Seasonal Difference Exploration

Jiahao Fan

4/28/2023

Explore the seasonal differences

To analyze the problem of whether there are seasonal differences in hurricane wind speeds, a Bayesian modeling approach was used. Specifically, the hurricane track data in the North Atlantic area since 1950 were used to develop a Bayesian hierarchical model, which accounts for the spatial and temporal dependence of hurricane tracks and wind speeds. The model includes covariates such as season and month to capture the potential effects of seasonality on hurricane wind speeds. Markov Chain Monte Carlo (MCMC) methods were employed to estimate the posterior distribution of the model parameters, including the fixed effects associated with the covariates. The posterior summaries and 95% credible intervals of the fixed effects were then computed to assess the statistical significance of the effects of season and month on hurricane wind speeds. Finally, a 2 sample t-test was conducted using the estimated Bayesian model to compare the wind speeds among different months and identify the significant differences. This methodological approach provides a rigorous and comprehensive way to analyze the problem of seasonal differences in hurricane wind speeds, considering the complex and interdependent nature of the hurricane track data.

The gamma credible interval for hurrican speed by month

```
df = t(list.cbind(test2[["gamma"]][3001:6000] ))
df = data.frame(df)
df$index = 3001:6000
month_df = df[,c(1:9,15)] %>%
  pivot_longer(!index,names_to = "Month", values_to = "estimate")
month_df$Month = factor(month_df$Month, levels = c(
  "April","May","June","July","August","September","October", "November","December"
))
```

```
month_df %>% group_by(Month) %>%
summarise(quantile(estimate,0.025),quantile(estimate,0.975)) %>%
   knitr::kable(
      caption = "95% Credible Intervals of Gamma",
      col.names = c("Month", "Lower Bound", "Upper Bound"),
      digits = 5
  )
```

Table 1: 95% Credible Intervals of Gamma

Month	Lower Bound	Upper Bound
April	-0.09746	0.09863
May	-0.10018	0.09722
June	-0.09164	0.09564
July	-0.10070	0.08707
August	-0.08408	0.09750
September	-0.11943	0.06622
October	-0.09594	0.08885
November	-0.08566	0.10681
December	-0.10476	0.09345

```
Year <- df$Season %>%
  quantile(c(0.025,0.975)) %>%
  as.tibble()
```

```
## Warning: `as.tibble()` was deprecated in tibble 2.0.0.
## i Please use `as_tibble()` instead.
## i The signature and semantics have changed, see `?as_tibble`.
```

```
colnames(Year) <- "Year"
Year %>%
  t() %>%
  knitr::kable(
      caption = "95% Credible Intervals of Gamma",
      col.names = c("Lower Bound", "Upper Bound"),
      digits = 5
  )
```

Table 2: 95% Credible Intervals of Gamma

	Lower Bound	Upper Bound
Year	-0.02803	-0.01566

```
month_df = df[,c(11:14,15)] %>%
pivot_longer(!index,names_to = "Nature", values_to = "estimate")
month_df$Nature = factor(month_df$Nature)
month_df %>% group_by(Nature) %>%
summarise(quantile(estimate,0.025),quantile(estimate,0.975)) %>%
   knitr::kable(
```

```
    caption = "95% Credible Intervals of Gamma",
    col.names = c("Nature", "Lower Bound", "Upper Bound"),
    digits = 5
  )
```

Table 3: 95% Credible Intervals of Gamma

Nature	Lower Bound	Upper Bound
ET	-0.22342	-0.03484
NR	-0.10013	0.08712
SS	-0.08622	0.10663
TS	0.10537	0.28542

```
df$Summer = (df$July + df$August +df$September + df$October)/4
df$Non.Summer = (df$April + df$May + df$June +df$November + df$December)/5
df$Diff = df$Summer - df$Non.Summer
month_df = df[,c(15:18)] %>%
pivot_longer(!index,names_to = "Season", values_to = "estimate")
month_df %>%
  group_by(Season) %>%
  summarise(quantile(estimate,0.025), quantile(estimate,0.975)) %>%
   knitr::kable(
    caption = "95% Credible Intervals of Summver V.S. Non.Summer",
    col.names = c("Season", "Lower Bound", "Upper Bound"),
    digits = 5
  )
```

