Habitat models for Gray Triggerfish collected in fishery-independent trap surveys off the southeastern United States

Barbara A. Muhling, Mitchell A. Roffer, Joseph C. Ballenger, Marcel J. M. Reichert,
Roger Pugliese, Tracey I. Smart

SEDAR41-DW11

Submitted: 21 July 2014

Please cite this document as:

Muhling, B. A., M. A. Roffer, J. C. Ballenger, M. J. M. Reichert, R. Pugliese, and T. I. Smart. 2014. Habitat models for Gray Triggerfish collected in fishery-independent trap surveys off the southeastern United States. SEDAR41-DW11. SEDAR, North Charleston, SC. 18 pp.

Habitat models for Gray Triggerfish collected in fishery-independent trap surveys off the southeastern United States

Barbara A. Muhling[1, 2], Mitchell A. Roffer[3], Joseph C. Ballenger[4], Marcel J. M. Reichert[4], Roger Pugliese[5], Tracey I. Smart[4]

1: Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami FL

2: Southeast Fisheries Science Center, NOAA, Miami, FL

3: Roffers Ocean Fishing Forecasting Service, Melbourne, FL

4: South Carolina Department of Natural Resources, Charleston, SC

5: South Atlantic Fishery Management Council

SEDAR41-DW-11

Abstract

Catches of Gray Triggerfish from trap surveys in the US South Atlantic currently are used to formulate a fishery-independent index of abundance. However, the impact of environmental variability on catch efficiencies and detectability is not currently well known. In this study, a predictive habitat model for Gray Triggerfish was developed using a boosted regression tree model, incorporating 11 spatiotemporal, environmental and biological predictor variables. Results showed that bottom temperature was the most important factor influencing the occurrence of Gray Triggerfish in traps, with intrusions of cold water from upwelling events leading to low catch rates. Annual time series of predicted and observed proportions of positive traps were in good agreement, including the successful prediction of low occupancy during a cold upwelling event in 2003. Future work will focus on investigating the feasibility of formally incorporating these results into the current index standardization.

Introduction

Gray Triggerfish (*Balistes capriscus*) are widely distributed on hard-bottom habitat along the eastern and western coasts of the Atlantic Ocean. In the western Atlantic, they are found from Nova Scotia in the north through to Argentina in the south, including the Gulf of Mexico (Robins & Ray, 1999), although they are comparatively rare north of Cape Cod (Collette & Hartel, 1988). Commercial and recreational interest in this species increased during the 1990s (Harper & McClennan, 1997), and they are currently managed in the US South Atlantic region through an Annual Catch Limit (ACL).

Fishery-independent trap surveys for reef fish species have been conducted between Cape Hatteras and St. Lucie Inlet, Florida, by the Southeast Reef Fish Survey (SERFS) program since the early 1970s. Prior to 2009, SERFS was funded exclusively by the Marine Resources Monitoring, Assessment, and Prediction program (MARMAP). Traditional MARMAP funding for the SERFS was enhanced by additional funding made available by the Southeast Area Monitoring and Assessment Program – South Atlantic (SEAMAP-SA) in 2009, and the Southeast Fishery Independent Survey (SEFIS) in 2010. The use of chevron traps for monitoring purposes commenced in 1990, and has continued up to the present, with surveys usually completed between May and September each year. Catches of Gray Triggerfish from these traps are used to formulate a fishery-independent index of abundance (e.g., Ballenger et al., 2013).

Periodic topographically-induced summer upwelling has been observed in the southern portion of the trap survey area, causing intrusions of cold water onto the continental shelf (Blanton et al., 1981). These events have the potential to displace cold-intolerant fish species (e.g. Buxton & Smale, 1989), and thus reduce the availability of the regional population to the trap surveys. Bottom temperature from CTD casts currently has been included in the fishery-independent index standardization as a coarsely binned variable (<20°C, 21 – 25°C, >25°C, Ballenger et al., 2013). However, the nature of the relationship between Gray Triggerfish and bottom temperature is not currently well known. In addition, interactions between spatiotemporal and environmental effects have not been fully explored.

In this study, we aimed to develop a predictive habitat model for Gray Triggerfish in the US South Atlantic region, incorporating spatiotemporal and environmental variables extracted from various data sources. The relative importance of these variables for driving observed distributions of Gray Triggerfish were defined, and the potential application of these results to the current fishery-independent index was explored.

Methods

Trap Sampling

Chevron traps have been deployed in a consistent manner in the survey region since 1990, mostly through the MARMAP program, but also supplemented by the SEAMAP-SA Reef Fish Survey and the SEFIS program in more recent years. Ballenger et al. (2013) provides a detailed description of sampling protocols and methods, and so only the essential points are repeated here.

Traps were deployed at randomly selected sampling stations, chosen from a database of >2,000 known live/hard bottom sites. Several surveys were completed each year between March and October, with most effort concentrated between May and August. Water depths ranged from 13 to 216m, however the vast majority of stations were situated in 100m of water or less. Traps were set in groups (generally of 6), with a minimum distance between sampling stations of 200m. Soak times ranged from 49 to 150 minutes. Catches of Gray Triggerfish were recorded as both abundances per trap, and by weight.

