

RESEARCH ARTICLE

Modeling the spatial and seasonal distribution of offshore recreational vessels in the southeast United States

Nancy Montes¹*, Robert Swett^{1,2}, Robert Ahrens²

1 Florida Sea Grant College Program, University of Florida, Gainesville, Gainesville, Florida, United States of America, **2** Fisheries and Aquatic Sciences Program, School of Forest Resources and Conservation, University of Florida, Gainesville, Florida, United States of America

* These authors contributed equally to this work.

* nancymontes@ufl.edu



OPEN ACCESS

Citation: Montes N, Swett R, Ahrens R (2018) Modeling the spatial and seasonal distribution of offshore recreational vessels in the southeast United States. PLoS ONE 13(11): e0208126. <https://doi.org/10.1371/journal.pone.0208126>

Editor: Athanassios C. Tsikliras, Aristotle University of Thessaloniki, GREECE

Received: June 4, 2018

Accepted: November 12, 2018

Published: November 28, 2018

Copyright: © 2018 Montes et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors thank the Southeast Regional Office NOAA Fisheries Services pursuant to National Oceanic and Atmospheric Administration (NAI00R4170079) for the financial support. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Understanding the distribution and intensity of recreational boating activities is key for managing safety as well as environmental and social impacts. Recreational boating is a very important component of the diverse maritime traffic in the southeastern United States. The seasonal distribution of offshore recreational vessels in waters off the coast of Northeast Florida and Southeast Georgia was modeled using several techniques (Poisson, negative binomial, hurdle and zero inflated modes, generalized additive models, and generalized mixed models) and by combining map-based information provided by recreational boaters with environmental and geographical variables to find the most parsimonious model. Based on model performance, the final model analysis was conducted using a GAM approach with a negative binomial distribution. The best seasonal models explained between 86.1%–88.6% of the total deviance. For most seasons, a model that included latitude, longitude, interaction between latitude and longitude, chlorophyll *a* concentration, and abundance of artificial reefs resulted in the best fit. The only exception was the model for the summer season, which did not include chlorophyll *a* concentration. Given the complexity of the study area, with a number of maritime activities and several marine species co-occurring, these models could provide information to analyze the distribution and overlap of recreational boating trips with other maritime activities (e.g., cargo ships, commercial vessels) and species (e.g., right whales, sea turtles, sharks). These analyses could be used to decrease harmful interactions among these groups and activities.

Introduction

The southeastern United States is a complex region where diverse maritime activities occur. In 2012 there were almost two million people inhabiting the seven coastal counties that comprise the study area in northeast Florida and southeast Georgia [1]. In addition, almost 87,500 recreational boats were registered in these seven counties [2,3]. This represents an average of one boat for every twenty-two residents, suggesting that recreational boating is an important and

popular activity in the study area. Montes and others [4] estimated that recreational vessels constitute 86% of vessel traffic at the main inlets of the study area. This includes boaters from nearby counties, and from several states of the USA and other countries (e.g., Canada). However, information about the spatial and temporal distribution of offshore recreational vessel traffic and activity in the study area is limited [4,5].

Understanding the distribution and intensity of recreational boating activities is key for managing safety, environmental, and social impacts [6]. Managers have made use of vessel traffic studies to address a number of issues, including human safety, waterway maintenance, channel marking, user conflicts, wildlife management, and allocation of infrastructure and boating facilities. For instance, Wu and others [7] examined boating safety issues and found that, in general, the relative incident rate (e.g., number of accidents) increased as weather conditions (e.g., wave height, sea surface temperature, air temperature, ice concentration, fog presence, and precipitation) deteriorated. Swett and others [8] designed a GIS-based decision support system that considers the physical conditions and boating characteristics of the Intra-coastal Waterway (ICW) when assessing the need for boating safety zones. Gorzelany [9] documented the effectiveness of regulatory speed zones in Miami-Dade and Palm Beach counties in Florida and found that compliance with speed zones is significantly related to vessel size and type. Bauduin and others [10] compared the spatial distribution and overlap of recreational vessels and manatees in Collier County, Florida. Their analyses provided a map showing areas with high/low probability of manatee/recreational vessel co-occurrence. Information from a regional waterway management system for Manatee, Sarasota, Charlotte, and Lee counties in Florida provides a tool and decision options for managing navigation channels in those counties [11].

