P8160 - Project 3
Baysian modeling of hurrican trajectories

Group 6

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

2022-05-09

Abstract

abstract

1. Introduction

1.1. Background

1.2. Objectives

2. Methods

2.1. Data Cleaning and Exploratory Analysis

There are 702 unique hurricanes in this dataset all the occurred in the north american region between the years 1950 to 2013. Data on the storm's location (longitude & latitude) and maximum wind speed were recorded every 6 hours. The number of observations we have for each storm range from 1 to 118 with a mean value of 31 observations per hurricane. Data is also collected on the storms month, and the nature of the hurricane; (Extra Tropical (ET), Disturbance (DS), Not Rated (NR), Sub Tropical (SS), and Tropical Storm (TS)).

To conduct our analysis we require at least 7 observations for each unique hurricane. In addition, we are only concerned about observations that occurred on 6 hour intervals. The dataset includes a couple observations between the 6 hour time periods. For our analysis we are only going to include observations that are recorded on hour 0, 6, 12, and 18. In addition we will exclude all hurricane IDs that have less then 7 observations. Through this process we remove 460 observations so we are left with 21578 observations and 681 unique hurricanes. In addition we also created lag (t to $t-6$) variables for latitude, longitude and wind speed.

Table 1: Data Characteristics shown by the Nature of the Hurricane variable

Characteristic	Overall, N = 21,578	DS, N = 963	ET, N = 2,149	NR, N = 96	SS, N = 750	TS, N = 17,620
Wind.kt	45 (30, 65)	25 (20, 30)	40 (30, 50)	25 (20, 25)	40 (30, 50)	50 (35, 70)
Latitude	26 (19, 34)	24 (17, 31)	44 (38, 50)	18 (14, 22)	32 (30, 36)	25 (18, 31)
Longitude	-64 (-78, -48)	-65 (-81, -45)	-46 (-62, -28)	-64 (-69, -55)	-65 (-74, -51)	-65 (-80, -51)
Month						
January	69 (0.3%)	5 (0.5%)	0 (0%)	0 (0%)	19 (2.5%)	45 (0.3%)
February	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
March	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
April	53 (0.2%)	0 (0%)	23 (1.1%)	0 (0%)	17 (2.3%)	13 (<0.1%)
May	274 (1.3%)	38 (3.9%)	18 (0.8%)	0 (0%)	57 (7.6%)	161 (0.9%)
June	795 (3.7%)	34 (3.5%)	113 (5.3%)	0 (0%)	58 (7.7%)	590 (3.3%)
July	1,478 (6.8%)	95 (9.9%)	103 (4.8%)	19 (20%)	41 (5.5%)	1,220 (6.9%)
August	5,101 (24%)	290 (30%)	319 (15%)	14 (15%)	108 (14%)	4,370 (25%)
September	8,810 (41%)	249 (26%)	854 (40%)	39 (41%)	160 (21%)	7,508 (43%)
October	3,717 (17%)	163 (17%)	547 (25%)	11 (11%)	146 (19%)	2,850 (16%)
November	1,047 (4.9%)	58 (6.0%)	158 (7.4%)	13 (14%)	86 (11%)	732 (4.2%)
December	234 (1.1%)	31 (3.2%)	14 (0.7%)	0 (0%)	58 (7.7%)	131 (0.7%)

2.2 Bayesian Model for Hurricane Trajectories

Climate researchers are interested in modeling the hurricane trajectories to forecast the winds peed. To model the wind speed of the ith 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. The $\epsilon_i(t) \sim N(0, \sigma^2)$ are 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 $\Sigma \in \mathbb{R}^{d \times d}$. For this model we will be assuming non-informative or weak prior distributions for our unknown parameters σ^2, μ and Σ .

$$\pi(\sigma^2) \propto \frac{1}{\sigma^2}, \qquad \pi(\mu) \propto 1, \qquad \pi\left(\Sigma^{-1}\right) \propto |\Sigma|^{-(d+1)} \exp\left\{-\frac{1}{2}\Sigma^{-1}\right\}$$