Environmental Variables

Temperature and salinity were measured throughout the water column *in situ* using a SeaBird SBE-19 or SBE-25 CTD. Oxygen was also available when the SBE-25 was used. For this study, we used temperature at the surface and at the bottom, and salinity and oxygen at the bottom as predictor variables. Positional variables included were latitude, longitude and water depth (m). Variability in sampling methods was described by the soak duration (minutes), and the deployment time of day. Day of the year also was included as a predictor variable.

Mean daily wind speeds and wave heights from buoy 41010 (78.464W, 28.903N) were extracted for the date of each trap deployment, and also were included in the predictive model. Future models may investigate the utility of using a spatial product such as the AVISO global gridded wind speed dataset instead. The last environmental predictor considered was the moon phase, defined as the fraction of the moon illuminated by date (%). These data were obtained from the US Navy Observatory.

The interaction of different species within traps is a potential contributor to observed within-trap assemblages. To investigate the effect of larger, predatory species on trap catches, the combined biomass in kilograms of five potential predators (Red Grouper, Snowy Grouper, Gag, Scamp and Red Snapper) also was included as an explanatory variable.

Although the habitat model structure selected does not require normally distributed predictor variables, highly skewed data can still affect model performance. Water depth, wind speed, wave height and predator biomass were highly right skewed, and so were $\log_{10}$ transformed before analysis ($\log_{10}$ (x+1) transformed in the case of predator biomass). In order to reduce over-specification of the habitat model from multicollinearity, the Pearson correlation coefficient among all predictor variables was first calculated. Where ρ was greater than 0.6 among any pair of predictor variables, one variable was selected for removal. The adequacy of this cutoff was also assessed by calculating the Variance Inflation Factors (VIFS) for each of the remaining predictor variables, using SigmaStat software.

Habitat Models

Habitat models were parameterized to predict the presence or absence of Gray Triggerfish in each chevron trap, using an optimum combination of all environmental variables. We chose to use boosted regression tree models, which were developed in DTREG software (Sherrod, 2003). These models are non-parametric, and can model highly non-linear relationships and interactions, including moderately zero-inflated data (Elith et al., 2008). Misclassification costs can be adjusted within presence/absence models, to give greater weight to placing positive occurrences in good habitat, which is useful for organisms with sub-optimal catch efficiencies (e.g. Muhling et al., 2012). Models were validated using v-fold cross validation (Sherrod, 2003), which divides the dataset into a pre-determined number of subsets (in this case, 8), and creates several different models, holding back one subset each time for validation. The final model for each run was selected automatically by maximizing the area under the Receive Operating Characteristic (ROC) curve (AUC). The misclassification cost was then adjusted iteratively over several model runs, until >80% of positive sampling stations were placed in predicted favorable habitat. The importance of each predictor variable to the final model was calculated by its importance in splitting data within trees, and was ranked out of a maximum possible score of 100.

The error of each model was quantified using probability calibration (Sherrod, 2003). The predicted probability of occurrence for each model was binned in 10% increments (i.e. 0 – 10% probability of occurrence, 10 – 20% etc), and the predicted and observed proportions of positive samples were compared. For example, if the mean predicted probability of occurrence within the 10 – 20% bin was 0.15, but the observed proportion of positive stations was 0.17, the error for that bin would be 0.02. Results were plotted in space using kriging, which was completed in Surfer 9 (Golden Software).

Remotely sensed vs. in situ variables

A disadvantage of using *in situ* environmental variables is the low spatial and temporal coverage of observations. Remotely-sensed satellite data are available for a wider range of locations and times, but are limited to surface observations. Ocean models provide data high resolution outputs throughout the water column, but their ability to reproduce observed conditions varies. To evaluate the potential of the global HYCOM 1/12° Model to replicate observations within the study area, the modeled bottom temperature was extracted for the same date and position of each CTD cast location completed since September, 2008 (the current available HYCOM data range). Predicted and observed bottom temperatures were then compared.

Results

Correlations among the 13 predictor variables were generally low, with the exception of longitude vs. latitude (ρ = 0.79), and yearday vs. surface temperature (ρ = 0.71) (**Table 1**). Longitude and yearday were therefore not included in the habitat model. VIFs of the remaining 11 variables were all low (<1.7).

The sensitivity of the habitat model was 82.1% during training, with a specificity of 69.5%, using a misclassification cost of 3 (**Table 2**). The area under the ROC curve was 0.85. The model lost some positive predictive power between training and validation, with a validation sensitivity of 74%, however specificity remained similar at 65.4%, with an AUC of 0.78. Of the 11 predictor variables included, bottom temperature was the most important, with a weight of 100, followed by latitude (78.0), water depth (62.9) and predator biomass (38.5). Most trap samples were assigned a moderate – low probability of Gray Triggerfish occurrence, with few samples showing a predicted probability of occurrence >50% (**Figure 1**). As a result, the error of the model increased with increasing probabilities of occurrence, exceeding 10% at the upper range. This suggested that the habitat model defined negative habitat very reliably, and positive habitat somewhat less so.