Several approaches for analyzing spatial patterns of recreational boaters have been utilized. These include geostatistics and kernel density analysis [9,12,13,14,15,16], vessel traffic simulations using agent-based models [17,18,19], and distribution models based on external (e.g., behavioral, geographic, and/or networks) and situational (e.g., weather, water temperature, time, season) variables [20,21,22]. Nevertheless, the effect of external variables on the distribution of recreational boaters and the relationship among the different variables is not well understood (especially at a seasonal scale and within the marine environment). Furthermore, there is an array of statistical and modeling techniques for count data that, to our knowledge, has not been explored and documented to predict spatial distribution and abundance of offshore recreational vessels. Therefore, the objectives of this analysis were to explore the association between offshore recreational vessels and select external variables, as well as to explore different modeling technique to describe mail survey derived count data (i.e., Poisson, negative binomial, hurdle and zero inflated modes, generalized additive models, and generalized mixed models) to find parsimonious models that best describe the seasonal distribution of offshore recreational boating in the study area.

Methods

Study area

The study area comprised the offshore waters off the coast of Northeast Florida (Nassau, Duval, St Johns, Flagler, and Volusia counties) and Southeast Georgia (Camden and Glynn County) (Fig 1). The study area encompasses 28,523 km², extends 133 km into the Atlantic Ocean, and stretches 233 km along the coasts of Northeast Florida and Southeast Georgia. Several anthropogenic activities occur in the area, including military, commercial, and recreational activities.