Our goal is to estimate $\Theta = (\boldsymbol{B}, \boldsymbol{\mu}, \Sigma^{-1}, \sigma^2)$. To do this we need to establish our likelihood and prior functions. Since this is a Bayesian model we have that the likelihood will be expressed as

$$L(\boldsymbol{Y} \mid \boldsymbol{\theta}) \propto \prod_{i=1}^{m} \left(\sigma^2\right)^{-\frac{n_i}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i\right)\right\}$$

where m is the number of hurricane and n_i is the number of observations for i^{th} hurricane. The joint prior distribution is expressed as

$$\pi\left(\boldsymbol{\theta} = (\boldsymbol{B}, \boldsymbol{\mu}, \Sigma^{-1}, \sigma^2)\right) \propto \left(\sigma^2\right)^{-1} \left|\Sigma^{-1}\right|^{d+1} \exp\left\{-\frac{1}{2} \operatorname{tr}\left(\Sigma^{-1}\right)\right\} \prod_{i=1}^{m} \left|\Sigma^{-1}\right|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}\right)^T \Sigma^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\right\}$$

where d is the dimension of $\boldsymbol{\mu}$. Using the likelihood and priors derived above we can calculate the posterior distribution.

$$\pi(\boldsymbol{\theta} \mid \boldsymbol{Y}) \propto \left(\sigma^2\right)^{-1} \left|\Sigma^{-1}\right|^{d+1} \exp\left\{-\frac{1}{2} + \operatorname{tr}\left(\Sigma^{-1}\right)\right\}$$

$$\times \prod_{i=1}^{m} \left(\sigma^2\right)^{-\frac{n_i}{2}} \left|\Sigma^{-1}\right|^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i\right)\right\} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}\right)^T \Sigma^{-1}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}\right)\right\}$$

$$= \left(\sigma^2\right)^{-\left(1 + \frac{\sum_{i=1}^{m} n_i}{2}\right)} \left|\Sigma^{-1}\right|^{d+1+\frac{m}{2}} \exp\left\{-\frac{1}{2} \operatorname{tr}\left(\Sigma^{-1}\right)\right\} \exp\left\{-\frac{1}{2}\sum_{i=1}^{m}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}\right)^T \Sigma^{-1}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}\right)\right\}$$

$$\times \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i\right)\right\}$$

Due to the number of unknown parameters and the complexity we will need to use method to approximate this distribution. Before describing that process we first need to calculate the conditional posterior distribution for each unknown parameter.

$$\boldsymbol{\beta}_i : \pi(\boldsymbol{\beta}_i \mid \boldsymbol{\theta}_{(-\boldsymbol{\beta}_i)}\boldsymbol{Y}) \propto \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}\right)^T \Sigma^{-1}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}\right) - \frac{1}{2\sigma^2}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i\right)\right\}$$

$$\boldsymbol{\mu} : \pi\left(\boldsymbol{\mu} \mid \boldsymbol{\theta}_{(-\boldsymbol{\mu})}, \boldsymbol{Y}\right) \sim N(\bar{\boldsymbol{\beta}}, \Sigma/m)$$

$$\sigma^2 : \pi\left(\sigma^2 \mid \boldsymbol{\theta}_{(-\sigma^2)}, \boldsymbol{Y}\right) \propto \left(\sigma^2\right)^{-\left(1 + \frac{\sum_{i=1}^{m} n_i}{2}\right)} \times \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i\right)\right\}$$

$$\Sigma^{-1} : \pi\left(\Sigma^{-1} \mid \boldsymbol{\theta}_{(-\Sigma^{-1})}, \boldsymbol{Y}\right) \sim \text{Wishart}\left(3d + 3 + m, \left(\boldsymbol{I} + \sum_{i=1}^{m}\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}\right)\left(\boldsymbol{\beta}_i - \boldsymbol{\mu}\right)^T\right)^{-1}\right)$$

To see a more detailed description of each conditional posterior please see Appendix A.

2.3 Markov Chain Monte Carlo (MCMC) Algorithm

Due to the complexity of the above posterior distribution we will use a Markov chain Monte Carlo (MCMC) process. Since we could generate full conditional posterior distribution for some parameters but not all we

will instead apply hybrid algorithm consisting with Metropolis-Hastings (MH) steps and Gibbs steps. We will preform a MH step for β_{ij}, σ^2 and sample from the $\Sigma^{-1(t+1)}$ distribution. We will also describe a gibbs step for μ using the β_{ij} gathered.

