

P8160 - Baysian modeling of hurrican trajectories

Jiahao Fan, Tianwei Zhao, Yijin Wang, Youlan Shen, Yujia Li

2023-05-02

Contents

Background and Objectives	2
Methods	2
Bayesian Model	2
MCMC Algorithm	2
MCMC Results	4
Results	4
Seasonal Analysis	4
Prediction	4
Discussion	7
Group Contributions	7
References	8
Appendices	9
Figures and Tables	9

Background and Objectives

As hurricanes affect people, the ability to forecast hurricanes is essential for minimizing the risks in suffered areas. Hereby, a hierarchical Bayesian strategy for modeling North Atlantic hurricane counts since 1950 is illustrated. Model integration would be expected to achieve through a Markov chain Monte Carlo algorithm. Contingent on the availability of values for the covariates, the models can be used to explore the seasonality among track data and make predictive inferences on hurricanes.

The data given has 703 observations and following features are recorded for each hurricanes in the North Atlantic:

- ID: ID of the hurricanes
- Season: 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

Methods

Baysian Model

MCMC Algorithm

Given the fact that we cannot directly sample from the joint posterior distribution calculated above. We choose to use Gibbs sampling. Gibbs sampling is a sequential sampler that samples from the conditional distribution of each parameter so that the final collection of samples forms a Markov chain whose stationary distribution is the joint distribution that we are interested in.

In our case, conditional posteriors are more manageable than the joint posterior. Conditional posteriors for each parameter are calculated as the following:

!!! ADD CONDITONAL POSTERIOR

Still, we could not find a regular distribution that has the density function that matches with the conditional posterior of σ^2 . Therefore, we incorporated Metropolis Hasting Random Walk into Gibbs sampling. Metropolis Hasting Random Walk finds approximated values of σ^2 by accepting and rejecting proposed value with an acceptance probability that helps maintain the detailed balance condition. In our algorithm, we work with log-scaled posterior values instead of the original-scaled posterior because $\frac{1}{\sigma^N}$ dominates the entire posterior value where N is the total number of observations in the data set.

In our investigation of hurricane wind speed, we used Gibbs sampling and Metropolis Hasting Random Walk as the following:

Algorithm 1 Gibbs Sampling

```

1: function GIBBS(starting_values, V,S,X,D,Y, nitr, $\nu$ ,H, hurricane_idx)
2:   create posts results lists to collect results for B,  $\mu$ , A,  $\gamma$ ,  $\sigma^2$ 
3:    $B^1 \leftarrow f(B|\text{start}_\mu, \text{start}_{\sigma^2}, \text{start}_A, \text{start}_\gamma, Y)$ 
4:    $\mu^1 \leftarrow f(\mu|B^1, \text{start}_{\sigma^2}, \text{start}_A, \text{start}_\gamma, Y)$ 
5:    $A^1 \leftarrow f(A|B^1, \mu^1, \text{start}_{\sigma^2}, \text{start}_\gamma, Y)$ 
6:    $\gamma^1 \leftarrow f(\gamma|B^1, \mu^1, \text{start}_{\sigma^2}, A^1, Y)$ 
7:    $\sigma^2_{\text{chain}} \leftarrow \text{metropolis\_hasting\_sigma\_squared}(\text{start}_{\sigma^2}, \text{iter} = 500, a = 0.8, Y, D, B^1, X, \gamma^1)$ 
8:    $(\sigma^2)^1 \leftarrow \text{mean}(\sigma^2_{\text{chain}}[401:500])$ 
9:   for i in 2:nitr do
10:     $B^i \leftarrow f(B|\mu^{i-1}, (\sigma^2)^{i-1}, A^{i-1}, \gamma^{i-1}, Y)$ 
11:     $\mu^i \leftarrow f(\mu|B^i, (\sigma^2)^{i-1}, A^{i-1}, \gamma^{i-1}, Y)$ 
12:     $A^i \leftarrow f(A|B^i, \mu^i, (\sigma^2)^{i-1}, \gamma^{i-1}, Y)$ 
13:     $\gamma^i \leftarrow f(\gamma|B^i, \mu^i, (\sigma^2)^{i-1}, A^i, Y)$ 
14:     $\sigma^2_{\text{chain}} \leftarrow \text{metropolis\_hasting\_sigma\_squared}((\sigma^2)^{i-1}, \text{iter} = 500, a = 0.8, Y, D, B^i, X, \gamma^i)$ 
15:     $(\sigma^2)^i \leftarrow \text{mean}(\sigma^2_{\text{chain}}[401:500])$ 
16:   end for
17:   return posts results lists that collect results for B,  $\mu$ , A,  $\gamma$ ,  $\sigma^2$ 
18: end function