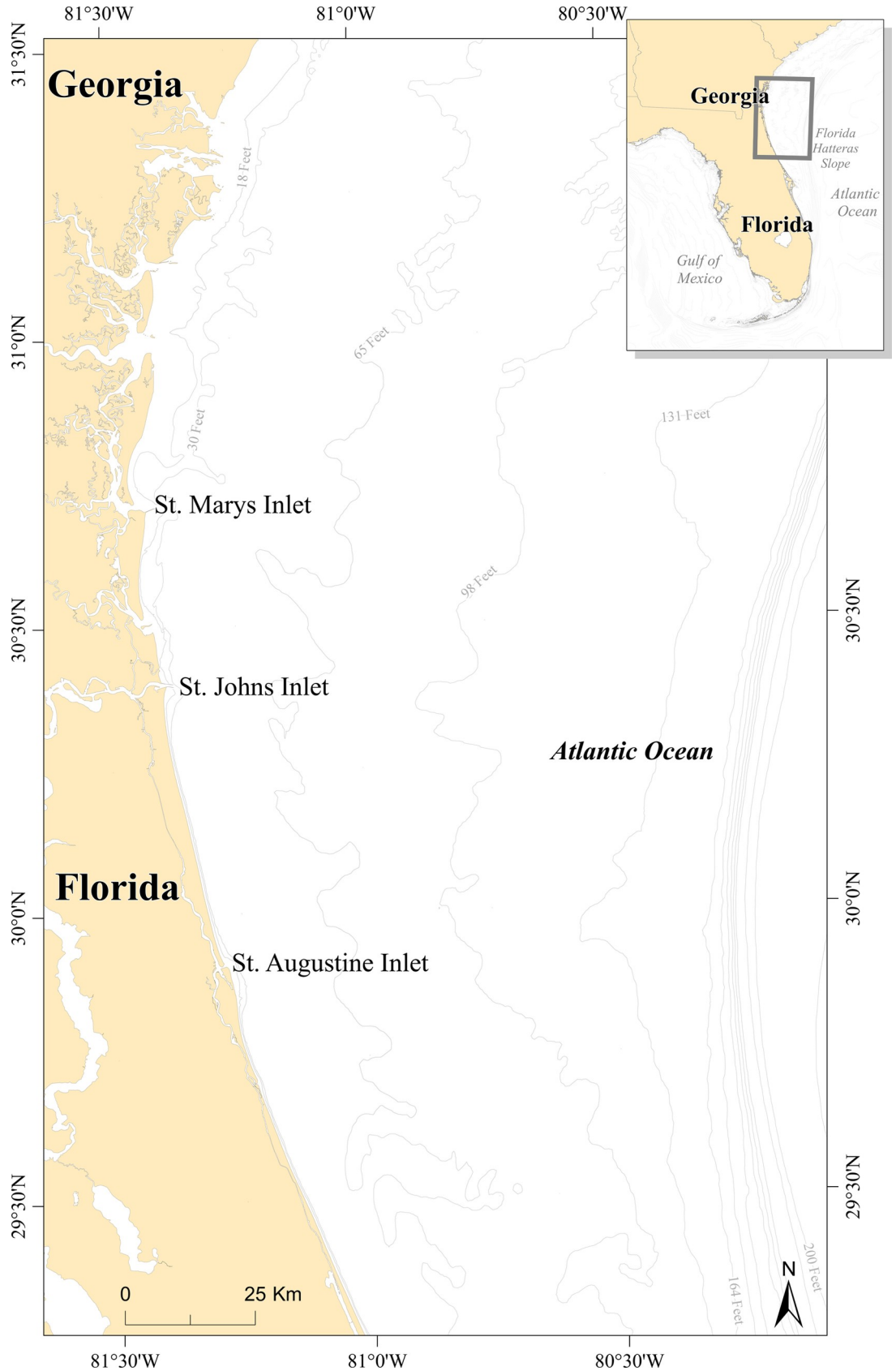


Fig 1. Study area. Northeast Florida (Nassau, Duval, St Johns, Flagler, and Volusia counties) and Southeast Georgia (Camden and Glynn County).

<https://doi.org/10.1371/journal.pone.0208126.g001>

Data collection

Recreational boating surveys. A sample of offshore recreational boating routes in the study area was obtained using a map-based questionnaire that was mailed to boat owners between September and October 2012. The mail survey design was based on previous vessel traffic studies [6,13,14,15,16]. The survey instrument and protocol were approved by the Institutional Review Board at the University of Florida. Survey participants received first an invitation/concent letter and survey questionnaire, followed by a reminder post-card (one-week later), and a replacement survey (two-weeks after). Boaters were asked to draw their entire on-the-water route for four boating trips that they had conducted: one each for winter, spring, summer, and fall. The sampling frame consisted of recreational boat owners whose vessels were recorded transiting the navigable inlets in the study area (St Marys, St Johns, and St Augustine) over a period of 17 months (2011–2012). Vessel owner names and addresses were obtained by linking vessel characteristics recorded at the inlets (e.g., registration number or vessel name and hailing port) with information contained in Florida's and Georgia's vessel registration systems, as well as the U.S. Coast Guard National Vessel Documentation Center database [23]. Surveys were distributed to all identified recreational boaters (N = 5,034). Completed surveys were digitized into ArcGIS 10.3 [24]. Non-response bias was evaluated by comparing the self-reported age of early (proxy for respondents) and late (proxy for non-respondents) respondents using a *t*-test [25].

Boating routes were digitized as polylines using an equidistant Universal Transverse Mercator (UTM) projection. A sampling grid cell (5.56 x 5.56 km cells) was used to account for disparities in our study area. The dependent variable (number of routes that intersected each grid cell) was calculated using a spatial join of the boating routes and the sampling grid layer. Additionally, we re-coded the dependent variable into a presence/absence matrix by routes and respondents to accommodate for the mixed model analysis of the winter routes.

Habitat variables. Several habitat variables were examined to model the relative probability that a recreational vessel will occupy any given grid cell in the study area. Habitat covariates included sea surface temperature (SST), chlorophyll *a* concentrations (Chl), water depth, abundance of artificial reefs, distance from each grid centroid to the nearest inlet (km), and latitude and longitude of each grid centroid.