Partial probabilities of occurrence plotted against the four most influential environmental predictors showed the strong relationship between Gray Triggerfish and bottom temperature (**Figure 2**). Probabilities of occurrence were less than 20% unless temperatures were >20°C. Habitat favorability then increased with temperature to a peak at 27°C before declining at very high temperatures. The model generally reproduced the observed relationship well, apart from over-predicting slightly at lower temperatures. The habitat model also reproduced the positive relationship between probabilities of occurrence and latitude, and showed the rarity of Gray Triggerfish in waters deeper than 90-100m (**Figure 2**). Somewhat surprisingly, the relationship between predator biomass and probability of occurrence was positive, particularly where predator biomass was at the upper end of the observed range.

Bottom temperature in the study region varied among years, with >10°C difference between two years observed across the region (**Figure 3**). Bottom temperatures during June - August 2002 were warm across much of the continental shelf, usually >22°C. In contrast, bottom temperatures during June – August 2003 were much cooler, <18°C across most of the outer shelf. These cooler temperatures were clearly associated with reduced Gray Triggerfish catches, and much lower predicted probabilities of occurrence (**Figure 3**). During spring (April – June), surface temperatures ranged between 16.6 and 29°C, while bottom temperatures were between 12.7 and 27.4°C (**Figure 4**). During summer (July – August), surface temperatures were warmer, between 23.7 – 30.1°C, with 85% of stations >27°C. However, bottom temperatures as cold as 9.4°C were recorded, with coldest values at stations deeper than 30m (**Figure 4**). Differences between surface and bottom temperatures sometimes exceeded 15°C during summer, indicating upwelling of cold water onto the shelf.

The modeled and observed probabilities of occurrence for Gray Triggerfish across all traps were generally quite comparable (**Figure 5**). Both modeled and observed probabilities were low during 1990 and 2003, and higher during 1991, 1995 and 1997. The observed time series was mostly within the error boundaries of the habitat model, except for in 1990 (model over-estimated) and 2002 (model under-estimated). Observed probabilities of occurrence were generally higher than the model predicted during the early portion of the time series (1991 – 1997), and lower than the model predicted in more recent years (2006 – 2011).

The current treatment of bottom temperature within the Gray Triggerfish index standardization involves binning into three categories: <20°C, 21 - 25°C, >25°C (Ballenger et al., 2013). When these categories were overlaid with observed and modeled probabilities of occurrence from this study, the potential for adjusted bin limits was evident (**Figure 6**). While the lowest bin generally contained the lowest probabilities of occurrence (<20%), the middle and upper bins both contained moderate and high probability samples.

The linear correlation between observed and modeled (HYCOM) bottom temperatures was positive, and significant (R^2 = 0.301, p<0.001). However, the variability and error (up to 10°C) between the two datasets suggested that modeled bottom temperatures from HYCOM are currently not accurate enough to reproduce observations. This result points to the potential need for use of a high-resolution nested oceanographic model for this purpose.

Discussion

Results from this study suggest that the primary factor influencing Gray Triggerfish catches in chevron traps in the US South Atlantic during recent decades is bottom temperature. Gray Triggerfish were rarely collected where bottom temperatures were <20°C, and were not recorded where temperatures were <14°C. These results are consistent with the documented affinity of Gray Triggerfish with warm water, and their potential sensitivity to cold-water episodes (Aggrey-Fynn, 2007; Simmons & Szedlmayer, 2012; Kacem & Neifar, 2014). Some species of triggerfish have been observed to migrate offshore to escape coldwater intrusions onto their usual benthic habitat (Essuman & Diakit, 1990). However, studies on Gray Triggerfish in the Gulf of Mexico show very high site fidelity, and low average movements (Ingram and Patterson, 2001). It is not clear at this stage whether the low catches of Gray Triggerfish at low temperatures are due to fish moving on/offshore, migrating upwards in the water column off the sea floor, or remaining in the area but lowering their activity levels, and thus their susceptibility to traps. However, the end result appears to be a lowered availability to chevron traps during cold events, leading to low abundance anomalies (e.g. 2003), which do not track expected stock biomass trends. The cold events observed were likely the result of periodic upwelling induced by the interaction of the Gulf Stream with topographical features (Blanton et al., 1981).

Including these episodes in index standardization may be complex. Although bottom temperature is included in the current fishery-independent index standardization, results presented here suggest that the bins currently used for this variable may not be capturing the responses of Gray Triggerfish effectively. Beyond this, while upwelling events may not affect the stock biomass of Gray Triggerfish, they are very likely to impact the availability of the stock to the trap survey. Studies on other species in other part of the US east coast have investigated more complex methods to incorporate bottom temperatures into management processes and indices. Kohut et al. (2013) used modeled, high resolution bottom temperature predictions to assess changes in suitable habitat for butterfish (*Peprilus triacanthus*) in the Mid Atlantic Bight, and thus the availability of the stock to bottom trawl surveys. A similar approach could be used for the US South Atlantic; however a high-resolution regional bottom temperature model would likely be required. A comparison completed here between HYCOM global 1/12° bottom temperatures and *in situ* bottom temperatures measured by CTDs suggested that HYCOM outputs would not be sufficient for habitat delineation purposes at this point. The observed discrepancies between *in situ* and HYCOM bottom temperatures are likely due to both model inaccuracies, and potentially to the relatively coarse vertical resolution of HYCOM model vertical layers in deeper waters.

