Ciguatera Fish Poisoning in Florida

Motti Goldberger[1] and Christopher Ryzowicz[2]

[1] University of Maryland, [2] New College of Florida

Graduate Assistant: Daniel Breininger, Dept. of Mathematical Sciences

Faculty Advisors: Prof. Michael Splitt, College of Aeronautics, Dr. Robert van Woesik, Dept. of Ocean Engineering and Marine Sciences, Dr. Nezamoddin N. Kachouie, Dept. of Mathematical Sciences

Florida Institute of Technology

Research Experiences for Undergraduates
National Science Foundation

Abstract

Ciguatera fish poisoning (CFP) is prevalent in southern Florida. We utilized monthly fish landings, sea surface temperature, degree heating week index, and storm intensity data to model ciguatera cases. Ultimately, we used a zero-inflated negative binomial model that predicts a PMF of CFP counts in each month.

Introduction

CFP is the most common fish-borne illness, affecting approximately 50,000 to 500,000 people in the world every year [1]. Ciguatoxin originates from dinoflagellates, which live on coral surfaces in reefs. Toxins accumulate in the fish that eat the algae and bioaccumulate up the food chain [2]. Humans eating fish with the ciguatoxin are susceptible to illness. Possible symptoms range from gastrointestinal to cardiovascular and, in the worst cases, neurological .

Datasets

- Monthly CFP calls from 1992-2019 (Figure 1)
- Commercial fish landings (Figure 2)
    - Grouper, Amberjack, Grunts, Hogfish, Snapper [3]
- Sea surface temperature from NOAA (Figure 2)
- Coral reef heat stress index from NOAA (Figure 2)
- Storm intensity metric over impact zone (Figure 2,3)

