beta_{ij}^{(t)} + a_{ij})$, where a_{ij} is the search window for β_{ij} . Since the proposals are symmetry, the accepting or rejecting the proposed β'_{ij} depends on the ratio of posterior distribution.
- Then, Gibb step for μ : Sample $\mu^{(t+1)}$ from $N(\bar{\beta}^{(t+1)}, \Sigma^{(t)}/m)$, where $\bar{\beta}^{(t+1)}$ is the average over $\beta_i^{(t+1)}$.
- Next, MH step to generate $\sigma^{2'}$ from $U(\sigma^{2(t)} - a_{\sigma^2}, \sigma^{2(t)} + a_{\sigma^2})$.
- Finally, we sample $\Sigma^{-1(t+1)}$ from
Wishart $\left(3d + 3 + m, \left(I + \sum_{i=1}^m (\beta_i^{(t+1)} - \mu^{(t+1)}) (\beta_i^{(t+1)} - \mu^{(t+1)})^T \right)^{-1} \right)$

Initial Starting Values

Initial Values:

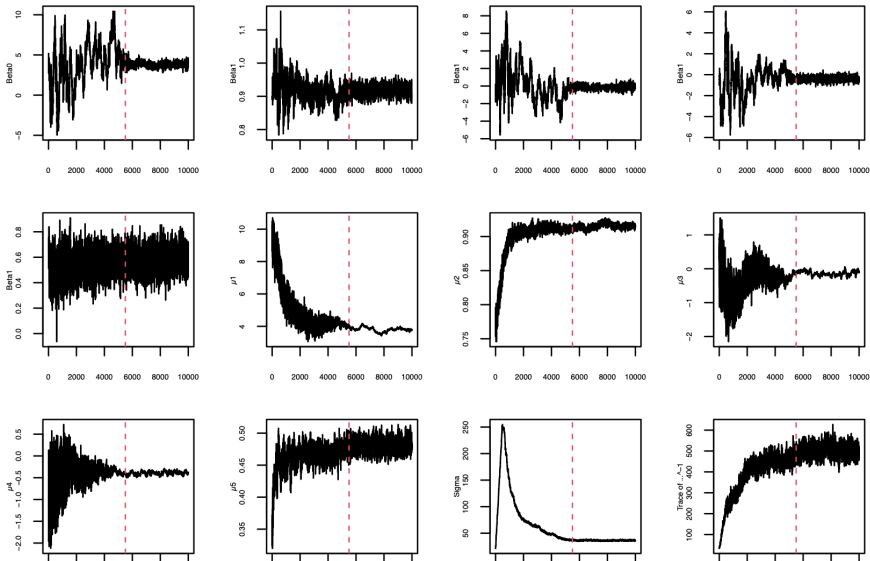
- β_i - Fit multivariate linear regression (MLR) for i^{th} hurricane and use the coefficients as β_{i0}
- μ - Average over all β_{i0} as μ_0
- σ^2 - $\hat{\sigma}_i^2$ is the mean square residuals of the MLR model on i^{th} hurricane. Take the mean over all $\hat{\sigma}_i^2$ as σ_0^2
- Σ^{-1} - Generate the covariance matrix of β_{i0} and take the inverse of the matrix as Σ_0^{-1}

Table 1: Range of Search Window and Acceptance Rate for parameters used MH step

	Search Window	Acceptance Rate (%)
β_0	1.1	45.87 - 51.36
β_1	(0.04, 0.1)	31.67 - 63.68
β_2	(0.8, 1.0)	38.60 - 45.60
β_3	(0.5, 0.6)	33.20 - 61.32
β_4	(0.4, 0.5)	34.95 - 60.45
σ^2	2.0	44.83

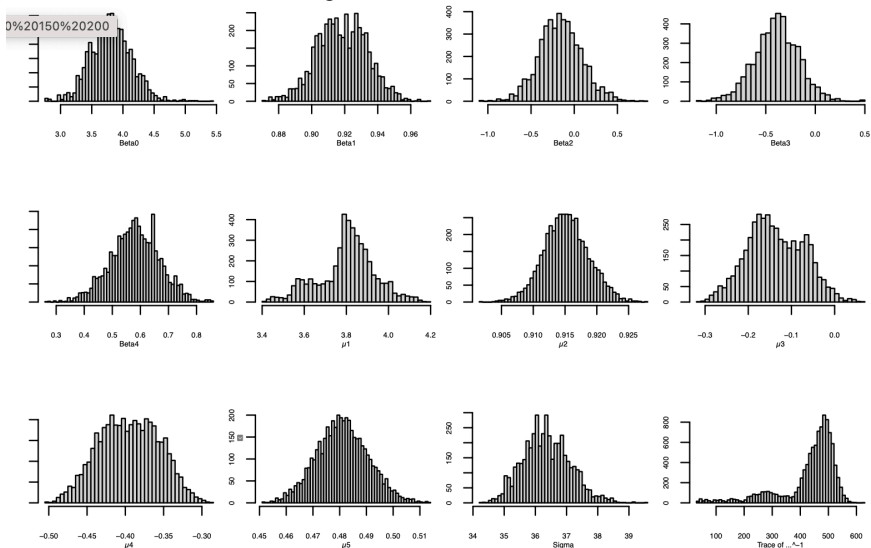
MCMC Model Convergence

Covergence Plots of Selected Parameters



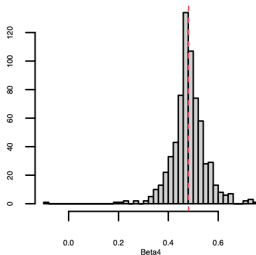
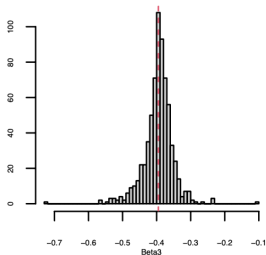
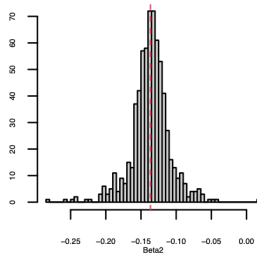
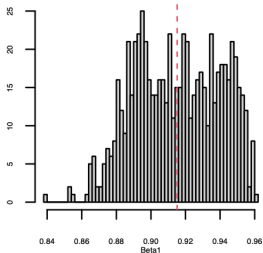
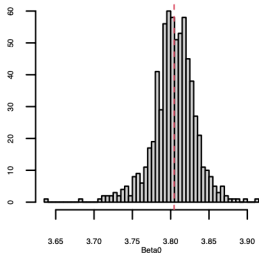
MCMC Model Convergence

Histograms of Selected Parameters



B Estimates

Histograms of estimated Betas of All Hurricanes



The μ and σ^2 Estiamtes

Model Performance

Understanding Seasonal differences in Hurricane Wind Speed

Exploring if wind speeds has increased over the years

Deaths and Damages Data Exploration

Hurricane Deaths

Hurricane Damages

Conclusions

- Larger Samples give better estimates
- High dim sample space, burn in takes longer than expected
 - very sensitive to starting values