Future research activities on this project will focus on the feasibility of including outputs from the habitat model in the Gray Triggerfish fishery-independent index standardization. Obstacles to overcome include addressing problems of independence created by including habitat model outputs, and assessment of different methods of quantifying error. If a successful approach is defined, it may be extended to other species of commercial interest collected in trap surveys, such as Black Sea Bass, Red Porgy and Vermillion Snapper.

References

Aggrey-Fynn, J. (2007). The fishery of *Balistes capriscus* (Balistidae) in Ghana and possible reasons for its collapse. Doctoral dissertation, University of Bremen, 122pp.

Ballenger, J., T. Smart, K. Kolmos, and M. Reichert. 2013. Trends in relative abundance of Gray Triggerfish in waters off the SE US based on fishery-independent surveys. SEDAR32-DW04. SEDAR, North Charleston, SC. 77 pp.

Blanton, J. O., Atkinson, L. P., Pietrafesa, L. J., & Lee, T. N. (1981). The intrusion of Gulf Stream water across the continental shelf due to topographically-induced upwelling. Deep Sea Research Part A. Oceanographic Research Papers, 28: 393-405.

Buxton, C. D., & Smale, M. J. (1989). Abundance and distribution patterns of three temperate marine reef fish (Teleostei: Sparidae) in exploited and unexploited areas off the southern Cape coast. Journal of Applied Ecology 26: 441-451.

Collette, B. B. & Hartel, K.E. (1988) An annotated list of the fishes of Massachusetts Bay. NOAA Technical Memorandum NMFS-F/NEC-51. 80pp.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. Journal of Animal Ecology, 77: 802-813.

Essuman K. & Diakit B. (1990). Utilization of triggerfish (*Balistes carolinensis* = *B. capriscus*) in Ghana and Senegal. FAO Fisheries Circular No. 828. 23 pp.

Harper, D., & McClellan, D. (1997). A review of the biology and fishery for Gray Triggerfish, *Balistes capriscus*, in the Gulf of Mexico. National Marine Fisheries Service, Southeast Fisheries Science Center, Miami Laboratory Contribution MIA-96/97-52, Miami, Florida.

Ingram Jr., G.W. & Patterson III, W.F. (2001) Movement patterns of Red Snapper (*Lutjanus campechanus*), Greater Amberjack (*Seriola dumerili*), and Gray Triggerfish (*Balistes capriscus*) in the Gulf of Mexico, and the utility of marine reserves as management tools. Proceedings of the Gulf and Caribbean Fisheries Institute 52: 686-699.

Kacem, H., & Neifar, L. The reproductive biology of the grey triggerfish *Balistes capriscus* (Pisces: Balistidae) in the Gulf of Gabès (south-eastern Mediterranean Sea). Journal of the Marine Biological Association of the United Kingdom, doi:10.1017/S0025315414000824.

Kohut, J., Palamara, L., Curchitser, E., Manderson, J., DiDomenico, G., & Oliver, M. J. (2013, June). Toward dynamic marine spatial planning tools: An example with butterfish assessment in the mid-Atlantic bight. In OCEANS-Bergen, 2013 MTS/IEEE (pp. 1-6). IEEE.

Muhling, B. A., Roffer, M. A., Lamkin, J. T., Ingram Jr, G. W., Upton, M. A., Gawlikowski, G., ... & Richards, W. J. (2012). Overlap between Atlantic bluefin tuna spawning grounds and observed Deepwater Horizon surface oil in the northern Gulf of Mexico. Marine Pollution Bulletin, 64: 679-687.

Robins, C. R., & Ray, G. C. (1999). A field guide to Atlantic Coast fishes: North America (Vol. 32). Houghton Mifflin Harcourt.

Sherrod PH (2003) DTREG: Classification and regression trees for data mining and modeling. www.dtreg.com/DTREG.pdf [June 2014].

Simmons, C. M., & Szedlmayer, S. T. (2012). Territoriality, reproductive behavior, and parental care in Gray Triggerfish, *Balistes capriscus*, from the northern Gulf of Mexico. Bulletin of Marine Science 88: 197-209.

Table 1: Pair-wise Pearson correlations among all predictor variables selected as candidates for inclusion in a predictive habitat model for Gray Triggerfish. Correlations >0.6 are highlighted.

