Generalized Additive Models for Categorical Count Data:

An Exploration of the Distribution of Queen Triggerfish (*Balistes vetula*) in the REEF Dataset

John Urquhart, Donald Olson, Elizabeth Babcock, Kathleen Sealey

Queen Triggerfish(*Balistes vetula*)

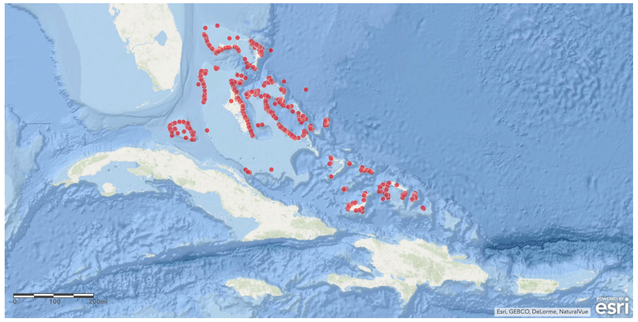
- Species threatened and in decline (IUCN)
- Local food fishery, but limited fisheries data
- Easily identifiable

B. vetula

REEF Data

- Logarithmic categorical (SFMA) : Single = 1, Few = 2-10, Many = 11-100 , Abundant = +100
- Categorical data is difficult to model over, and difficult to interpret
- Data requires modeling to remove observer and environmental effects

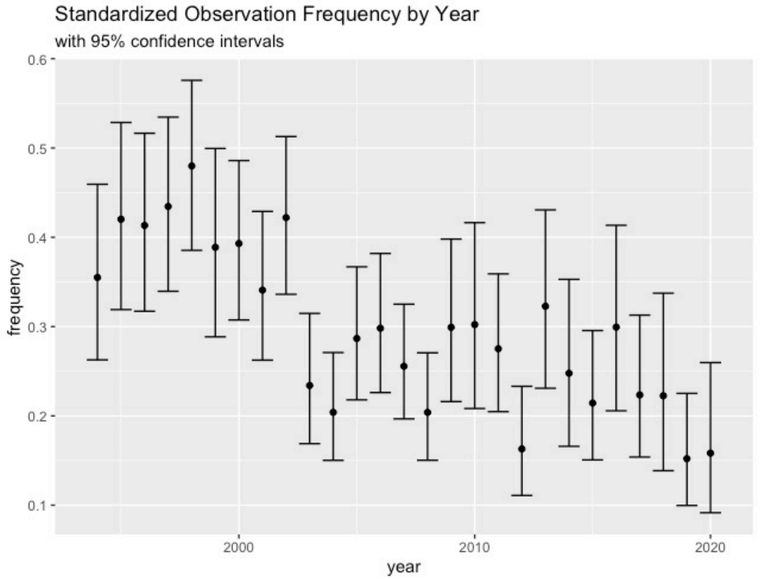
Map of REEF dive sites



Methods

- Count data was converted to numeric using established methodology, then fit with GAM
- Negative binomial model trained on simulated count data, both categorized and uncategorized
- Presence absence (Binomial), abundance (negative binomial), and multinomial model trained on REEF data

Results



- Simulated data: Model fit to categories only slightly underperformed model fit to counts
- REEF data: Abundance model produced similar results to presence-absence, was able to account for ~17% of variance in data.
- Abundance model was slightly worse at predicting correct categories than multinomial model, but is easier to interpret and can be fit with more variables
- Significant decline in both standardized observation frequency and standardized relative abundance score over years sampled

Conclusions

- Abundance model was able to fit to categorical data with minimal bias on simulated data, showing methodology is effective
- Models performed well on actual REEF data, found significant decline in species abundance and frequency over last 20 years, with implications for the conservation of this species

Acknowledgments

I would like to thank Dr. Donald Olson for guiding me through this project, and for his patience with my writing during the editing of this paper. I would like to thanks Dr. Elizabeth Babcock for technical assistance, and inspiration for many of the above analyses and insights. My thanks to Dr. Kathleen Sullivan Sealey for access to REEF data. I would like to thank Dr. Michael Schmale for introducing me to the potential of the REEF dataset. Finally, I'd like to thank my parents, for sharing their love of diving and constant support for interests in aquatic sciences.