First is the MH steps for β_{ij}. Sampling proposed β'_{ij}, $j = 0, 1...4$ for i^{th} hurricane from proposal distribution $U\left(\beta_{ij}^{(t)} - a_{ij}, \beta_{ij}^{(t)} + a_{ij}\right)$, 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. Some of the parameters in $\boldsymbol{\theta}$ could be cancelled out, so the ratio simplified to be:

$$\frac{\pi\left(\boldsymbol{\beta}'_i, \boldsymbol{\theta}_{(-\boldsymbol{\beta}_i)}^{(t)} \mid \boldsymbol{Y}\right)}{\pi\left(\boldsymbol{\beta}_i^{(t)}, \boldsymbol{\theta}_{(-\boldsymbol{\beta}_i)}^{(t)} \mid \boldsymbol{Y}\right)} = \frac{\exp\left\{-\frac{1}{2\sigma^{2(t)}}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}'_i\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}'_i\right) - \frac{1}{2}\left(\boldsymbol{\beta}'_i - \boldsymbol{\mu}^{(t)}\right)^T \Sigma^{-1(t)}\left(\boldsymbol{\beta}'_i - \boldsymbol{\mu}^{(t)}\right)\right\}}{\exp\left\{-\frac{1}{2\sigma^{2(t)}}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^{(t)}\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^{(t)}\right) - \frac{1}{2}\left(\boldsymbol{\beta}_i^{(t)} - \boldsymbol{\mu}^{(t)}\right)^T\Sigma^{-1(t)}\left(\boldsymbol{\beta}_i^{(t)} - \boldsymbol{\mu}^{(t)}\right)\right\}}$$

where $\boldsymbol{\beta}'_i$ consisting with β'_{ij}, $\beta_{ik}^{(t)}$ for $k > j$ and $\beta_{ik}^{(t+1)}$ for $k < j$. The log of the ratio is

$$\log\frac{\pi\left(\boldsymbol{\beta}'_i, \boldsymbol{\theta}_{(-\boldsymbol{\beta}_i)}^{(t)} \mid \boldsymbol{Y}\right)}{\pi\left(\boldsymbol{\beta}_i^{(t)}, \boldsymbol{\theta}_{(-\boldsymbol{\beta}_i)}^{(t)} \mid \boldsymbol{Y}\right)} = -\frac{1}{2}\left(\frac{\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}'_i\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}'_i\right)}{\sigma^{2(t)}} + \left(\boldsymbol{\beta}'_i - \boldsymbol{\mu}^{(t)}\right)^T \Sigma^{-1(t)}\left(\boldsymbol{\beta}'_i - \boldsymbol{\mu}^{(t)}\right)\right)$$
$$+ \frac{1}{2}\left(\frac{\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^{(t)}\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^{(t)}\right)}{\sigma^{2(t)}} + \left(\boldsymbol{\beta}_i^{(t)} - \boldsymbol{\mu}^{(t)}\right)^T \Sigma^{-1(t)}\left(\boldsymbol{\beta}_i^{(t)} - \boldsymbol{\mu}^{(t)}\right)\right)$$

Then we randomly sample u from $U(0, 1)$ and compare $\log(u)$ with the log ratio. If the $\log(u)$ is smaller, we accept $\beta'_{ij} = \beta_{ij}^{(t+1)}$, otherwise we reject β'_{ij} and $\beta_{ij}^{(t)} = \beta_{ij}^{(t+1)}$.

Then, Gibb step for $\boldsymbol{\mu}$: Sample $\boldsymbol{\mu}^{(t+1)}$ from $N\left(\bar{\boldsymbol{\beta}}^{(t+1)}, \Sigma^{(t)}/m\right)$, where $\bar{\boldsymbol{\beta}}^{(t+1)}$ is the average over $\boldsymbol{\beta}_i^{(t+1)}$.

