derive posterior distribution

Renjie Wei

2022-05-02

The suggested Bayesian model

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

where $Y_i(t)$ 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.

In the model, $\boldsymbol{\beta}_i = (\beta_{0,i}, \beta_{1,i}, ..., \beta_{5,i})$ are the random coefficients associated the ith hurricane, we assume that

$$\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

follows a multivariate normal distributions with mean $\boldsymbol{\mu}$ and covariance matrix Σ.

We assume the following non-informative or weak prior distributions for σ^2, $\boldsymbol{\mu}$ and Σ.

$$P(\sigma^2) \propto \frac{1}{\sigma^2}; \quad P(\boldsymbol{\mu}) \propto 1; \quad P(\Sigma^{-1}) \propto |\Sigma|^{-(d+1)} \exp(-\frac{1}{2}\Sigma^{-1})$$

d is dimension of β.

Posterior Distributions

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

Let

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

where $\boldsymbol{X}_i(t) = (1, Y_i(t), \Delta_{i,1}(t), \Delta_{i,2}(t), \Delta_{i,3}(t))$, $\boldsymbol{\beta}_i = (\beta_{0,i}, \beta_{1,i}, \beta_{2,i}, \beta_{3,i}, \beta_{4,i})$

then, we can find that

$$Y_i(t+6) \sim N(\boldsymbol{X}_i(t)\boldsymbol{\beta}_i^\top, \sigma^2)$$

For i^{th} hurricane, there may be m_i times of record (excluding the first and second observation), let

$$\boldsymbol{Y}_i = \begin{pmatrix} Y_i(t_0 + 6) \\ Y_i(t_1 + 6) \\ \vdots \\ Y_i(t_{m_i-1} + 6) \end{pmatrix}_{m_i \times 1}$$

denotes the m_i-dimensional result vector for the ith hurricane Therefore, since $Y_i(t)$'s are independent across t, we can show that the conditional distribution of $\boldsymbol{Y}_i$ is

$$\boldsymbol{Y}_i \mid \boldsymbol{X}_i, \boldsymbol{\beta}_i, \sigma^2 \sim N(\boldsymbol{X}_i\boldsymbol{\beta}_i^\top, \sigma^2 I)$$

where

$$\boldsymbol{X}_i = \begin{pmatrix} 1 & Y_i(t_0) & \Delta_{i,1}(t_0) & \Delta_{i,2}(t_0) & \Delta_{i,3}(t_0) \\ 1 & Y_i(t_1) & \Delta_{i,1}(t_1) & \Delta_{i,2}(t_1) & \Delta_{i,3}(t_1) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & Y_i(t_{m_i-1}) & \Delta_{i,1}(t_{m_i-1}) & \Delta_{i,2}(t_{m_i-1}) & \Delta_{i,3}(t_{m_i-1}) \end{pmatrix}_{m_i \times d}$$

and the pdf of $\boldsymbol{Y}_i$ is

$$f(\boldsymbol{Y}_i \mid \boldsymbol{\beta}_i, \sigma^2) = \det(2\pi\sigma^2 I_{(m_i \times m_i)})^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\sigma^2 I_{(m_i \times m_i)})^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)\}$$

$$= (2\pi\sigma^2)^{-m_i/2} \exp\left\{-\frac{1}{2}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\sigma^2 I_{(m_i \times m_i)})^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)\right\}$$

Since

$$\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

Therefore the pdf of $\boldsymbol{\beta}_i$ is

$$\pi(\boldsymbol{\beta}_i \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \det(2\pi\boldsymbol{\Sigma})^{-\frac{1}{2}} \exp(-\frac{1}{2}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top)$$

Notice that $\mathbf{B} = (\boldsymbol{\beta}_1^\top, ..., \boldsymbol{\beta}_n^\top)^\top$, i.e.