	Water depth (log)	Latitude	Longitude	Yearday	Time of day	Bottom temperature	Bottom salinity	Bottom oxygen	Predator biomass (log)	Moon phase	Wind speed (log)	Wave height (log)
Latitude	-0.19											
Longitude	-0.02	**0.79**										
Yearday	-0.01	0.05	0.02									
Time of day	-0.06	0.03	-0.01	-0.01								
Bottom temperature	-0.43	0.31	0.11	0.40	0.01							
Bottom salinity	0.13	-0.01	0.06	0.04	-0.04	-0.02						
Bottom oxygen	0.02	-0.39	-0.38	-0.07	-0.01	0.01	0.00					
Predator biomass (log)	0.26	-0.02	0.06	-0.03	0.00	-0.04	0.05	0.09				
Moon phase	0.04	0.02	0.00	0.00	0.00	0.03	0.02	-0.03	0.00			
Wind speed (log)	-0.07	-0.03	-0.03	-0.13	0.01	-0.08	-0.01	0.07	-0.02	0.01		
Wave height (log)	-0.02	-0.18	-0.11	-0.16	-0.03	-0.04	-0.01	0.24	0.02	-0.07	0.28	
Surface temperature	0.23	-0.14	-0.05	**0.71**	0.04	0.24	0.11	0.02	0.03	0.06	-0.19	-0.19

Table 2: Specifications and predictor variable importance for Gray Triggerfish habitat model, built using a boosted classification tree.

Model parameters	
Training sensitivity (%)	82.1
Training specificity (%)	69.5
Training AUC	0.85
Validation sensitivity (%)	74.0
Validation specificity (%)	65.4
Validation AUC	0.78
Variable Importance	
Bottom temperature	100
Latitude	78.0
Water depth (log)	62.9
Predator Biomass (log)	38.5
Time of day	35.2
Surface temperature	33.8
Soak Time	29.9
Wave height (log)	26.9
Wind speed (log)	25.9
Moon phase	25.0
Bottom salinity	15.9

Figure 1: Percentage of data points (open diamonds), and model error (closed circles) from boosted regression tree predictive habitat model for Gray Triggerfish. Parameters are shown based on predicted probabilities of occurrence from the habitat model, binned to 10% intervals.

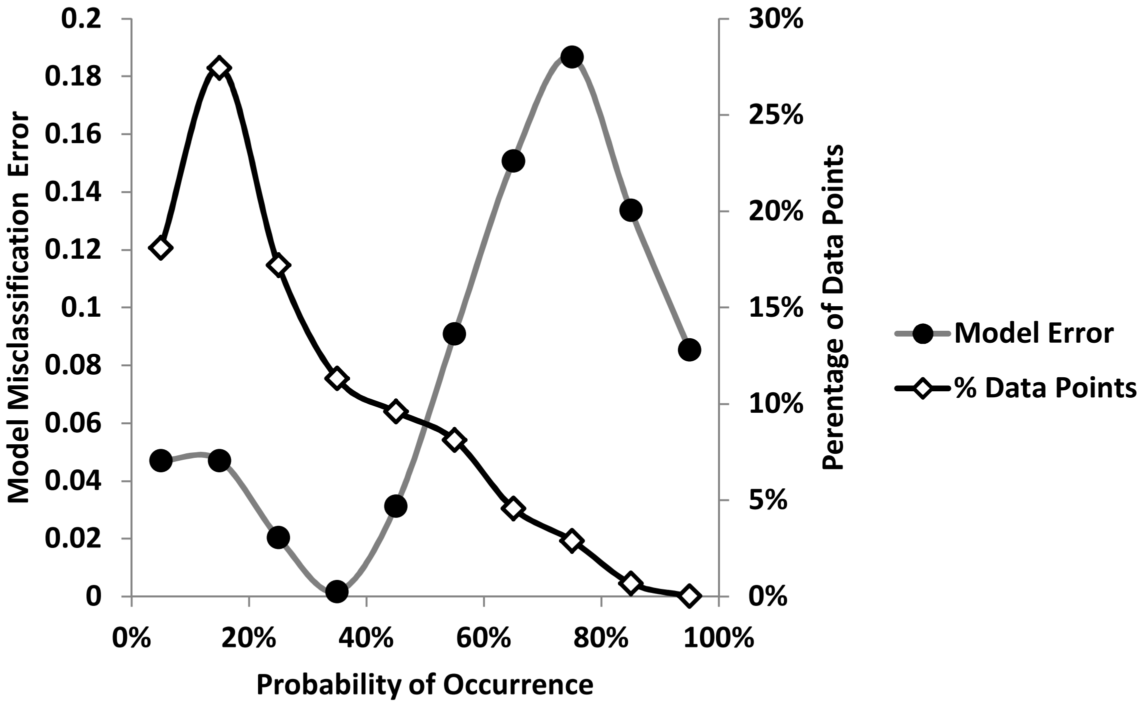

Figure 2: Partial plots of predicted and observed probability of occurrence of Gray Triggerfish vs. the four most influential environmental predictors, as shown by the habitat model results.

