Unveiling the Mysteries of Coral Bleaching: A Bayesian Approach

Authors: Tyler Pollard, Rachel Hardy, Hanan Ali

Date: May 6, 2024

1. Introduction

Coral bleaching is a pressing environmental concern that threatens the health of marine ecosystems. This study leverages Bayesian modeling to better understand the causality behind coral bleaching events in Florida. The dataset analyzed consists of 2,394 observations collected between 2006 and 2016 from the Florida Reef Resilience Program. Key factors such as temperature, turbidity, and spatial-temporal trends are examined to identify the primary drivers of coral bleaching.

2. Data Overview

The dataset consists of coral bleaching observations across multiple transects along the Florida coastline, collected by the Florida Reef Resilience Program from 2006 to 2016. The primary response variable is **Percent Bleaching**, measured as the proportion of coral bleached within each transect.

Data Preprocessing
The **distribution of Percent Bleaching** exhibits a **right-skewed shape**, with a substantial number of **zero bleaching observations**. Since beta regression requires values strictly between (0,1), **0% and 100% bleaching values were adjusted to 0.001 and 0.999**, respectively.

Spatial Structure
Coral bleaching observations were **geographically distributed across Florida's reef systems**. Mapping Percent Bleaching reveals **spatial clustering**, with certain areas experiencing more severe bleaching than others. These spatial effects were captured using a **tensor product spline (`t2(Lat, Lon)`)** to model continuous variation.

Temporal Structure
The dataset spans **2006 to 2016**, allowing an analysis of **temporal bleaching trends**. Boxplots of Percent Bleaching across years show **variation over time**, with city-specific differences. To account for **nonlinear temporal trends**, we incorporated a **Gaussian Process (`gp(Date_Year, by = City_Town_Name)`)**, allowing each city to have its own temporal trend.

3. Model Specification

To model the proportion of coral bleaching (mu_i), a Bayesian Beta regression
with a logit link function is used:

logit(mu_i) = f(Date_Year_i, City_Town_Name_i) + g(Lat_i, Lon_i) + X_i * beta

where:
- f(Date_Year_i, City_Town_Name_i) is a **Gaussian Process (`gp(Date_Year, by = City_Town_Name)`)**
  that models **nonlinear temporal trends** specific to each city.
- g(Lat_i, Lon_i) is a **tensor product spline (`t2(Lat, Lon)`)** that captures **spatial variation**.
- X_i * beta represents environmental predictors (e.g., Turbidity, Depth, Cyclone Frequency).
- The **Beta distribution** accounts for the bounded nature of bleaching proportions:

  y_i ~ Beta(mu_i * phi, (1 - mu_i) * phi)

where phi is the precision parameter.

4. Model Comparison

Multiple models were tested to determine the best approach for capturing
temporal and spatial variation in coral bleaching. The candidate models included:

- **Linear models** with Date_Year as a fixed effect.
- **Random effects models**, where City_Town_Name was treated as a random intercept.
- **Gaussian Process (GP) models**, both with and without city-specific trends.
- **Spatial models**, incorporating either **Lat and Lon as fixed effects** or a **smooth
  spatial term (`t2(Lat, Lon)`)**.

The final model was selected using **Leave-One-Out Cross-Validation (LOO-CV)**, ensuring it
provided the best balance between fit and complexity.

5. Model Comparison Results

Model	Temporal Structure	Spatial Structure	LOO Score
Model A	Linear (Date_Year)	None	-XXXX
Model B	Linear (Date_Year)	Lat/Lon as Fixed	-XXXX
Model C	Global GP	None	-XXXX
Model D	City-Specific GP	None	**Lowest LOO**
Final Model	City-Specific GP	Spatial Smoother (t2)	Slightly lower LOO than Mode

6. Final Model Code in `brms`

```
formulaBleaching_beta <- bf(
  PercentBleachingBounded ~
    gp(Date_Year, by = City_Town_Name) +
    t2(Lat, Lon) +
    Distance_to_Shore + Exposure + Turbidity + Cyclone_Frequency +
    Depth_m + Windspeed + ClimSST + SSTA + TSA + TSA_DHW,
  family = Beta(link = "logit")
)

priors <- c(
  prior(normal(0, 5), class = "b"),
  prior(gamma(2, 1), class = "gp", coef = "Date_Year"),
  prior(half_cauchy(0, 2), class = "gp", coef = "Date_Year"),
  prior(normal(0, 3), class = "sds", coef = "t2LatLon"),
  prior(gamma(0.1, 0.1), class = "phi")
)

betaFit <- brm(
  formulaBleaching_beta,
  data = procData2,
  prior = priors,
  chains = 4, iter = 2000, warmup = 1000,
  control = list(adapt_delta = 0.95)
)
```