$$\mathbf{B} = \begin{pmatrix} \beta_{0,1} & \beta_{1,1} & \beta_{2,1} & \beta_{3,1} & \beta_{4,1} \\ \beta_{0,2} & \beta_{1,2} & \beta_{2,2} & \beta_{3,2} & \beta_{4,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta_{0,n} & \beta_{1,n} & \beta_{2,n} & \beta_{3,n} & \beta_{4,n} \end{pmatrix}_{n \times d}$$

So, by using Bayesian rule, we can show the posterior distribution of $\boldsymbol{\Theta}$ is,

$$\pi(\mathbf{\Theta}|\boldsymbol{Y}) = \pi(\mathbf{B}^\top, \boldsymbol{\mu}^\top, \sigma^2, \boldsymbol{\Sigma} \mid Y) \propto \prod_{i=1}^{n} f(\boldsymbol{Y}_i \mid \boldsymbol{\beta}_i, \sigma^2) \prod_{i=1}^{n} \pi(\boldsymbol{\beta}_i \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) P(\sigma^2) P(\boldsymbol{\mu}) P(\boldsymbol{\Sigma}^{-1})$$

$$\propto \prod_{i=1}^{n} \left\{ (2\pi\sigma^2)^{-m_i/2} \exp\left\{ -\frac{1}{2}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top (\sigma^2 I)^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)\right\}\right\}$$

$$\times \prod_{i=1}^{n} \left\{ \det(2\pi\boldsymbol{\Sigma})^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top\right\}\right\}$$

$$\times \frac{1}{\sigma^2} \times \det(\boldsymbol{\Sigma})^{-(d+1)} \exp\left\{ -\frac{1}{2}\boldsymbol{\Sigma}^{-1}\right\}$$

Markov chain Monte Carlo Algorithm to Generate the posterior distribution

Due to the high-dimensional problem of the full joint posterior distribution, considering the computational complexity plus we actually well know the form of the joint posterior, we suggest to use Gibbs sampling algorithm instead of Metropolis-Hastings algorithm.

To apply MCMC using Gibbs sampling, we need to find conditional posterior distribution of each parameter, then we can implement Gibbs sampling on these conditional posterior distributions.

Since our suggested model mainly focus on the parameter $\mathbf{B}$, we decided to derive its conditional posterior distribution.

1. The posterior distribution of $\mathbf{B}$

Since finding the posterior distribution of $\mathbf{B}$ is the same to find the posterior distribution of $\boldsymbol{\beta}_i$, we try to derive the conditional distribution $\pi(\boldsymbol{\beta}_i|\boldsymbol{Y}, \boldsymbol{\mu}^\top, \sigma^2, \boldsymbol{\Sigma})$

$$\pi(\mathbf{B}|\boldsymbol{Y}, \boldsymbol{\mu}^\top, \sigma^2, \boldsymbol{\Sigma}) \propto L_Y(\mathbf{B}^\top, \sigma^2) \times \pi(\mathbf{B} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\propto \prod_{i=1}^{n} f(\boldsymbol{Y}_i \mid \boldsymbol{\beta}_i, \sigma^2) \prod_{i=1}^{n} \pi(\boldsymbol{\beta}_i \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\propto \prod_{i=1}^{n} \left\{ (2\pi\sigma^2)^{-m_i/2} \exp\left(-\frac{1}{2}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top (\sigma^2 I)^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)\right)\right\}$$

$$\times \prod_{i=1}^{n} \left\{ \det(2\pi\boldsymbol{\Sigma})^{-\frac{1}{2}} \exp(-\frac{1}{2}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top)\right\}$$

$$\propto \prod_{i=1}^{n} \exp\left\{-\frac{1}{2}\left((\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top (\sigma^2 I)^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top) + (\boldsymbol{\beta}_i - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top)\right)\right\}$$