```

Figure 1: Gibbs Sampling

Algorithm 2 Metropolis Hasting sigma squared

```

1: function METROPOLIS HASTING SIGMA SQUARED(start_sigma_squared, nitr, a, B, X,D,Y, H,
   hurricane_idx)
2:   Initialize chain
3:   chain[1]  $\leftarrow$  start_sigma_squared
4:   for i in 2:itr do
5:     curr  $\leftarrow$  chain[i - 1]
6:     prop  $\leftarrow$  a sample from Unif(curr-a, curr+a)
7:     prop_posterior  $\leftarrow$  sigma_squared_log_posterior(prop,...)
8:     curr_posterior  $\leftarrow$  sigma_squared_log_posterior(curr,...)
9:     posterior_ratio  $\leftarrow$  exp(prop_posterior - curr_posterior)
10:     $\alpha \leftarrow \min(\text{posterior\_ratio}, 1)$   $\triangleright$  acceptance probability
11:    u  $\leftarrow$  runif(1)
12:    if u <  $\alpha$  then
13:      chain[i]  $\leftarrow$  prop  $\triangleright$  accept proposal
14:    else
15:      chain[i]  $\leftarrow$  curr  $\triangleright$  reject proposal
16:    end if
17:  end for
18:  return chain
19: end function

```

Figure 2: Metropolis Hasting Random Walk

We decide starting values of each parameters by fitting a linear mixed effects model that is the same as the proposed model in a frequentist way and then extract its corresponding coefficients. Starting values for γ are the fixed effects coefficients in the linear mixed effects model.

Table 1: Starting Values for Gibbs Sampling

Terms	Starting Value
γ	Fixed effect coefficients in the fitted Linear Mixed Effects Model
β_{i_1, \dots, i_n}	Random effect coefficients in the fitted Linear Mixed Effects Model
σ^2	Residual Variance in the fitted Linear Mixed Effects Model
β_0	Sum of random effect and fixed effect intercepts in the fitted Linear Mixed Effects Model
μ	Mean of the β_{i_1, \dots, i_n} for each predictor
$A = \Sigma^{-1}$	Mean of the Wishart distribution with $\nu = 5$ and scale matrix with 0.7 on the main diagonal and 0.2 on off-diagonal

With the stating values above, we assume that the true σ^2 will be around 27 and tried a few values for window size a . Finally, with $a = 0.8$, σ^2 samples converge pretty quickly and the acceptance rate is around 46% with 500 iterations. Therefore, we set $a = 0.8$ in each iteration of Gibbs Sampling and burn the first 400 samples of σ^2 .

Although our algorithm is not the most efficient algorithm, it has two main advantages: 1) It is simple to compute and does not rely on the calculation of gradient of log densities as Hamiltonian Monte Carlo; 2) By incorporating the Random Walk, we have the flexibility to control the range of sampling area when the conditional distribution of parameter of interest is difficult to simulate from and when our initial prior is not informative .

MCMC Results

We ran our algorithm with these starting values for 6000 iterations and decided to burn the first 3000 samples for each parameters for analysis in the rest of this project. We use diagnostic plots to evaluate the quality of these samples. We created trace plots for all 6000 samples for each parameter, and autocorrelation plots and histogram plots for the second half of the samples after burn-in.