Next, is the MH step to generate $\sigma^{2'}$ from $U\left(\sigma^{2(t)} - a_{\sigma^2}, \sigma^{2(t)} + a_{\sigma^2}\right)$. Firstly, check whether $\sigma^{2'}$ is positive, if not, we reject $\sigma^{2'}$. Then, we randomly sample u from $U(0, 1)$ and compare $\log(u)$ with the log posterior ratio. The log posterior ratio

$$\log\frac{\pi\left(\sigma^{2'}, \boldsymbol{\theta}_{(-\sigma^2)}^{(t)} \mid \boldsymbol{Y}\right)}{\pi\left(\sigma^{2(t)}, \boldsymbol{\theta}_{(-\sigma^2)}^{(t)} \mid \boldsymbol{Y}\right)} = -\left(1 + \frac{M}{2}\right)\log(\sigma^{2'}) - \frac{1}{2\sigma^{2'}}\sum_{i=1}^{m}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^{(t+1)}\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^{(t+1)}\right)$$
$$+ \left(1 + \frac{M}{2}\right)\log(\sigma^{2(t)}) + \frac{1}{2\sigma^{2(t)}}\sum_{i=1}^{m}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^{(t+1)}\right)^T \left(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^{(t+1)}\right)$$

where M is total number of observation for all hurricanes. If the $\log(u)$ is smaller, we accept $\sigma^{2'} = \sigma^{2(t+1)}$.

Finally, we sample $\Sigma^{-1(t+1)}$ from Wishart $\left(3d + 3 + m, \left(\boldsymbol{I} + \sum_{i=1}^{m}\left(\boldsymbol{\beta}_i^{(t+1)} - \boldsymbol{\mu}^{(t+1)}\right)\left(\boldsymbol{\beta}_i^{(t+1)} - \boldsymbol{\mu}^{(t+1)}\right)^T\right)^{-1}\right)$.

2.4 Initial Starting Values for MCMC

To initiate the MCMC process we need to specify the starting values. The choice of starting value is important to help with the converence time of our algorithm. For each paramater we will chose the start values to be:

- $\boldsymbol{\beta}_i$: Fit multivariate linear regression (MLR) for i^{th} hurricane and use the coefficients as $\boldsymbol{\beta}_{i0}$

4

Table 2: Range of Search Window and Acceptance Rate for paraemters 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

- $\boldsymbol{\mu}$: Average over all $\boldsymbol{\beta}_{i0}$ as $\boldsymbol{\mu}_0$

- 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 covaraiance matrix of $\boldsymbol{\beta}_{i0}$ and take the inverse of the matrix as Σ_0^{-1}

There is a MH step in the MCMC algorithm so the choice of search window is important. The acceptance rate of the MH step is around 0.317 to 0.637, so the search window of our algorithm is appropriate. We tune the search windows multiple times to achieve this result. Table 1 demonstrates the range of search windows a and associated acceptance rates for $\boldsymbol{\beta}$ and σ^2.

2.5 MCMC Model Performance

To assess the quality of the proposed model we will assess the overall adjusted R^2 value from the Bayesian model. In addition the model performance will be assessed for each hurricane by the adjusted R^2 value as well as a goodness of fit test. The goodness of fit test we will use residuals of Bayesian estimates, r_{ij}, of the jth observation in i^{th} hurricane, to calculate the test statistics. $\chi^2_{stat} = \frac{\sum_{j=1}^{n_i} r_{ij}^2}{\sigma^2}$, where σ^2 is the estimate σ^2 from MCMC. Based on the normal assumption in intro, $\chi^2_{stat} \sim \chi^2_{n_i}$, where n_i is the number of observation for i^{th} hurricane. Visual inspection can also be used for each individual hurricane plotting their observed wind speed with the model's predicted wind speed.

2.6 Models to Explore Seasonal Differences and Yearly Trends

In the dataset information about the hurricanes start month, year and the nature of the hurricane is recorded. It is of interested to explore how these three variables affect the wind speed. Specifically we are interested in exploring the seasonal differences, and if there any evidence supporting the statement that "the hurricane wind speed has been increasing over years". To do visual inspection can first be used to help us understand how these three variables impact the beta estimates. Then we can see if there is a linear trend given the three variables using each β estimates gained from the Bayesian model described above. For each β value we fit a different linear model. We can first fit a model using the three variables (month, year, and nature) as predictors and the 5 different β values as the outcome. This will result in 5 different linear models, one for each β value.

$$Y_{ji} = \alpha_{0j} + \alpha_{1j} \times \text{Decade}_i + \alpha_{(k+1)j} I(\text{Nature} = k)_i + \alpha_{(l+5)j} I(\text{Month} = l)_i + \epsilon_{ji}$$

Where i is the hurricane, j is the Beta model (0 through 4), $k \in (ET, NR, SS, TS)$ making DS the reference group. Let l in (April - December) where January is the reference group. We chose to include nature and month as categorical variables and year (transformed into decade) as a continuous variable.