$$= \exp\left\{-\frac{1}{2}\left(\boldsymbol{Y}_i^\top (\sigma^2 I)^{-1}\boldsymbol{Y}_i + \boldsymbol{\beta}_i\boldsymbol{X}_i^\top(\sigma^2 I)^{-1}\boldsymbol{X}_i\boldsymbol{\beta}_i^\top - \boldsymbol{Y}_i^\top(\sigma^2 I)^{-1}\boldsymbol{X}_i\boldsymbol{\beta}_i^\top\right.\right.$$

$$\left.\left. - \boldsymbol{\beta}_i\boldsymbol{X}_i^\top(\sigma^2 I)^{-1}\boldsymbol{Y}_i + \boldsymbol{\beta}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}_i^\top + \boldsymbol{\mu}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}^\top - \boldsymbol{\mu}\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}_i^\top - \boldsymbol{\beta}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}^\top\right)\right\}$$

$$= \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}_i(\boldsymbol{\Sigma}^{-1} + \boldsymbol{X}_i^\top(\sigma^2 I)^{-1}\boldsymbol{X}_i)\boldsymbol{\beta}_i^\top - 2(\boldsymbol{Y}_i^\top(\sigma^2 I)^{-1}\boldsymbol{X}_i + \boldsymbol{\mu}\boldsymbol{\Sigma}^{-1})\boldsymbol{\beta}_i^\top + \boldsymbol{C}\right)\right\}$$

where,
$$\boldsymbol{C} = \boldsymbol{Y}_i^\top (\sigma^2 I)^{-1}\boldsymbol{Y}_i + \boldsymbol{\mu}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}^\top$$

By re-writing the conditional posterior distribution, and ignoring some constant terms, we can show that

$$\pi(\mathbf{B}|\boldsymbol{Y}, \boldsymbol{\mu}^\top, \sigma^2, \boldsymbol{\Sigma}) \propto \prod_{i=1}^{n} \exp\{(\boldsymbol{\beta}_i^\top - \hat{\boldsymbol{\beta}}_i)^\top (\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}_i})^{-1}(\boldsymbol{\beta}_i^\top - \hat{\boldsymbol{\beta}}_i)\}$$

3

Hence, each $\boldsymbol{\beta}_i$ has a conditional posterior multivariate normal distribution

$$\pi(\boldsymbol{\beta}_i|\boldsymbol{Y},\boldsymbol{\mu}^\top,\sigma^2,\boldsymbol{\Sigma}) \sim \mathcal{N}(\hat{\boldsymbol{\beta}}_i, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}_i})$$

where

$$\hat{\boldsymbol{\beta}}_i = (\boldsymbol{\Sigma}^{-1} + \boldsymbol{X}_i^\top(\sigma^2 I)^{-1}\boldsymbol{X}_i)^{-1}\boldsymbol{Y}_i^\top(\sigma^2 I)^{-1}\boldsymbol{X}_i + \boldsymbol{\mu}\boldsymbol{\Sigma}^{-1}$$
$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}_i} = (\boldsymbol{\Sigma}^{-1} + \boldsymbol{X}_i^\top(\sigma^2 I)^{-1}\boldsymbol{X}_i)^{-1}$$

2. The posterior distribution of $\pi(\sigma^2|\boldsymbol{Y},\mathbf{B}^\top,\boldsymbol{\mu}^\top,\boldsymbol{\Sigma})$

$$\pi(\sigma^2|\boldsymbol{Y},\mathbf{B}^\top,\boldsymbol{\mu}^\top,\boldsymbol{\Sigma}) \propto L_Y(\mathbf{B}^\top,\sigma^2) \times \pi(\boldsymbol{\sigma}^2)$$

$$\propto \frac{1}{\sigma^2} \prod_{i=1}^n \left\{ (2\pi\sigma^2)^{-m_i/2} \exp\left(-\frac{1}{2}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\sigma^2 I)^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top) \right) \right\}$$

$$\propto \frac{1}{\sigma^2}(\frac{\sum_{i=1}^n m_i}{2} + 1) \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)\}$$

which follows the form of pdf of inverse gamma distribution

$$f(x;\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{x} \exp\{-\frac{\beta}{x}\}$$

in this case, x is replaced by σ^2, α is replaced by $\frac{1}{2}\sum_{i=1}^n m_i$, β is replaced by $\frac{1}{2}\sum_{i=1}^n (\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)$

i.e.