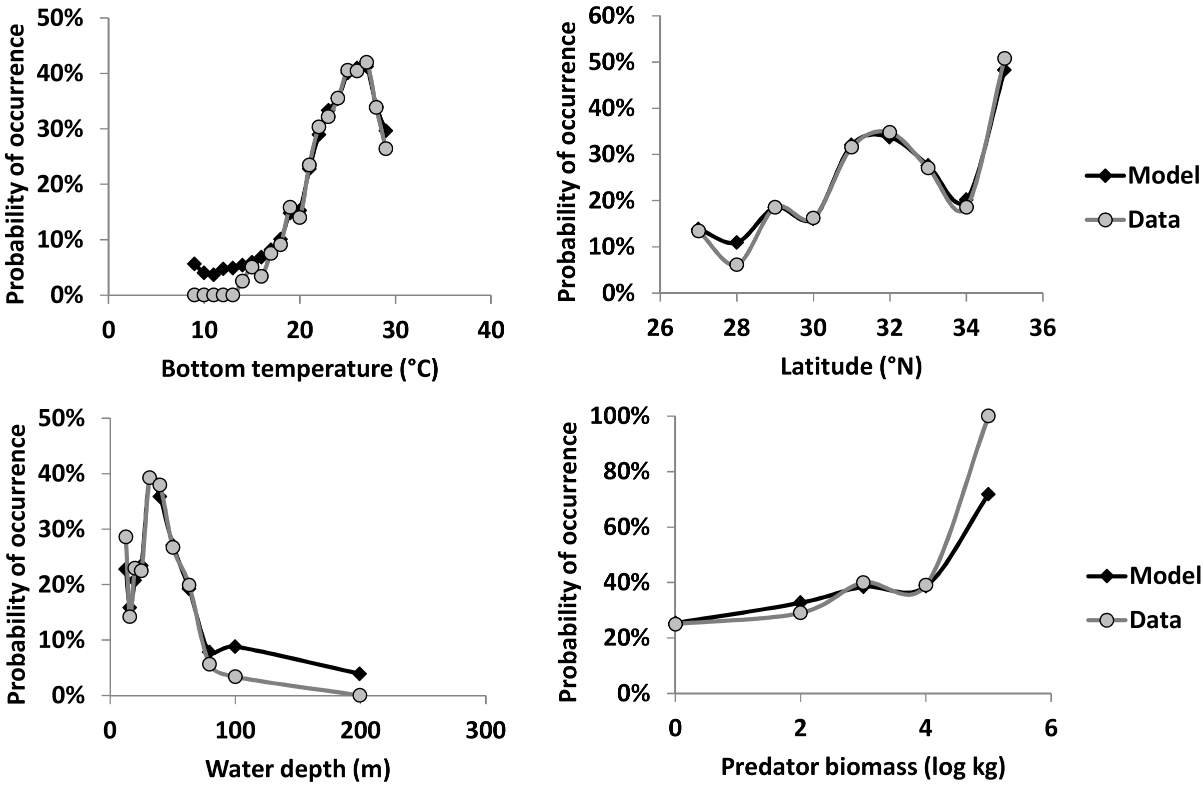

Figure 3: Mean abundance of Gray Triggerfish per trap at 0.1° latitude-longitude points for June – August 2002, and June – August 2003. Catches are overlaid on predicted probabilities of occurrence from the habitat model (top), and bottom temperatures from CTD (bottom).