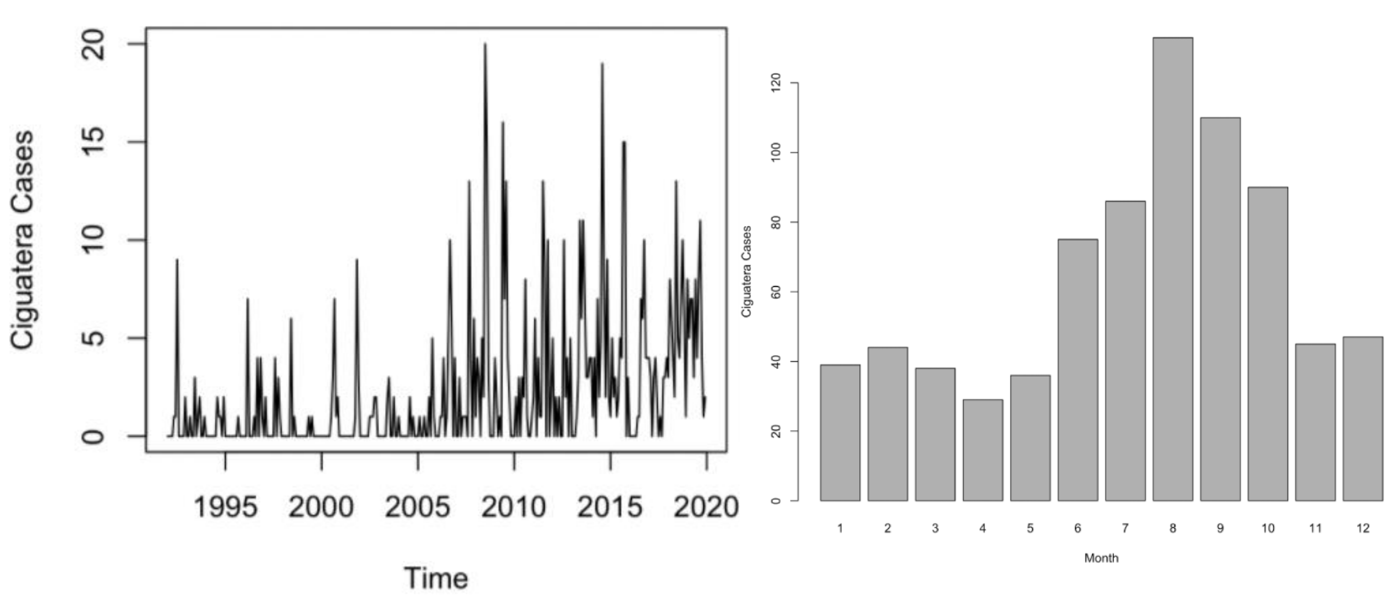

Fig. 1. (Left) CFP cases time-series (Right) Bar plot of monthly ciguatera cases

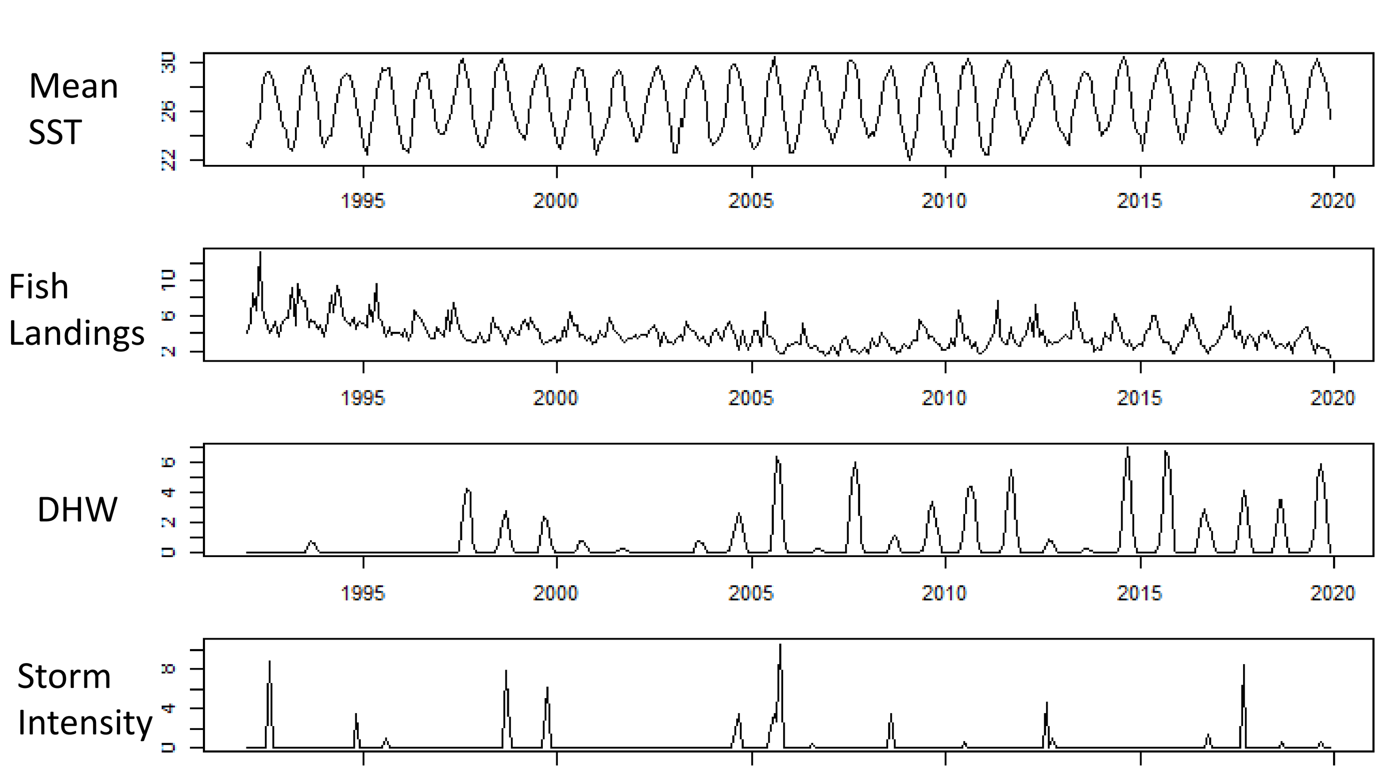

Fig. 2. Explanatory variables time-series

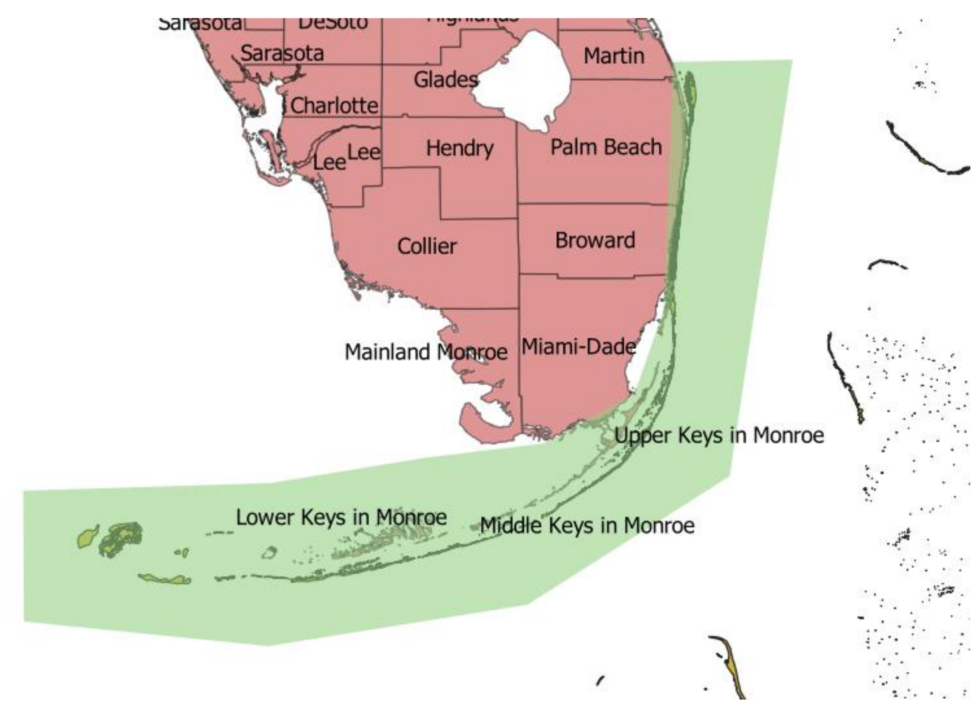

Fig. 3. South Florida storm impact zone

Methods

Data Collection ⟶ Modeling ⟶ Results and Discussion

Data Collection
- CFP data
- Storm intensity
- Mean SST
- DHW index
- Fish landings

Modeling
- Multiple regression
- Poisson regression
- Negative Binomial regression
- Zero-inflated negative binomial regression

Results and Discussion
- Identified significant factors
- How well our models can predict
- Future work

Fig. 4. Flow chart of modeling process

Count data regression models can be estimated for cases in which the response is a discrete quantitative variable that can assume non-negative values. Traditional linear regression models fail to consider the presence of discrete and non-negative values of the response. Wooldridge [4] indicates that a general regression model for count data can be described using Expression (1) as follows:

$$\lambda_i = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_K X_{Ki}), \quad (1)$$

where λ_i is the expected number of occurrences. β_0 is the intercept, β_i $(i = 1, 2, 3, \ldots, k)$ are the coefficients estimated for each predictor X_i ; k is the number of predictors in the model; and i indicates each observation of a given sample.

Poisson models assume: $E(Y_i) = Var(Y_i) = \lambda_i$