$$\pi(\sigma^2|\boldsymbol{Y},\mathbf{B}^\top,\boldsymbol{\mu}^\top,\boldsymbol{\Sigma}) \sim IG(\frac{1}{2}\sum_{i=1}^n m_i, \frac{1}{2}\sum_{i=1}^n (\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top))$$

3. The posterior distribution of $\pi(\boldsymbol{\Sigma}|\boldsymbol{Y},\mathbf{B}^\top,\boldsymbol{\mu}^\top,\boldsymbol{\sigma^2})$

$$\pi(\boldsymbol{\Sigma}|\boldsymbol{Y},\mathbf{B}^\top,\boldsymbol{\mu}^\top,\boldsymbol{\sigma^2}) \propto \pi(\mathbf{B} \mid \boldsymbol{\mu},\boldsymbol{\Sigma})\pi(\boldsymbol{\Sigma}^{-1})$$

$$\propto \prod_{i=1}^n \left\{ \det(2\pi\boldsymbol{\Sigma})^{-\frac{1}{2}} \exp(-\frac{1}{2}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top) \right\} |\boldsymbol{\Sigma}|^{-(d+1)} \exp(-\frac{1}{2}\boldsymbol{\Sigma}^{-1})$$

$$\propto |\boldsymbol{\Sigma}|^{-(n+d+1+d+1)/2} \exp\{-\frac{1}{2}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top - \frac{1}{2}\boldsymbol{\Sigma}^{-1}\}$$

$$\propto |\boldsymbol{\Sigma}|^{-(n+d+1+d+1)/2} \exp\{-\frac{1}{2}tr(\boldsymbol{S}\boldsymbol{\Sigma}^{-1})\}$$

where

$$\boldsymbol{S} = \boldsymbol{I} + \sum_{i=1}^n (\boldsymbol{\beta}_i - \boldsymbol{\mu})(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top$$

which is the form of pdf of the inverse wishart distribution Inverse Wishart$(\mathbf{V},\boldsymbol{S})$, where $\mathbf{V} = n + d + 1$, i.e.

$$\pi(\boldsymbol{\Sigma}|\boldsymbol{Y},\mathbf{B}^\top,\boldsymbol{\mu}^\top,\boldsymbol{\sigma^2}) \sim IW(n + d + 1, \ \boldsymbol{I} + \sum_{i=1}^n (\boldsymbol{\beta}_i - \boldsymbol{\mu})(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top)$$

4. The posterior distribution of $\pi(\boldsymbol{\mu}|\boldsymbol{Y},\mathbf{B}^\top,\boldsymbol{\sigma^2},\boldsymbol{\Sigma})$

$$\pi(\boldsymbol{\mu}|\boldsymbol{Y}, \mathbf{B}^\top, \boldsymbol{\sigma^2}, \boldsymbol{\Sigma}) \propto \pi(\mathbf{B} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})\pi(\boldsymbol{\mu})$$

$$= \prod_{i=1}^{n} \left\{ \det(2\pi\boldsymbol{\Sigma})^{-\frac{1}{2}} \exp(-\frac{1}{2}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top) \right\}$$

$$\propto \exp\{-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top\}$$

$$\propto \exp\{-\frac{1}{2}\Big(\sum_{i=1}^{n}\boldsymbol{\beta}_i\,\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}_i^\top + n\boldsymbol{\mu}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}^\top - 2\sum_{i=1}^{n}\boldsymbol{\beta}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}^\top\Big)\}$$