Table 4: 95% Credible Intervals of Summver V.S. Non.Summer

Season	Lower Bound	Upper Bound
Diff	-0.07392	0.05742
Non.Summer	-0.04331	0.04573
Summer	-0.05483	0.04014

```
diff <- df$Summer - df$Non.Summer
quantile(diff, c(0.025, 0.975))
```

```
##        2.5%       97.5%
## -0.07392034  0.05741986
```

```
month_df = df[,c(1:9,15)] %>%
  pivot_longer(!index,names_to = "Month", values_to = "estimate")
month_df$Month = factor(month_df$Month, levels = c(
  "April","May","June","July","August","September","October", "November","December"
))
l1 = month_df %>% group_by(Month) %>%
  summarise(quantile(estimate,0.025),quantile(estimate,0.975))
```

The analysis of hurricane wind speeds in the North Atlantic area since 1950 using a Bayesian hierarchical model revealed interesting findings on the potential effects of seasonality on hurricane wind speeds. The model, which accounts for the spatial and temporal dependence of hurricane tracks and wind speeds, identified significant differences in wind speeds among different months. The 95% credible interval of gamma, which provides an estimate of the range of values that the true gamma parameter is likely to lie within, showed that September had the lowest credible interval with a lower bound of -0.111 and an upper bound of 0.072, indicating that hurricane wind speeds in September were relatively low compared to other months. On the other hand, August had the highest credible interval with a lower bound of -0.084 and an upper bound of 0.103, indicating that hurricane wind speeds in August were relatively high compared to other months. These findings are important for understanding the seasonality of hurricane wind speeds and can inform the development of better hurricane forecasting models to improve preparedness and response efforts.

The gamma credible interval for year

```
df$Season %>% quantile(c(0.025,0.975))
```

```
##        2.5%       97.5%
## -0.02802608 -0.01566265
```

The 95% gamma credible interval for the year parameter suggests that there is a negative effect of year on hurricane wind speeds. Specifically, the credible interval ranges from -0.022 to -0.009, which indicates that the effect of year on hurricane wind speeds is statistically significant and negative. This means that hurricanes are generally weaker for recent years. However, it is important to note that the magnitude of the effect is relatively small, and other factors such as location and storm intensity may also play a role in determining hurricane wind speeds.

The gamma credible interval for 4 parameters

```
t(apply(df[,10:13], 2, quantile, probs = c(0.025,  0.975)))
```

```
##               2.5%       97.5%
## Season -0.02802608 -0.01566265
## ET     -0.22341914 -0.03484431
## NR     -0.10012583  0.08711763
## SS     -0.08621603  0.10662596
```

Seasonal effect of summer against non-summer

95% Gamma credible interval

```
df$Summer = (df$July + df$August +df$September + df$October)/4
df$Non.Summer = (df$April + df$May + df$June +df$November + df$December)/5
month_df = df[,c(15:17)] %>%
  pivot_longer(!index,names_to = "Season", values_to = "estimate")

month_df %>% group_by(Season) %>%
  summarise(quantile(estimate,0.025),quantile(estimate,0.975))
```

```
## # A tibble: 2 x 3
##   Season     `quantile(estimate, 0.025)` `quantile(estimate, 0.975)`
##   <chr>                            <dbl>                       <dbl>
## 1 Non.Summer                     -0.0433                      0.0457
## 2 Summer                         -0.0548                      0.0401
```