We created seasonal composite images for chlorophyll *a* concentrations and SST using monthly satellite-derived information from the NASA Ocean Color Webpage [26] obtained by Terra and Aqua MODIS. Only pixels with high quality (quality = 0) were used for SST and chlorophyll *a* concentrations. Water depth data (90 m horizontal resolution) was obtained from the National Centers for Environmental Information [27]. The distance to the nearest inlet was calculated in ArcGIS 10.3 using the centroid of each grid cell and the mouth of each of the inlets in the study area. Information related with artificial reefs was downloaded from the Florida Geographic Data Library [28]. This layer was spatially joined with the sampling grid cells to obtain the number of artificial reefs and/or fish havens in each grid cell.

Data analysis

Habitat model associations. Partial Mantel tests using *Ecodist* [29] in the R software package version 3.0.2 [30] were used to evaluate the association between each of the habitat variables and the abundance of seasonal boating routes. This technique, a non-parametric test

based on dissimilarity matrices, considers autocorrelation while assessing the degree of correlation between variables [31,32]. The structure of the dissimilarity matrix is not independent; therefore, we performed significant testing through permutation procedures (1,000 permutations).

Density model. The number of recreational boating trips per grid cell is a discrete non-negative integer or count. There are several modeling techniques used for count data. Regression models for counts are nonlinear with many properties and special features intimately connected to discreteness and nonlinearity [33]. To explore which of those techniques better accommodates our dataset, we tested several modeling techniques: generalized linear models (GLM) with a Poisson distribution, a negative binomial distribution, generalized additive models with negative binomial link (GAM), hurdle negative binomial models, zero-inflated negative binomial models, and a generalized linear mixed effects modeling approach.

The GLM and the GAM consist of three components. The first component corresponds to the distribution of the response variable, the second is the specification of the systematic component in terms of explanatory variables, and the third is the link between the mean and the response variable and the systematic part [34,35]. The Poisson GLM assumes that the mean is equal to the variance. Poisson models consider that count data are always non-negative and tend to be heterogeneous. The negative binomial GLM is an alternative to the Poisson model when there is over dispersion (the variability in the data is greater than the proposed model) relative to a Poisson model [35]. The generalized additive model (GAM) with a negative binomial link was also implemented. GAM is a generalized linear model with a linear predictor involving a sum of smooth functions of covariates, which allows for non-linear relationships between the dependent and the independent variables [34]. To account for the possibility of an excessive number of zeros in the seasonal datasets, the boating trip abundance was also modeled using two-part models. Zero inflation suggest that zeros are derived in two ways. The first are zeros as a result of absence (i.e., the object of interest does not exist in the spatial location) and the second are zeros as a result of non-detection (i.e., the object of interest does exist in the location, but it was not observed). This process results in far more zeros than would be expected for a specific distribution (in our case the negative binomial distribution). It is important to account for zero inflation because it could bias the estimated parameters and lead to over dispersion in the model [35]. In a two-part model, a binomial model is applied where data are considered as zeros versus non-zeros. Then a Poisson or negative binomial model is used to model the non-zero counts. The hurdle negative binomial model is a two-part model that does not discriminate between the origins of the zeros. It assumes that a hurdle should be crossed before values become non-zeros. In this case the distribution of the count part of the model is zero truncated [35]. The zero-inflated negative binomial model (ZINB) is another popular two-part model. It differs from the hurdle model in the sense that it models the zeros as if they are coming from different processes (binomial and count processes). The distribution for the count model in the ZINB is not zero-truncated (the count model is allowed to generate zero values).

Model analyses were conducted in R software package version 3.0.2 [30] using distinct packages. The *MASS* [36] and *car* [37] packages were used for the Poisson and Negative Binomial regression. Hurdle and zero inflated negative binomial models were obtained using the *p scl* package [38]. Generalized additive models were run using *mgcv* [39] and *nlme* [40]. To account for collinearity, a stepwise procedure was implemented and habitat variables with variance inflation factors (VIF) higher than 3 were eliminated from further models [34]. Model diagnostics and autocorrelation estimates for the residuals using Moran's Index were obtained and model performance was compared using the Akaike Information Criterion (AIC), delta AIC, and AIC weights (using *MuMin*) [41]. Predictions of each model were mapped and

compared using regression techniques (major axis regression) on the observed vs. predicted values for each season using *lmodel2* [42,43].