We are also going to consider 5 different models just using decade as a predictor. $Y_i = \alpha_{0i} + \alpha_{1i} \times \text{Decade}$ where Y_i is each β_i and $i \in (0, \ldots, 4)$.

2.7 Models to Explore Death and Damages

FILL IN

3. Results

3.1 MCMC Model Convergence

Using the starting values describe above we can now examine the model performance. We are particular interested in the parameter convergence.
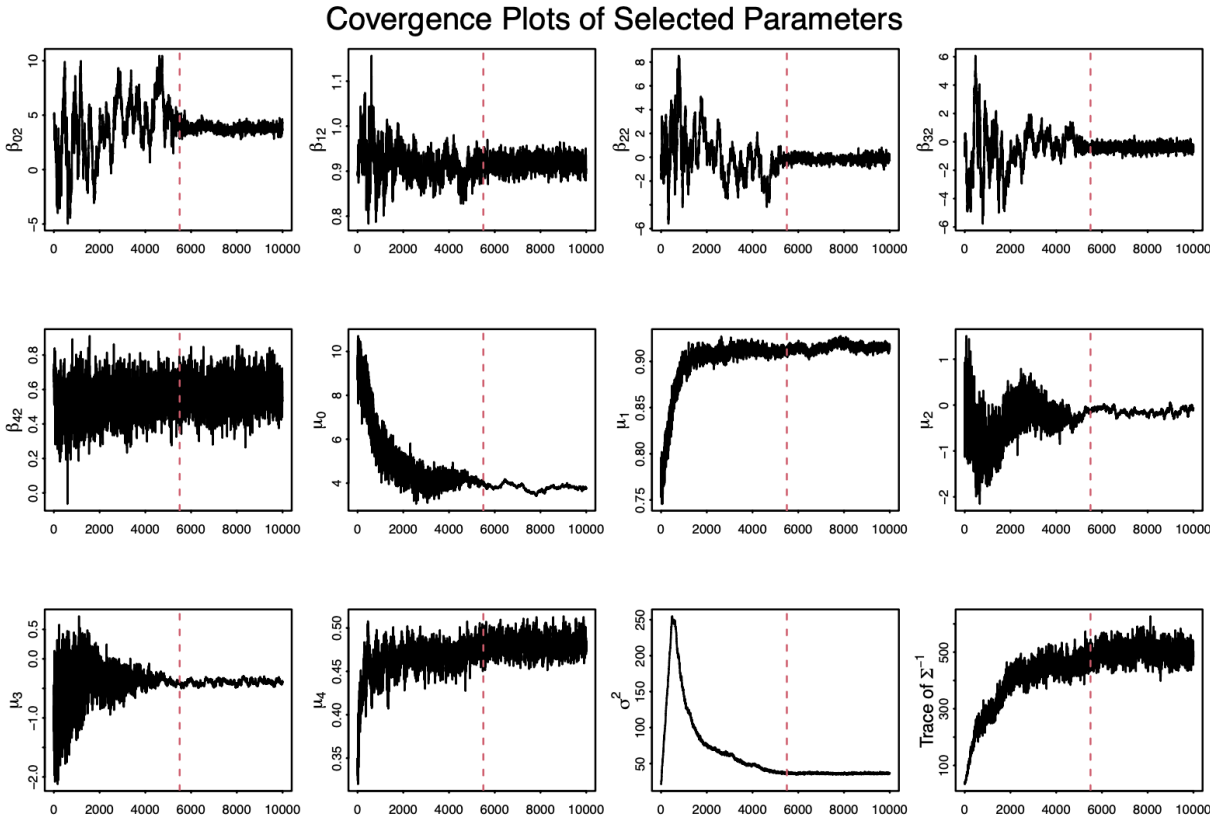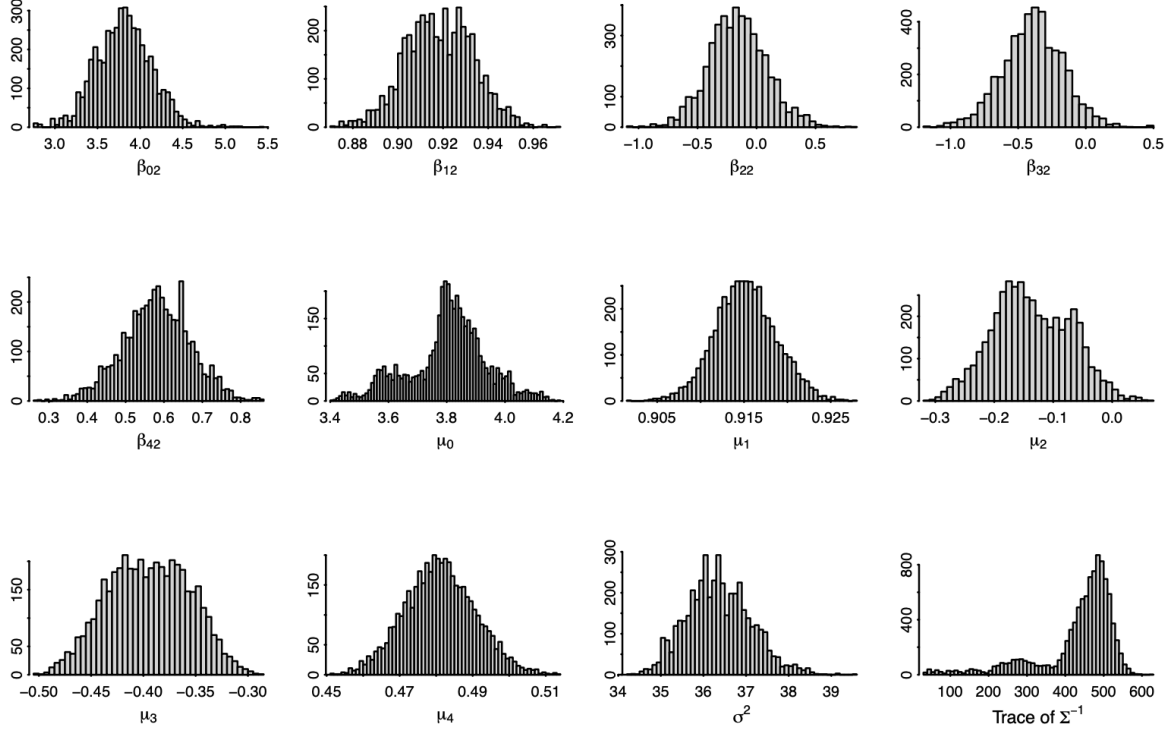