Results

Seasonal Analysis

Prediction

Hurricane ID	Number of Observation	Prediction RMSE
SIXTEEN.2008	6	1.095
FABIAN.1991	9	1.245
FOUR:UNNAMED.1988	5	1.297
TEN.2005	19	1.420
LORENZO.2001	14	1.471
FIFTEEN:UNNAMED.1988	4	1.857
FABIAN.1997	15	1.933
ANDREW.1986	14	1.947
CINDY.1987	20	1.984
GUSTAV.1996	27	2.000

Figure 3: Prediction Standard Error of Wind Speed in Ascending Order

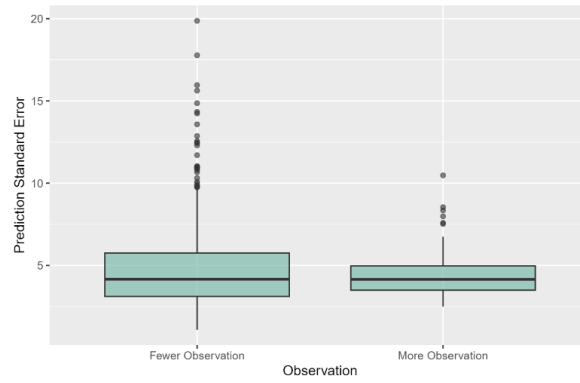