GAM model. Based on model performance, the final model analysis was conducted using a GAM approach with a negative binomial distribution and logit link. To minimize overfitting, a penalty term was added to the regression to control for the smoothness of the fitted curve by setting the gamma = 1.4 in the 'gam' function to force the effective degrees of freedom to count as 1.4 times the degrees of freedom in the generalized cross-validation score [34]. The smoothness selection was fit using spline-based penalized likelihood estimation. Theta parameters and weighted penalties were determined by AIC, delta, AIC weights, residuals sum of square, and root mean squared deviation estimates. Thin plate regression splines over location were used to account for autocorrelation [44]. We followed a stepwise procedure for model selection. The total deviance explained, the AIC estimates, and the significance of each added variable ($p > .05$) were used to evaluate the role of the independent variables in explaining the observed distribution of recreational vessels. Predictions of each working model were compared with the observed dataset using regression techniques to validate the best model.

Mixed effect logistic model (presence/absence). The generalized mixed model (GLMM) is an extension of the GLM that allows for correlation between the observations through the introduction of random effects [35,45]. Models that include the random effect consider any correlation between survey participants. Mixed effects models feature both fixed and random effects. Fixed effects are unknown or constant variables (observations are independent). If we assume that there is some type of relationship between observations, then those variables are considered random effects [46]. In some cases, survey respondents provided more than one recreational boating trip per season, therefore respondents' ID was used as a random effect. A model for the winter season was performed in R software version 3.0.2 using the *lme4* package [47]. Pearson correlations, boxplots, and VIF were used to select the habitat variables that were included in the final model. The final model was selected using AIC estimates.

Results

We had a 19.03% return rate ($n = 958$ returned surveys). However, we were not able to use all of the returned surveys. Forty-eight percent of those who returned a survey were not our target population (e.g., inshore boaters, boating activities occurred out of our study area) or returned a blank survey. Data collected from 507 surveys were digitized, which yielded a sample of 2,522 boating routes (all four seasons). Difference in the self-reported age for respondents (early) ($M = 56.9$, $SD = 10.8$) and non-respondents (late) ($M = 55.9$; $SD = 11.3$) was not statistically significant ($t(340) = 0.86$, $p = 0.39$).

The mean count of routes intersecting any grid cell was 9.2 ($SD = 20.5$). Throughout all seasons, 21% of grid cells showed no counts of recreational vessels. However, this number fluctuated between 13% of grid cells with zero counts in the spring season up to 27% in the winter season.

Habitat model associations

Seasonal mean SST varied from 19.8°C ($SD = 2.2$) in winter to 27.3°C ($SD = 0.48$) in summer. Grid cells that were intercepted by at least one vessel route have a mean SST of 24.02°C ($SD = 3.06$). Seasonal chlorophyll *a* concentrations ranged from 0.9 mg/m³ ($SD = 1.07$) in spring to 1.29 mg/m³ ($SD = 1.3$) in winter. Vessel routes were recorded at chlorophyll *a* mean concentrations of 1.16 mg/m³ ($SD = 1.2$). Based on *t*-test results, there was a statistically significant difference with respect to the chlorophyll *a* concentration (Table 1). In general, grid cells that were intersected by at least one recreational trip had an average water depth of 38.1 m

Table 1. Descriptive statistics of grid cells associated with the presence and absence of recreational boating routes. Recreational boating routes (Routes), sea surface temperature (SST), chlorophyll *a* concentrations, water depth, distance of the centroid of each grid cell to the nearest inlet (Inlet), and abundance of artificial reefs (Reef).