$$= \exp\{-\frac{1}{2}\Big(\boldsymbol{\mu}(n\boldsymbol{\Sigma}^{-1})\boldsymbol{\mu}^\top - 2(\sum_{i=1}^{n}\boldsymbol{\beta}_i\boldsymbol{\Sigma}^{-1})\boldsymbol{\mu}^\top + \boldsymbol{C'}\Big)\}$$

$$\propto \exp\{-\frac{1}{2}(\boldsymbol{\mu} - \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\beta}_i)(n\boldsymbol{\Sigma}^{-1})(\boldsymbol{\mu} - \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\beta}_i)^\top\}$$

where

$$\boldsymbol{C'} = \sum_{i=1}^{n}\boldsymbol{\beta}_i\,\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}_i^\top$$

Hence

$$\pi(\boldsymbol{\mu}|\boldsymbol{Y}, \mathbf{B}^\top, \boldsymbol{\sigma^2}, \boldsymbol{\Sigma}) \sim \mathcal{N}(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\beta}_i, \frac{1}{n}\boldsymbol{\Sigma})$$

Markov Chain Monte Carlo

Because our hierarchical Bayesian Model exploited non-informative priors for four parameters, the Gibbs Sampling method would be implemented, updating parameters in the following order from their conditional posteriors distributions, $\mathbf{B}$, σ^2, $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}$.

Algorithm Implemtation and Estimation

Since we have derived the conditional posterior of the four parameters, we implemented Gibbs sampling algorithm, updating parameters by randomly generating samples from their conditional posterior distributions.

The update of parameters is component wise, at $(t+1)^{\text{th}}$ step, updating parameters in the following the order:

- Sample $\mathbf{B}^{(t+1)}$, i.e., sample each $\boldsymbol{\beta}_i^{(t+1)}$ from $\mathcal{N}(\hat{\boldsymbol{\beta}}_i^{(t)}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}_i}^{(t)})$

- Then, sample σ^2 from $IG(\frac{1}{2}\sum_{i=1}^{n}m_i, \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^{(t+1)\top})^\top(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^{(t+1)\top}))$

- Next, sample $\boldsymbol{\Sigma}^{(t+1)}$ from $IW(n + d + 1, \boldsymbol{I} + \sum_{i=1}^{n}(\boldsymbol{\beta}_i^{(t+1)} - \boldsymbol{\mu}^{(t)})(\boldsymbol{\beta}_i^{(t+1)} - \boldsymbol{\mu}^{(t)})^\top)$

- Finally, sample $\boldsymbol{\mu}^{(t+1)}$ from $\mathcal{N}(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\beta}_i^{(t+1)}, \frac{1}{n}\boldsymbol{\Sigma}^{(t+1)})$

Inital Values

For a good performance and to speed up the convergence of our algorithm, also keep some uncertainty in the MCMC process, we delicately designed the initial values.

For initial value of $\mathbf{B}$, we run multivariate linear regressions for each hurricane and use the regression coefficients $\boldsymbol{\beta}_i^{MLR}$ as the initial value for $\boldsymbol{\beta_i}$. Then, the initial value of $\mathbf{B}$ can be represented as $\mathbf{B}_{init} = (\boldsymbol{\beta}_1^{MLR^\top}, \ldots, \boldsymbol{\beta}_n^{MLR^\top})^\top$.

For initial value of $\boldsymbol{\mu}$, we take the average of $\boldsymbol{\beta}_i^{MLR}$, that is $\boldsymbol{\mu}_{init} = \frac{1}{n} \sum\limits_{i=1}^{n} \boldsymbol{\beta}_n^{MLR}$

For initial value of σ^2, we take the average of the MSE for i hurricanes.