Negative Binomial model assumes overdispersion of the response variable, conditional to predictor variables [5], i.e., $\mu_i = E(Y_i) < Var(Y_i)$, where $E(Y_i) = \mu_i$ is mean, $Var(Y_i) = \mu_i + \varphi\mu_i^2$ is variance, and φ represents overdispersion in the count data.

$$\mu_i = \exp(\beta 0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_K X_{Ki}).$$

Zero-inflated negative binomial regression is a model that considers excess zeros in data. It does so by having two models, a zero inflated model and a count model. The zero inflated model gives a probability that a given data point is an excess zero. The count model is a negative binomial model that is a posterior model based on the probability of excess zero produced by the zero inflated model.

Results

We tested all combinations of variables and interaction terms to acquire the optimal model based on AIC and residuals.

Optimal Zero-Inflated Negative Binomial Model:

Count			
	Predictors	Coefficient	P-value
	Fish Weight	-0.610	<0.001
	MeanSST	0.310	<0.001
	Storm Intensity	-0.937	0.003
	DHW*Storm	-0.180	0.014
	Storm*Fish Weight	-1.16	0.002
Zero-Inflated			
	MeanSST	-0.71	0.0015

AIC: 1223.526
Sum of Residuals: 704.43

Optimal Negative Binomial Model:

Predictors	Coefficients	P-values
MeanSST	0.53	<0.001
Storm Intensity	-0.78	0.0095
Fish Weight	-0.56	<0.001
Storm*Fish Weight	-0.63	0.21

AIC: 1235.8
Sum of Residuals: 718.91

Table 1. (Left) Zero-inflated model output coefficients and p-values. (Right) Negative binomial model output coefficients and p-values

