

---

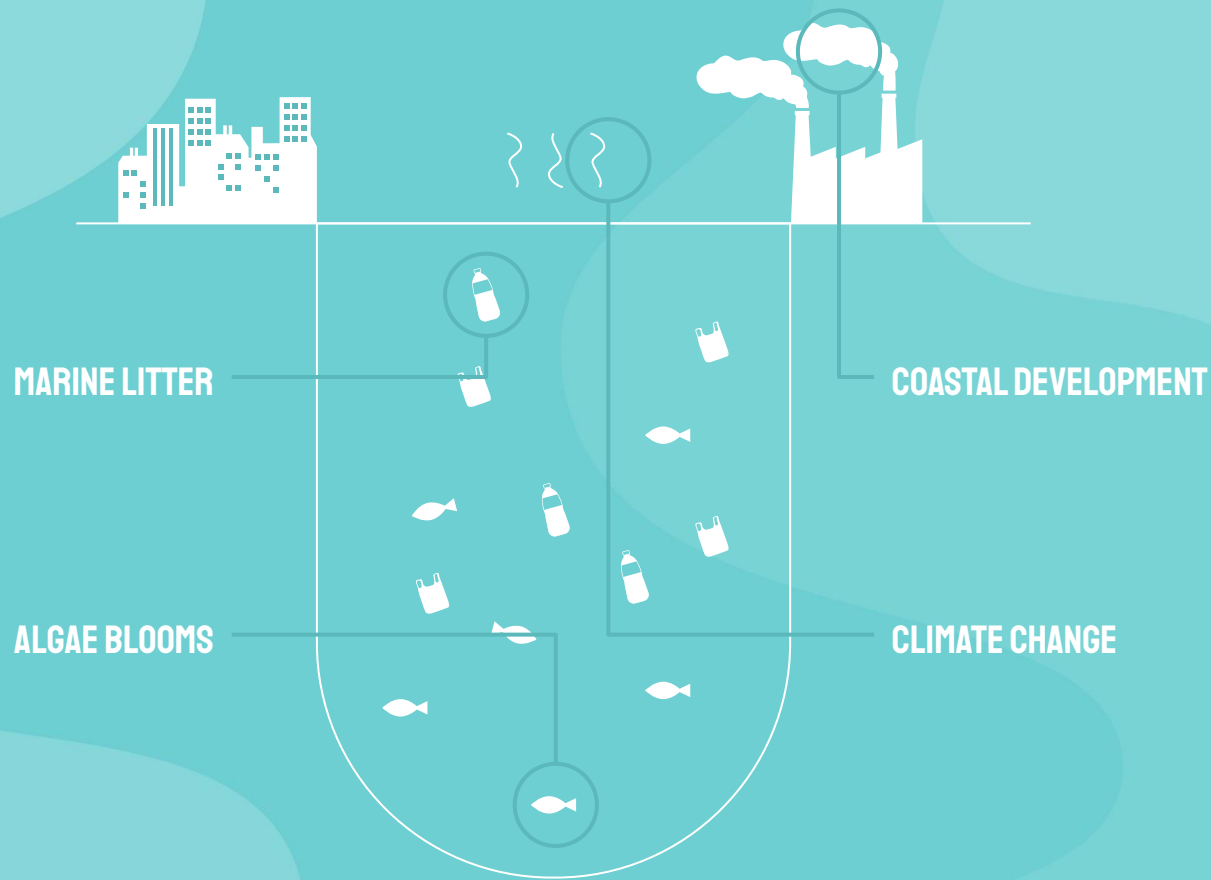
# ANALYZING THE ENVIRONMENTAL FACTORS OF U.S. BEACHES

Calleigh Smith | Duke University, STA 440L | Spring 2021

---



# THREATS TO AQUATIC BODIES



# PERSONAL MOTIVATION



## 1. Measuring Risk Factors

Can we predict the number of days under beach actions (beach-specific advisories or closings issued by the reporting state or local governments) based on a beach's characteristics? Similarly, are certain variables more valuable in predicting beach actions?

## 2. Measuring Pollution

How are the different types of pollution (algal, animal, sewer line, etc.) distributed across U.S. waters? Are certain regions more prone to certain pollution types of pollution than others?