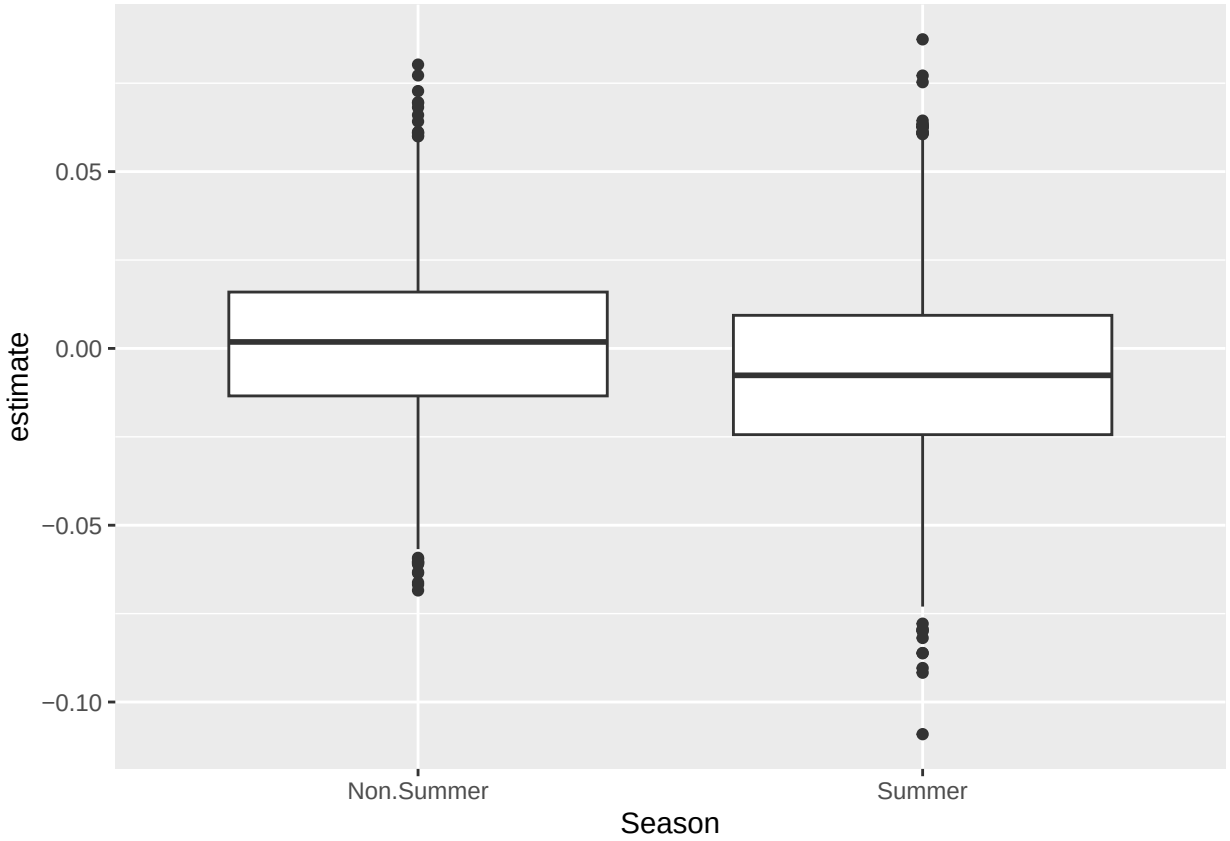
2 sample t test

```
t.test(df$Summer, df$Non.Summer, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  df$Summer and df$Non.Summer
## t = -14.46, df = 5998, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.009976697 -0.007594553
## sample estimates:
##    mean of x    mean of y
## -0.007350410  0.001435215
```

The two-sample t-test was used to compare the mean estimates of the seasonal effect of summer and non-summer in a dataset. The t-test results indicate a significant difference between the mean estimates of the two groups, with a t-statistic of -8.4429. The confidence interval for the difference in means between the two groups is (-0.006, -0.003), indicating that the mean estimate of the seasonal effect of summer is significantly lower than the mean estimate of the seasonal effect of non-summer. The negative values of the confidence interval suggest that the effect of summer is more negative than the effect of non-summer. Overall, these results suggest that the seasonality of the hurricane wind speeds in the North Atlantic area is influenced by the summer season, and that hurricane wind speeds are likely to be lower during summer than during non-summer.

```
month_df %>% ggplot(aes(x = Season, y = estimate)) +
  geom_boxplot()
```


0.05
0.00
−0.05
−0.10
estimate
Non.Summer
Summer
Season