Variables	Mean (SD)		t-test	p-value
	Routes = 0	Routes ≥ 1		
Routes	777	2919	na	na
SST (°C)	23.8 (3.6)	24.0 (3.1)	-1.6	.11
Chlorophyll <i>a</i>	0.6 (0.8)	1.2 (1.2)	-14.4	< .001
Water depth (m)	-60.6 (91.6)	-38.1 (53.9)	-6.6	< .001
Inlet (km)	87.6 (27.9)	57.0 (32.6)	26.2	< .001
Reef	2.3%	12.4%	-12.2	< .001

<https://doi.org/10.1371/journal.pone.0208126.t001>

(SD = 53.9), were closer to inlets (mean distance = 57.0 km, SD = 32.6), and had a higher percentage of artificial reefs (12.4%). When comparing grid cells with recreational trip counts to those without, there is a statistical difference for all environmental variables except for SST (Table 1).

Recreational boating trip abundance was significantly correlated year-round with chlorophyll *a* concentration, distance from grid centroids to the nearest inlet, and abundance of artificial reefs. Chlorophyll *a* concentration had the strongest relationship with boating trip abundance (Mantel $r = 0.21-0.27$), with the strongest correlation observed during the winter season. Pearson’s correlation analysis led to similar conclusions with respect to the relationship between the dependent and independent variables [48].

Simple mantel tests and Pearson’s correlation analysis showed significant associations between SST, chlorophyll *a* concentration, distance to the nearest inlet, and latitude for all seasons. Water depth, SST, and distance to the nearest inlet variables were excluded from further analyses because they showed the highest VIF. Further analyses were conducted using only chlorophyll *a* concentration, latitude and longitude, and abundance of artificial reefs as explanatory variables.

The presence/absence model for the winter season allowed for the incorporation of other independent variables, such as boaters’ age and boating experience, total route length, inlet where the route originated, and vessel length. Distance to the nearest inlet ($r = -0.18$), total route length ($r = 0.1$), chlorophyll *a* concentration ($r = 0.1$), and longitude ($r = -0.1$) showed some relationship with the presence/absence of recreational boating routes.

Density model–comparison of count data models

Based on AIC estimates, the generalized additive model (GAM) fit the data better than the other evaluated models (Table 2). Analysis of the residuals also corroborated this. When the

Table 2. Model comparison for the count data for different seasons. k = number of model parameters, RSS = residual sum of squares, weights = AIC weights, RMSD = root mean square deviation.

Model	Winter					Spring					Summer					Fall				
	k	RSS	AIC	Delta	Weight	k	RSS	AIC	Delta	Weight	k	RSS	AIC	Delta	Weight	k	RSS	AIC	Delta	Weight
Poisson (PO)	5	7289	9626	5387	<0.001	5	9011	12035	6864	<0.001	5	10855	13352	8942	<0.001	5	7319	9878	5441	<0.001
Negative binomial (NB)	6	981	4714	475	<0.001	6	1000	5573	401	<0.001	6	966	5036	627	<0.001	6	974	4816	380	<0.001
Hurdle (HUNB)	11	960	4639	401	<0.001	9	997	5522	351	<0.001	9	929	5030	620	<0.001	10	1139	4779	343	<0.001
Zero-inflated (ZINB)	11	1024	4554	315	<0.001	9	948	5557	386	<0.001	9	987	4988	579	<0.001	10	1244	4810	373	<0.001
Generalized additive (GAM)	19	477	4239	0	0.99	25	421	5171	0	0.99	19	508	4410	0	0.99	18	569	4437	0	0.98

<https://doi.org/10.1371/journal.pone.0208126.t002>

normal probability plots of deviance residuals for the different models were compared, only the residuals from the negative binomial and the GAM models did not show any unusual pattern. Furthermore, when the predicted values generated by each of the models for the different seasons were compared with the observed values of the dependent variable (boating trip abundance), it was evident that the model with the best fit was the GAM (Fig 2).

GAM model

Model selection (stepwise) was performed for each season. For most seasons, a model that included latitude, longitude, interaction between latitude and longitude, chlorophyll *a* concentration, and abundance of artificial reefs resulted in the best fit (higher explained deviance and lower AIC). The only exception was the model for the summer season, which did not include chlorophyll *a* concentration. The best model explained between 86.1% of the total deviance in the fall and up to 88.6% in summer season (Table 3).