Covergence Plots of Selected Parameters

Histograms of Selected Parameters

	μ_0	μ_1	μ_2	μ_3	μ_4	σ^2	σ^2_{00}	σ^2_{11}	σ^2_{22}	σ^2_{33}	σ^2_{44}
Estimates	3.8	0.92	-0.14	-0.39	0.48	36.36	0.063	0.003	0.047	0.042	0.018

$$\hat{\Sigma}^{-1} = \begin{bmatrix} 16.07 & 7.63 & 0.34 & -1.78 & 0.47 \\ 7.63 & 381.32 & 5.39 & 3.96 & -6.32 \\ 0.34 & 5.39 & 21.43 & 1.48 & 1.14 \\ -1.78 & 3.96 & 1.48 & 24.35 & -3.3 \\ 0.47 & -6.32 & 1.14 & -3.3 & 55.12 \end{bmatrix} \quad \hat{\rho} = \begin{bmatrix} 1 & -0.101 & -0.018 & 0.094 & -0.011 \\ -0.101 & 1 & -0.057 & -0.043 & 0.043 \\ -0.018 & -0.057 & 1 & -0.067 & -0.042 \\ 0.094 & -0.043 & -0.067 & 1 & 0.089 \\ -0.011 & 0.043 & -0.042 & 0.089 & 1 \end{bmatrix}$$