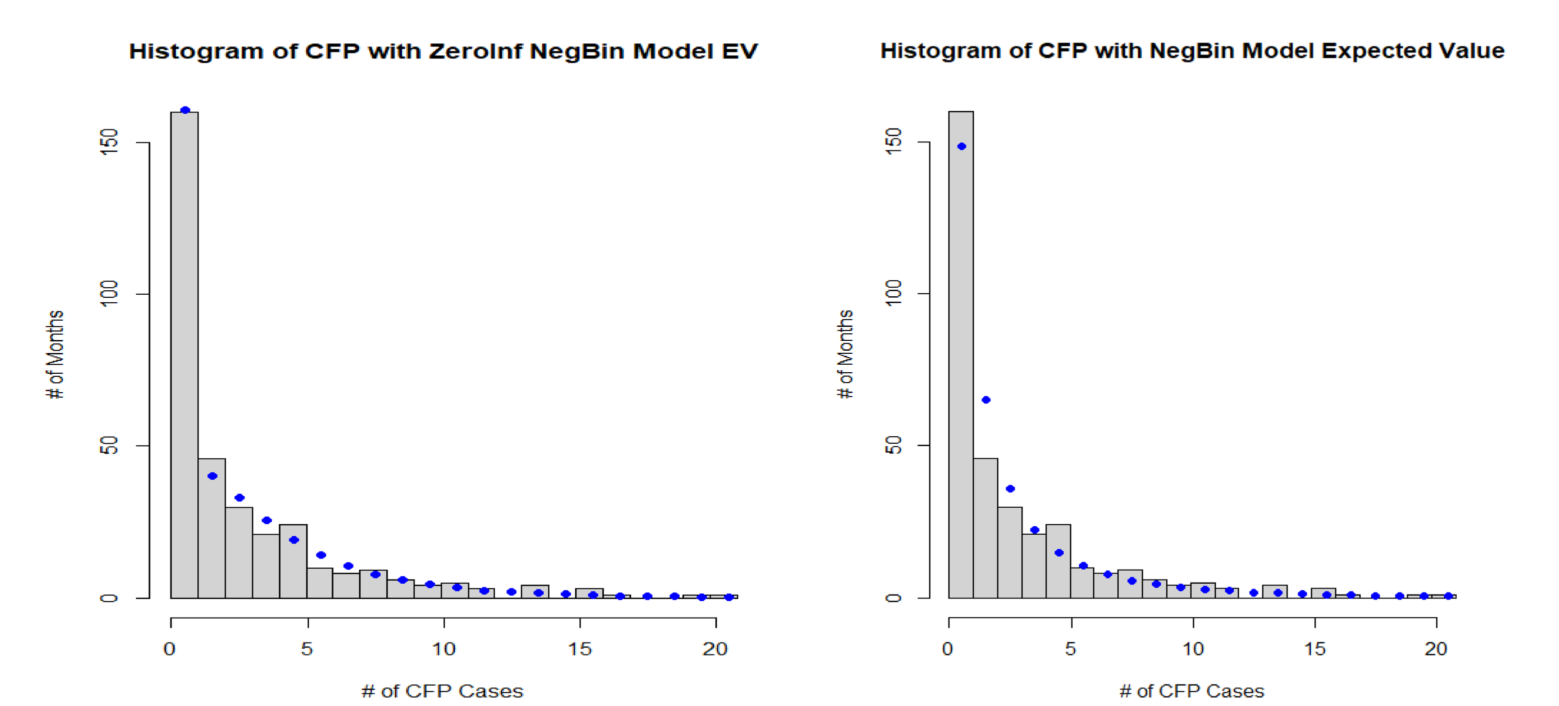

Fig. 5. (Left) Rootogram of optimal ZINB model (Right) Rootogram of optimal NB model

Conclusions and Discussions

We found mean SST has a strong positive relationship with CFP cases in Florida. In contrast, storm intensity has a negative relationship with CFP cases. Our ZINB model is the best model in terms of quality of fit (Table 1). It can predict a PMF of CFP cases in a given month. There is an upward trend in CFP cases that is not completely captured by our model. Thus, there are additional underlying climate factors that we did not account for that influence the number of CFP cases. More research needs to be done to determine what these factors might be.

References

[1] Gingold, D.B., Strickland, M.J. and Hess, J.J., 2014. Ciguatera fish poisoning and climate change: analysis of National Poison Center data in the United States, 2001–2011. *Environmental Health Perspectives*, *122*(6), pp.580-586.

[2] Barrett, J.R., 2014. Under the weather with ciguatera fish poisoning: Climate variables associated with increases in suspected cases.

[3] Ault, J.S., Smith, S.G. and Bohnsack, J.A., 2005. Evaluation of average length as an estimator of exploitation status for the Florida coral-reef fish community. *ICES Journal of marine Science*, *62*(3), pp.417-423.

[4] Wooldridge, J.M. Econometric Analysis of Cross Section and Panel Data, 2nd ed.; MIT Press: Cambridge, MA, USA, 2010; p. 1095.

[5] Hilbe, J.M. Modeling Count Data; Cambridge University Press: Cambridge, MA, USA, 2014; p. 300.