For initial value of $\boldsymbol{\Sigma}$, we just set it to a simple diagonal matrix, i.e. $\boldsymbol{\Sigma}_{init} = diag(1, 2, 3, 4, 5)$

MCMC Results

Model Convergence   The figures below show the estimates of each parameters over 10000 iterations. From the trace plots below, we can see that each parameters converge quickly. We take the first 5000 iterations as our burn-in period and use the last 5000 iterations (iterations 5001 to 10000) to do the posterior parameter estimates and inferences. Figure display the histogram of parameters based on the last 5000 MCMC samples. From the histograms, we found that for β_i and μ, the distributions are relatively normal. However, we also found some skewness in the distributions of Σ, e.g. Σ_{11}, Σ_{12}, Σ_{33} and Σ_{44}.
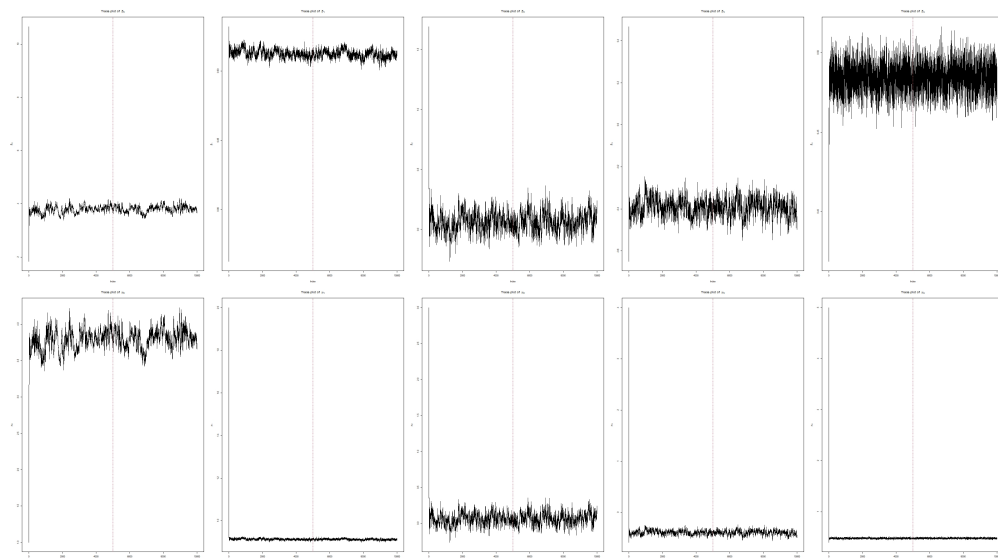

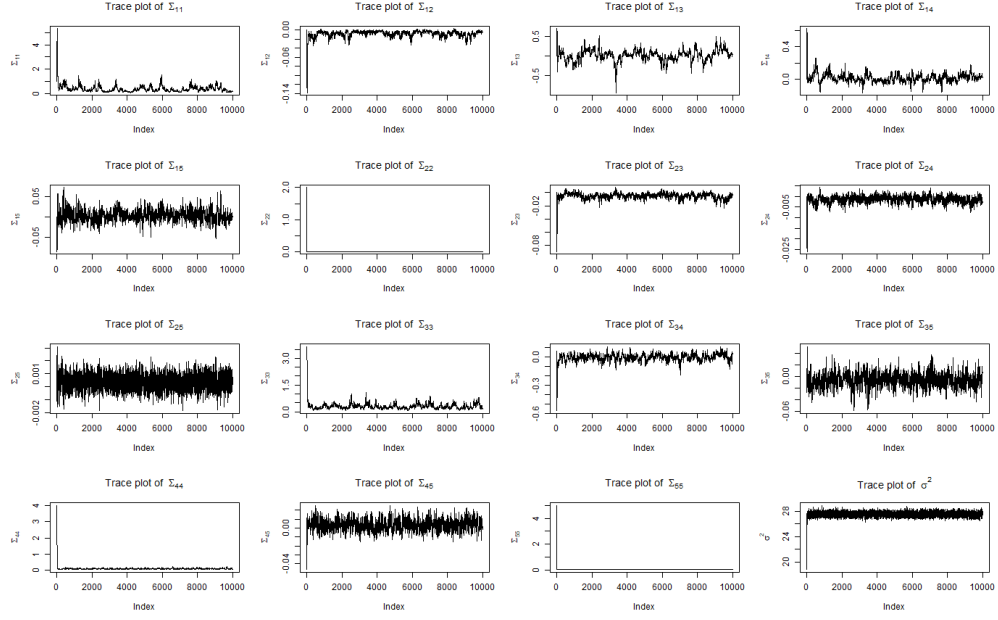Figure 1: Trace plots of model parameters, based on 10000 MCMC sample