3.2 MCMC Model Performance

3.3 Seasonal Differences and Yearly Trends Results

3.5 Death and Damages Results

4. Discussion

4.1. Summary of Findings

4.2. Limitations

4.3 Future Work

4.4. Group Contributions

References

[1] one

[2] two

Appendices

Appendix A

Calculating the conditional Priors for each paramater.

$$\boldsymbol{\beta}_i : \pi(\boldsymbol{\beta}_i \mid \boldsymbol{\theta}_{(-\boldsymbol{\beta}_i)}\boldsymbol{Y}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^T \Sigma^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu}) - \frac{1}{2\sigma^2}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i)^T(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i)\right\}$$

$$\boldsymbol{\mu} : \pi\left(\boldsymbol{\mu} \mid \boldsymbol{\theta}_{(-\boldsymbol{\mu})}, \boldsymbol{Y}\right) \propto \exp\left\{-\frac{1}{2}\sum_{i=1}^{m}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^T \Sigma^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\right\}$$

$$(\boldsymbol{\beta}_i - \boldsymbol{\mu})^T \Sigma^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu}) = \operatorname{tr}\left((\boldsymbol{\beta}_i - \boldsymbol{\mu})^T \Sigma^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\right)$$

$$= \operatorname{tr}\left(\Sigma^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})(\boldsymbol{\beta}_i - \boldsymbol{\mu})^T\right)$$

$$\pi\left(\boldsymbol{\mu} \mid \boldsymbol{\theta}_{(-\boldsymbol{\mu})}, Y\right) \propto \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}\sum_{i=1}^{m}(\boldsymbol{\beta}_i - \boldsymbol{\mu})(\boldsymbol{\beta}_i - \boldsymbol{\mu})^T\right)\right\}$$

$$= \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}m(\boldsymbol{\mu} - \bar{\boldsymbol{\beta}})(\boldsymbol{\mu} - \bar{\boldsymbol{\beta}})^T\right)\right\}$$

$$= \exp\left\{-\frac{1}{2}(\boldsymbol{\mu} - \bar{\boldsymbol{\beta}})^T \Sigma^{-1}m(\boldsymbol{\mu} - \bar{\boldsymbol{\beta}})\right\}$$

$$\Rightarrow \boldsymbol{\mu} \mid \boldsymbol{\theta}_{(-\boldsymbol{\mu})}, \boldsymbol{Y} \sim N(\bar{\boldsymbol{\beta}}, \Sigma/m)$$

$$\sigma^2 : \pi\left(\sigma^2 \mid \boldsymbol{\theta}_{(-\sigma^2)}, \boldsymbol{Y}\right) \propto \left(\sigma^2\right)^{-\left(1 + \frac{\sum_{i=1}^{m} n_i}{2}\right)} \times \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i)^T(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i)\right\}$$

$$\Sigma^{-1} : \pi\left(\Sigma^{-1} \mid \boldsymbol{\theta}_{(-\Sigma^{-1})}, \boldsymbol{Y}\right) \propto \left|\Sigma^{-1}\right|^{d+1+\frac{m}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}\right)\right\} \exp\left\{-\frac{1}{2}\sum_{i=1}^{m}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^T \Sigma^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\right\}$$

$$= \left|\Sigma^{-1}\right|^{d+1+\frac{m}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}\right) - \frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}\sum_{i=1}^{m}(\boldsymbol{\beta}_i - \boldsymbol{\mu})(\boldsymbol{\beta}_i - \boldsymbol{\mu})^T\right)\right\}$$

$$= \left|\Sigma^{-1}\right|^{d+1+\frac{m}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}\left(I + \sum_{i=1}^{m}(\boldsymbol{\beta}_i - \mu)(\boldsymbol{\beta}_1 - \mu)^T\right)\right)\right\}$$

$$\Rightarrow \Sigma^{-1} \mid \boldsymbol{\theta}_{(-\Sigma^{-1})}, \boldsymbol{Y} \sim \text{Wishart}\left(3d + 3 + m, \left(\boldsymbol{I} + \sum_{i=1}^{m}(\boldsymbol{\beta}_i - \boldsymbol{\mu})(\boldsymbol{\beta}_i - \boldsymbol{\mu})^T\right)^{-1}\right)$$