## 3. Measuring Government Communication with Locals

How do local governments communicate with citizens when there is a beach advisory or closing? Are governments transparent in reporting the risks associated with local bodies of water to residents?

# DATA

- Environmental Protection Agency (EPA) - BEach Advisory and Closing Notification (BEACON)
- Tier 1 Beaches
  - Highest risk
  - Regularly monitored
  - Stratified sample based on bodies of water
- “A Criteria-Based Evaluation of Environmental Literacy Plans in the United States,” Karena Ruddiero, University of Tennessee, Knoxville (2016)



## West Coast and Gulf of Mexico Have Most Number of Days Under Action

Year: 2000



Number of Days Under Action

100 200 300

## Number of Pollutants Increases in California Since 2016

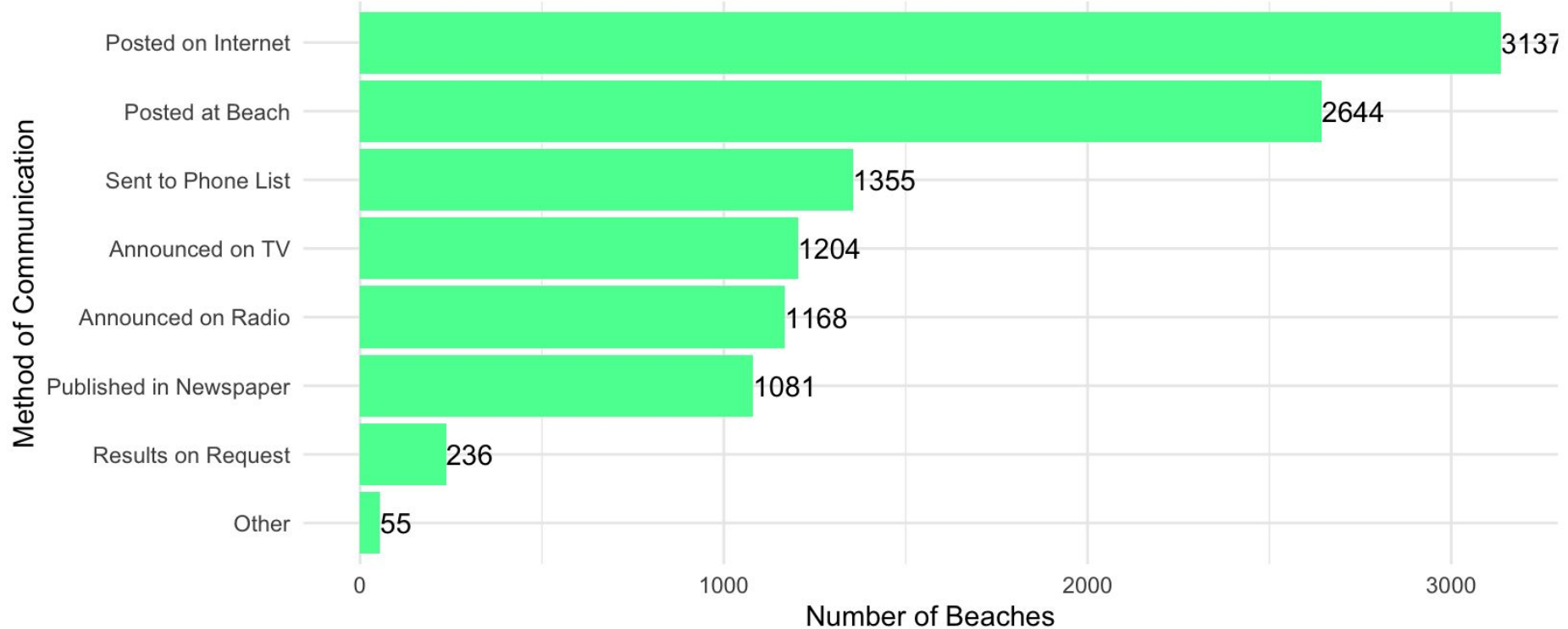
Year: 2000



Number of Pollutants

1.0 1.5 2.0 2.5 3.0

## Most Common Method of Communicating Beach Actions to Citizens Is via Posting on the Internet







## MEASURING RISK

- Outcome: number of days under beach action for most recent Tier 1 beach observations ( $n = 547$ )
- Features: beach length, beach ownership, water body, days in swim season
- Negative Binomial Model
  - Over-dispersion
  - Zero-inflated?
  - Satisfactory diagnostics
- Offset term for “eligible” beach actions?
- Goals: inference (primary) + prediction (secondary)