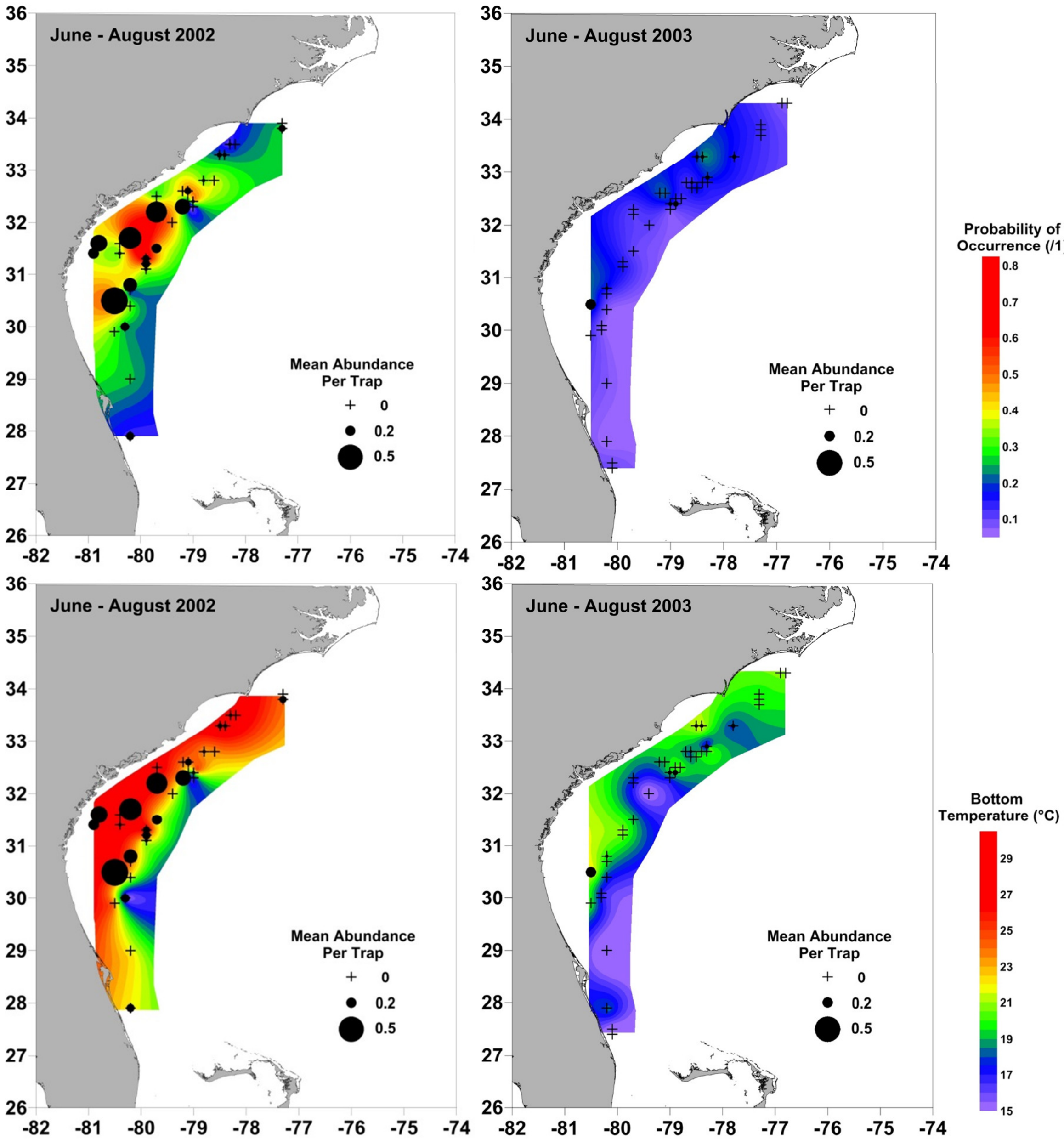

Figure 4: Scatterplots between surface and bottom temperatures at all sampled trap survey stations in spring (top) and summer (bottom), for locations with depths shallower than 30m (left), 30 – 50m (middle), and deeper than 50m (right). Upwelling anomalies are shown by points in the upper left of each plot (i.e. high surface temperatures, low bottom temperatures).

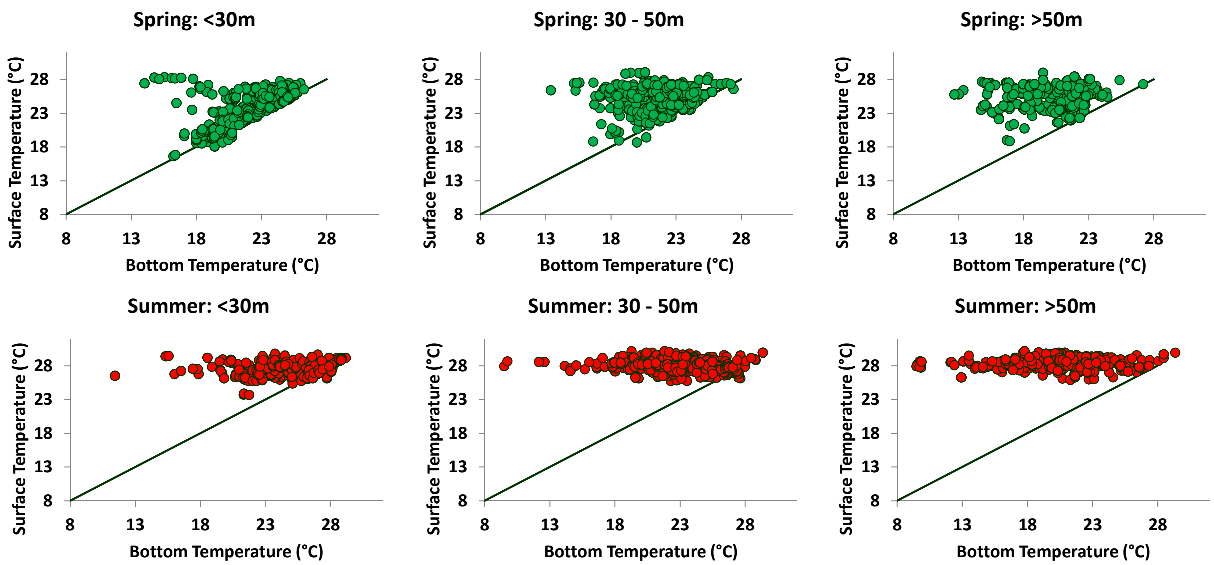

Figure 5: Time series of mean probabilities of Gray Triggerfish occurrence in traps from observations (red series), and the habitat model (black series). The habitat model series is bounded by upper and lower error bounds from the boosted classification tree model.