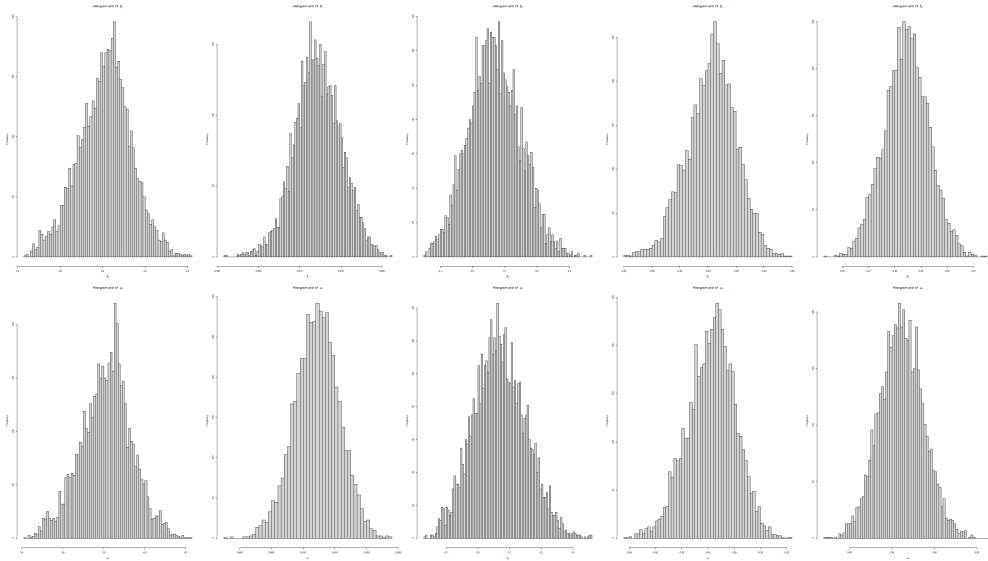Figure 2: Trace plots of variance parameters, based on 10000 MCMC sample

Figure 3: Histograms of model parameters, based on last 5000 MCMC sample
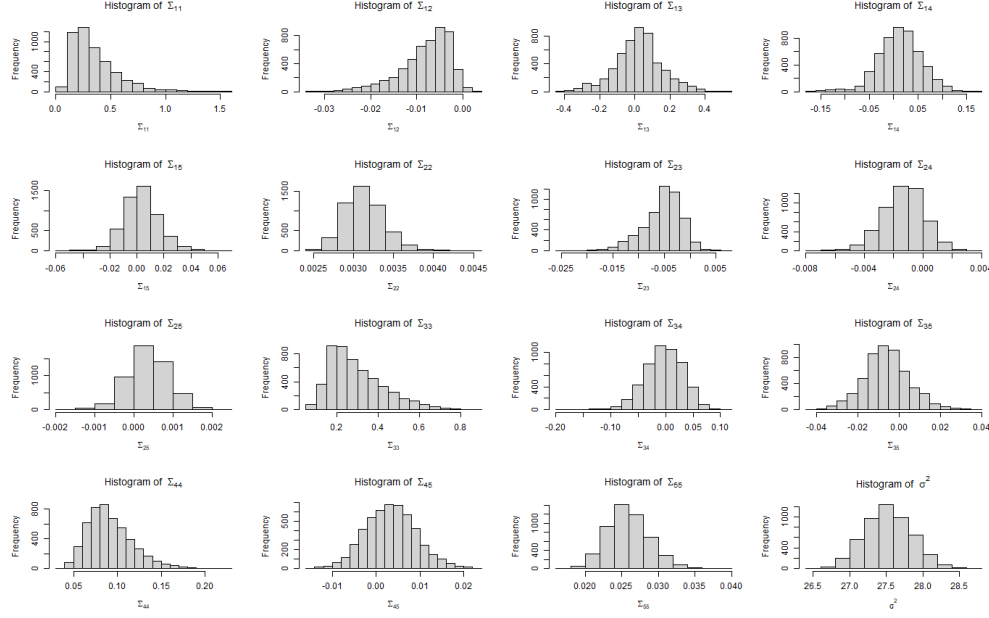
8