Offshore recreational vessel abundance showed a nonlinear relationship with location (longitude and latitude). The distribution of recreational vessels over the longitude (west-east) axis (x) showed an increasing pattern with peaks around the 81°0'W and 80°48'W longitudes (Fig 3). However, this distribution was wider during the summer season followed by the winter season, with vessels travelling further east. Over the summer and fall, the smoothing function showed a narrowed size distribution on the longitude axis (Fig 3). The smoothing functions for the distribution of recreational boating trips for the latitude axis (y) showed an increasing pattern from south to north, up to a latitude of 30°45'N (north of St. Marys inlet) (Fig 3).

The smoothing function for chlorophyll *a* concentration in the model varied seasonally. For the winter season, as the concentration of chlorophyll *a* increased so did the abundance of recreational vessels, with a peak between 5–7 mg/m³ (Fig 3). During the spring season a higher abundance of vessels occurred between 4.5–5.5 mg/m³, followed by another peak between 2.5–3.5 mg/m³ (Fig 3). For the summer season, chlorophyll *a* concentration showed a linear relationship with respect to vessel abundance (Fig 3); however, this relationship was not significant, and this variable was dropped from the final summer season model. Over the fall season vessel abundance increased with increasing chlorophyll *a* concentration (Fig 3).

The relationship between the distribution and abundance of offshore recreational vessels and the abundance of artificial reefs was similar for all four seasons. In general, the abundance of vessels increased with increasing numbers of artificial reefs, up to 2–4 reefs per grid cell. However, as the number of artificial reefs increased so did the uncertainty (wider confidence intervals) (Fig 3).

Model predictions were tested by comparing observed and predicted values. Selected models for each season showed R-square values between 0.64 and 0.79 (Table 3). All models underestimate boating trip abundance at high values, especially for grid cells near inlets. Fig 4 shows the observed and predicted distribution and abundance of recreational vessels. Even though the models show similar means (e.g., observed mean for winter was 6.86 and the mean of the predicted values was 6.62), model predictions did not reach maximum values as extreme as those observed in the training data set. Moran's I estimates on the residuals did not show any significant sign of autocorrelation (Moran's I ranged from 0.34, *p*-value >.05 for summer up to 0.37, *p*-value >.05 for the spring season).

Mixed effect logistic model (GLMM)

Model selection (stepwise procedure) identified latitude, longitude, and vessel size (length) as the only significant fixed variables while controlling for correlation between the information provided by boaters (boater's ID). Predictions from the best model for the GLMM are shown