Figure 4: Prediction Standard Error of Wind Speed by Number of Observation

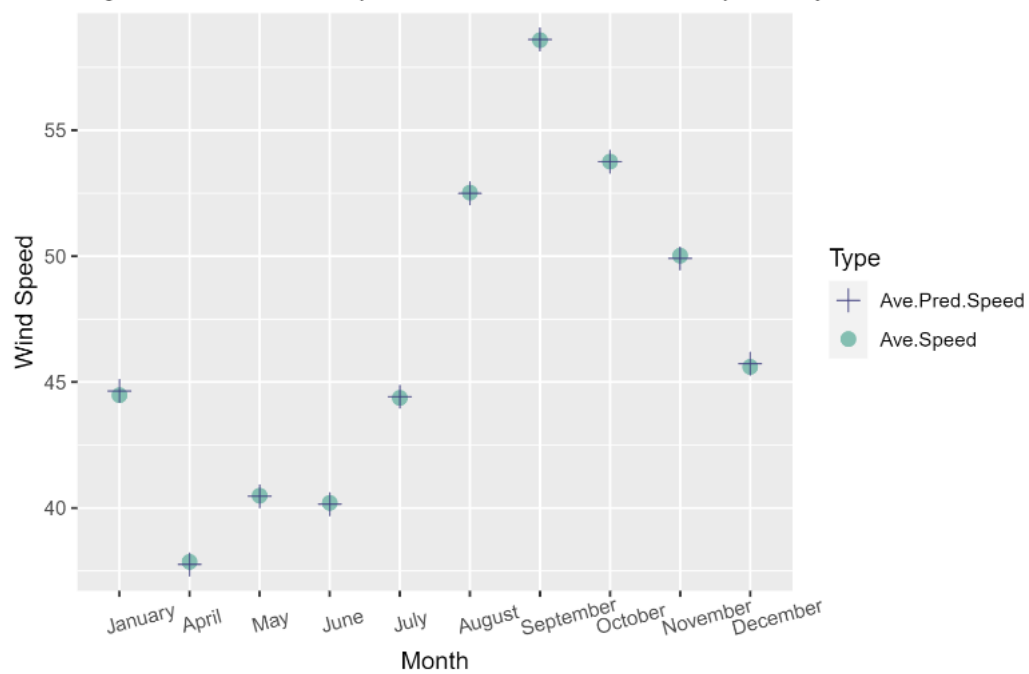


Figure 5: Actual Wind Speed vs. Predicted Wind Speed by Month

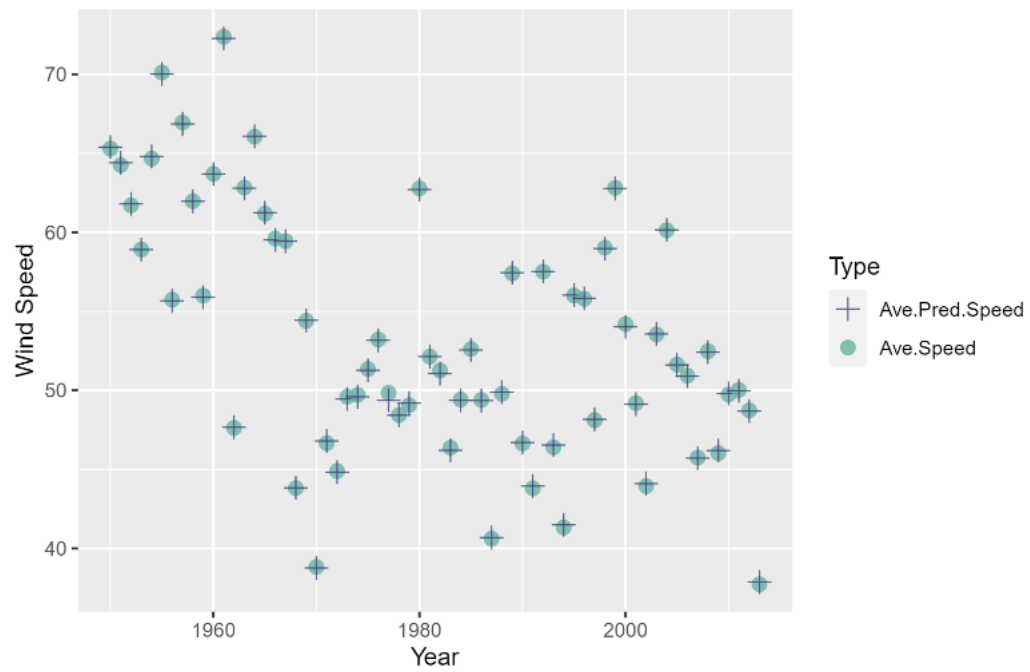


Figure 6: Actual Wind Speed vs.Predicted Wind Speed by Year

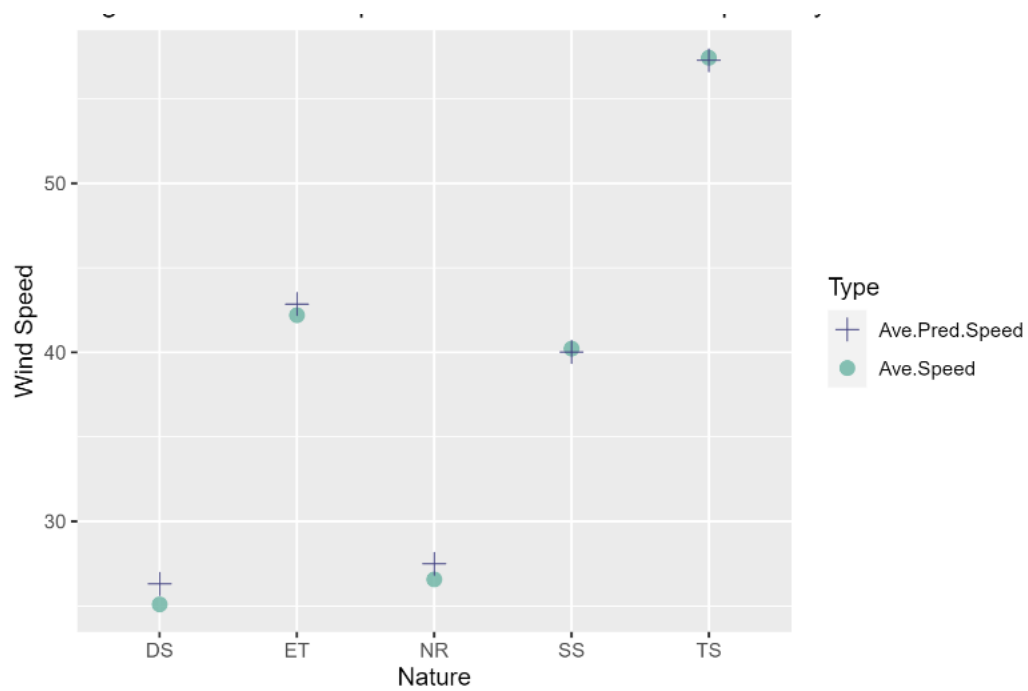


Figure 7: Actual Wind Speed vs.Predicted Wind Speed by Nature

Discussion

Group Contributions

References

Appendices

Figures and Tables