Figure 4: Histograms of variance parameters, based on last 5000 MCMC sample

Parameter Estimates As we stated above, we take the first 5000 iterations as our burn-in period and use the last 5000 iterations (iterations 5001 to 10000) to do the posterior parameter estimates and inferences.

We use the posterior mean as the posterior estimates for each parameter, and the 95% confidence interval is derived based on empirical rule, since the distributions of parameters are almost normal. Notice that the estimates of **B** including all $\boldsymbol{\beta}_i$'s for all hurricanes, to do the estimation and inference, we take the average value of all $\boldsymbol{\beta}_i$'s in each last 5000 MCMC iterations and then take the average of these results as the estimate of **B**. So the estimate of **B** reflects the sample mean response on the change of the hurricane speed across the predictors. The following tables shows the posterior estimates of the parameters

Variables	$\bar{\beta}_i$	$\mathrm{Var}(\bar{\beta}_i)$	95% CI of $\bar{\beta}_i$	$\bar{\mu}$	$\mathrm{Var}(\bar{\mu})$	95% CI of $\bar{\mu}$
intercept	3.8252	0.0185	(3.5588,4.0916)	3.8166	0.0190	(3.5468,4.0865)
Wind_prev	0.9118	0.0000	(0.9059,0.9177)	0.9121	0.0000	(0.9049,0.9194)
Lat_change	0.0744	0.0060	(-0.0776,0.2264)	0.0720	0.0065	(-0.0857,0.2298)
Long_change	-0.4014	0.0015	(-0.4771,-0.3257)	-0.3968	0.0016	(-0.4759,-0.3177)
Wind_change	0.4841	0.0001	(0.4674,0.5009)	0.4847	0.0001	(0.464,0.5053)

Parameters	Estimates	Variance	95% CI
Σ_{11}	0.3493	0.0435	(-0.0595,0.7581)
Σ_{12}	-0.0081	0.0000	(-0.0189,0.0027)
Σ_{13}	0.0201	0.0176	(-0.2399,0.2801)
Σ_{14}	0.0131	0.0019	(-0.0725,0.0987)
Σ_{15}	0.0035	0.0002	(-0.0215,0.0285)
Σ_{22}	0.0031	0.0000	(0.0026,0.0036)
Σ_{23}	-0.0053	0.0000	(-0.0125,0.0019)
Σ_{24}	-0.0013	0.0000	(-0.0041,0.0014)
Σ_{25}	0.0004	0.0000	(-7e-04,0.0014)
Σ_{33}	0.2960	0.0176	(0.0362,0.5558)
Σ_{34}	-0.0031	0.0012	(-0.0716,0.0653)
Σ_{35}	-0.0060	0.0001	(-0.0276,0.0156)
Σ_{44}	0.0918	0.0007	(0.0412,0.1424)
Σ_{45}	0.0034	0.0000	(-0.008,0.0148)
Σ_{55}	0.0258	0.0000	(0.0203,0.0313)
σ^2	27.5247	0.1030	(26.8957,28.1538)

The posterior estimates of $\boldsymbol{\mu}$ reflects the population average of our model parameters. For $\boldsymbol{\mu}_0$, which is the parameter of the intercept,