$$\begin{aligned} \mu_i = & \exp(\ln(t_i) + \beta_0 + \beta_1 \times \text{beach length}_i + \beta_2 \times I(\text{water body name}_i = \text{Gulf of Mexico}) + \beta_3 \times I(\text{water body name}_i = \text{Great Lakes}) \\ & + \beta_4 \times I(\text{water body name}_i = \text{Long Island Sound}) + \beta_5 \times I(\text{water body name}_i = \text{Pacific Ocean}) \\ & + \beta_6 \times I(\text{beach owner}_i = \text{Public}) + \beta_7 \times \text{days in swim season}_i). \end{aligned}$$

\*where  $i$  represents the  $i$ -th beach



# MEASURING POLLUTION

- Outcome: number of pollutants at a given beach from 2015-2020 (note: pollutants can return annually)
- Obstacle: pollutants are likely similar within the same body of water
- Random Intercepts Linear Mixed Model
  - Fixed Effects: beach length, beach ownership, water body, days in swim season
  - **Random Effect (grouping factor):** body of water

$$Y_{ij} = (\gamma_{00} + u_{0j}) + \gamma_{10}\text{Beach Length}_{ij} + \gamma_{20}I(\text{Beach Owner}_{ij} = \text{Public}) + \gamma_{30}\text{Swim Season Beach Days}_{ij}$$

\*where  $(i, j)$  represents the  $i$ -th beach for the  $j$ -th body of water and  $u_{0j}$  is the random effect term for the  $j$ -th body of water



## GOVERNMENT COMMUNICATION

- Shiny App
- For each state in mainland USA...
  - Proportion of beaches in that state which use the selected method of communication
  - Environmental literacy rating
- Note: most recent beach observations used to understand more up-to-date methods of communication



# MEASURING RISK

<i>Predictors</i>	<b>Number of Days Under Beach Action</b>		
	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
Intercept	15.4972	6.1023 – 39.3566	<b>&lt;0.001</b>
Beach Length (mi)	0.9919	0.9190 – 1.0706	0.835
Water Body (Great Lakes)	1.2297	0.7259 – 2.0831	0.442
Water Body (Gulf of Mexico)	2.2213	1.4461 – 3.4119	<b>&lt;0.001</b>
Water Body (Long Island Sound)	0.7765	0.3645 – 1.6541	0.512
Water Body (Pacific Ocean)	5.9923	4.0595 – 8.8452	<b>&lt;0.001</b>
Beach Owner (Public)	0.5630	0.2686 – 1.1802	0.128
Swim Season (day)	1.0010	0.9991 – 1.0029	0.314
Observations	547		
R <sup>2</sup> / R <sup>2</sup> adjusted	0.997 / 0.997		