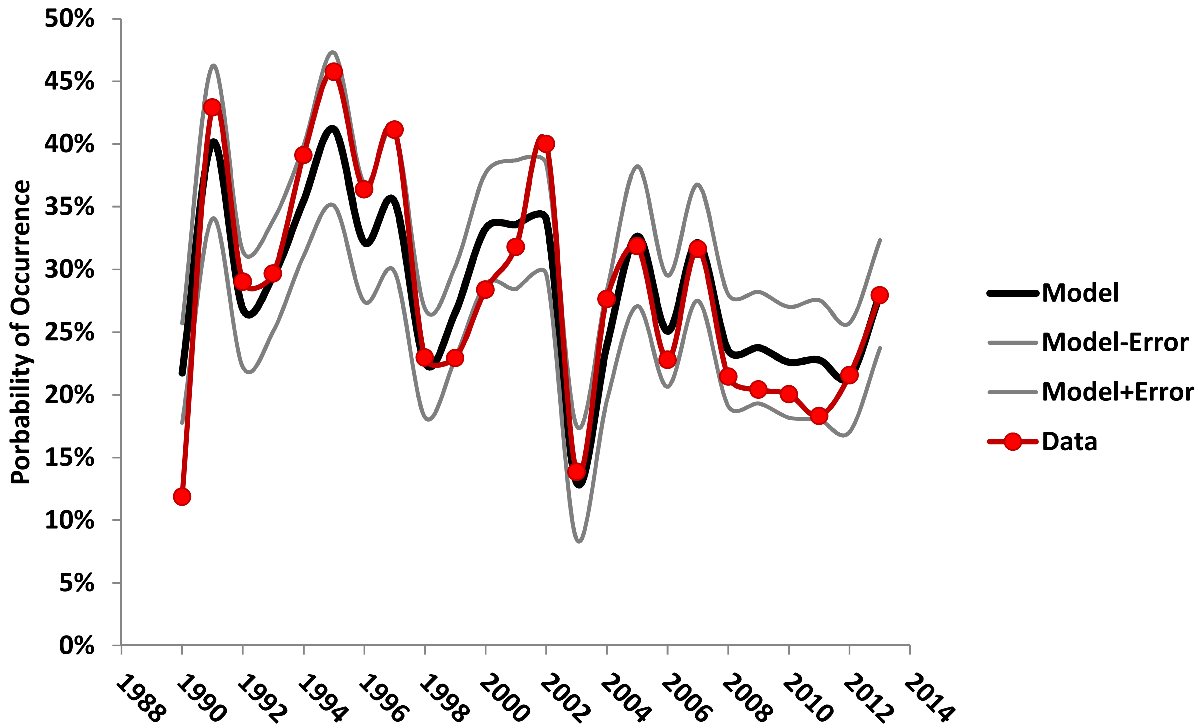

Figure 6: Predicted (line) and observed (dots) probabilities of occurrence of Gray Triggerfish in trap surveys, color-coded to show the three bins currently used for bottom temperature in the fishery-independent index standardization.

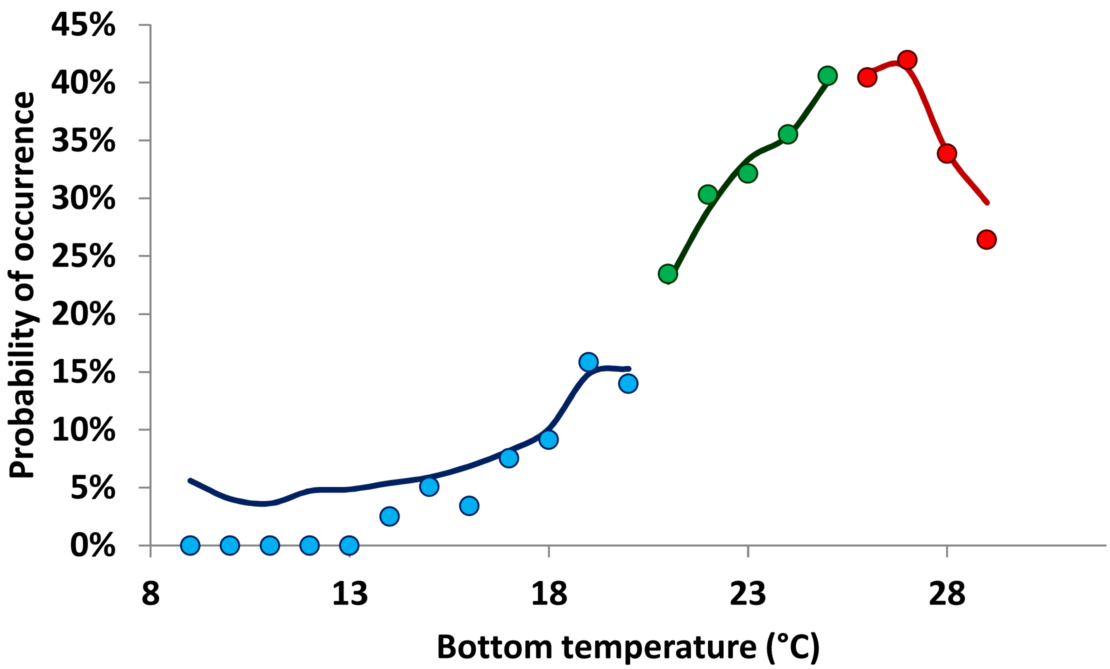

Figure 7: Observed bottom temperature values from CTD casts vs. predicted bottom temperatures from the global 1/12° HYCOM ocean model, September 2008 – October 2013. Each dot represents one cast location in time and space, with the corresponding HYCOM model value extracted.