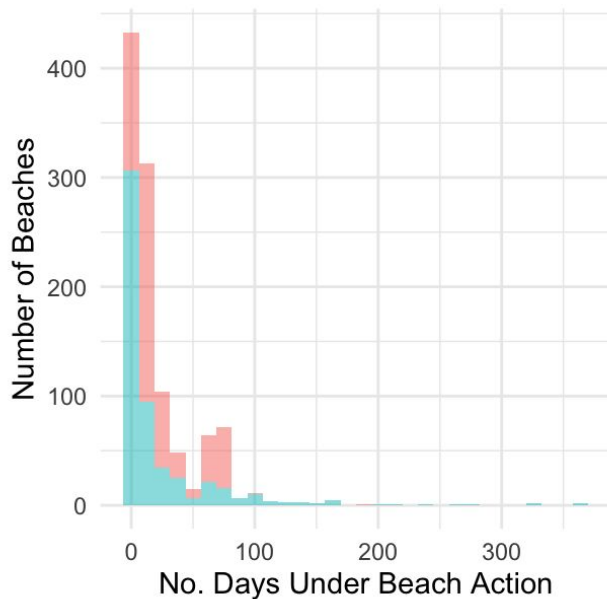


# MEASURING RISK

## Evaluating NB Model's Predictive Accuracy

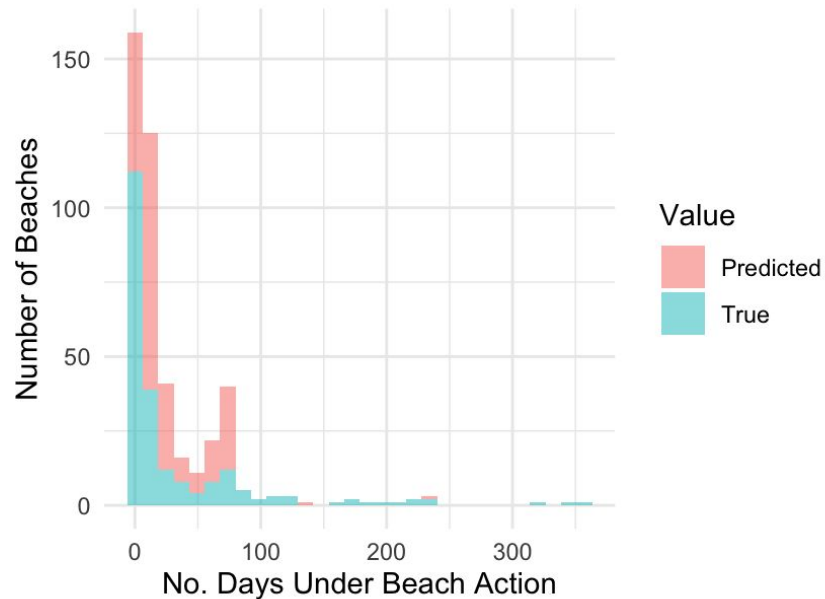
A

Training Set



B

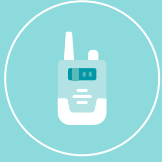
Testing Set





## MEASURING POLLUTION

- INCONCLUSIVE
- Range for number of pollutants over a five year period is between 5-10
  - Most beaches only have 1 pollutant/year
- Proposed LMM and other traditional modeling techniques not appropriate
- Remedies:
  - Data collection with pollution monitoring as primary goal
    - Consistent, descriptive, and quantitative measurements



## GOVERNMENT COMMUNICATION

- DEMO
- Commonly used forms of communication:
  - Posting on Internet (exceptions: CA, MA)
  - Posting on Beach (exceptions: CA)
- Less common forms of communication:
  - Results on Request (exceptions: LA, SC, MA)
  - Announced on TV (exceptions: NC, NY)
  - Announced on Radio (exceptions: PA, NY, NC)
- North Carolina:
  - Posted on Internet, Posted at Beach, Announced in Newspaper, Announced on Radio, Announced on TV
  - Good environmental literacy rating (63.75%, #10)

# **TAKEAWAYS**

**RISK IS MAINLY  
DEPENDENT ON BODY OF  
WATER AND PERHAPS  
BEACH OWNERSHIP**

**NEED BETTER DATA  
RELATED TO POLLUTANTS  
ACROSS U.S. BEACHES**

**GOVERNMENTS  
COMMUNICATE WITH  
CITIZENS DIFFERENTLY...  
SHOULD THIS BE MORE  
CONSISTENT?**



## LIMITATIONS

- Used Tier 1 beaches in the mainland U.S. (already at high risk) to build risk model
  - Generalizability
- Data is centered around beach closings and advisories and thus data relating to pollutants is not that reliable
- Despite use of stratification and randomization, spatial dependencies likely exist

## FUTURE DIRECTIONS

- Robust pollution data collection
- Used most recent observations for analysis, but temporal effects would be interesting
- What is the most effective means of governmental communication with locals?
  - Should communication be standardized?



# THANK YOU

---

This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**.