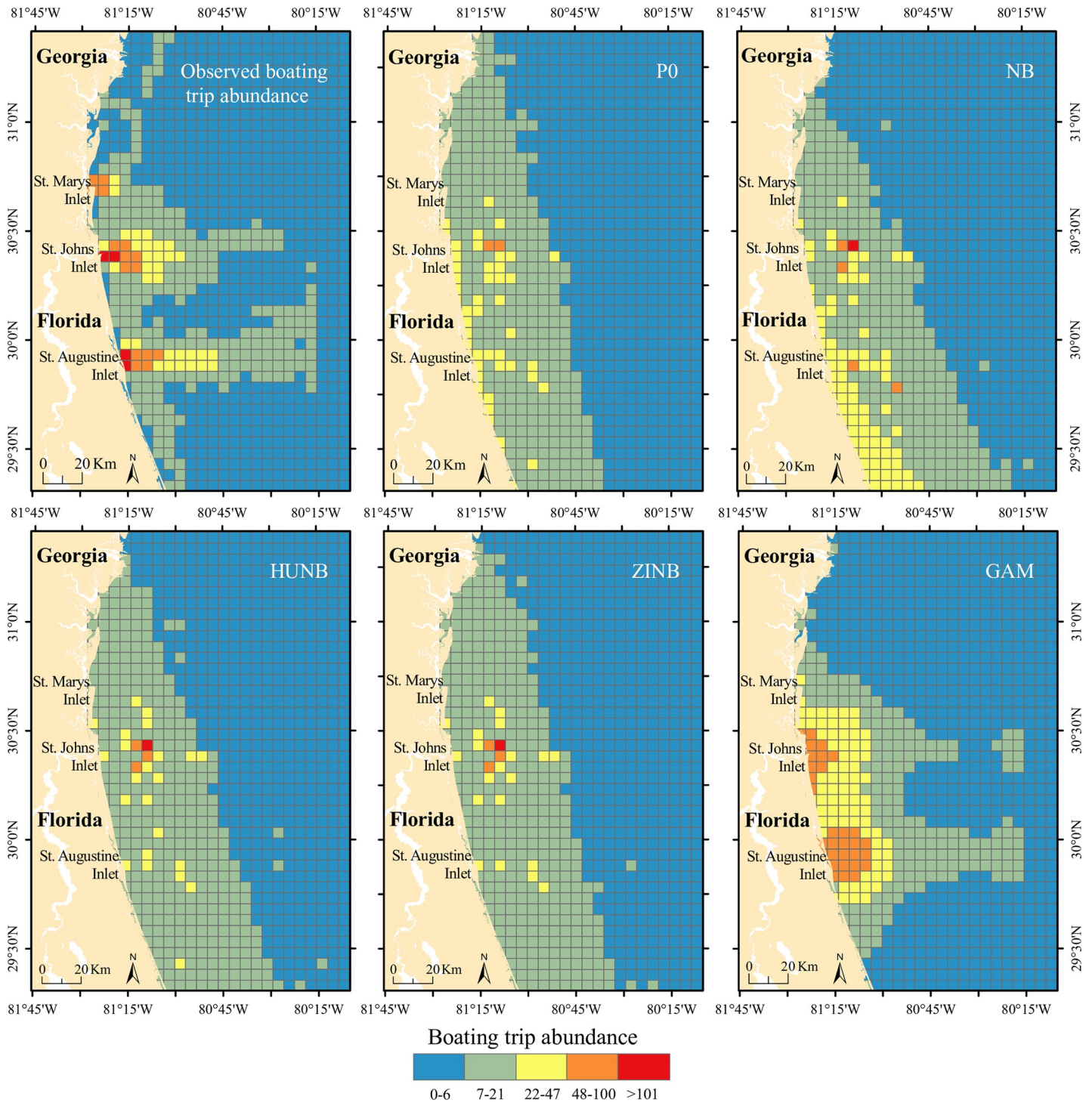


Fig 2. Observed and predicted boating trip abundance generated by different models for the winter season: Poisson model (P0), negative binomial model (NB), hurdle negative binomial model (HUNB), zero inflated negative binomial model (ZINB), and generalized additive model (GAM).

<https://doi.org/10.1371/journal.pone.0208126.g002>

in Fig 5. In this case, the cumulative predicted values reached maximum values closer to the observed (cumulative value predicted by GLMM = 154; Max value observed = 174). Areas near the St. Augustine, St. Johns, and St. Marys inlets show the highest predicted values.

Table 3. Model selection to test for the spatial distribution of offshore recreational boating trips for each season. x = longitude, y = latitude, x:y = interaction between longitude and latitude, chl = chlorophyll a concentrations, ref = reef abundance per sampling unit, and s = thin plate regression spline fit, k = number of model parameters, RSS = residual sum of squares, weights = AIC weights, RMSD = root mean square deviation.

Model	Explanatory variables					Model parameters							Observed vs Predicted	
	s(x:y)	s(x)	s(y)	s(chl)	s(ref)	k	RSS	AICc	Delta	Weight	% deviance explained	R-Sq (adj)	R-sq	RMSD
WINTER														
5	+	+	+	+	+	51	683	3644	0	0.99	87.1	0.61	0.64	9.0
4	+	+	+	+		53	696	3661	17	0	86.9	0.61	0.65	9.0
3	+	+	+			47	724	3675	32	0	86.3	0.57	0.61	9.0
2	+	+				36	957	3884	240	0	81.9	0.39	0.41	11.3
1	+					24	1165	4066	422	0	78.0	0.38	0.39	11.5
SPRING														
5	+	+	+	+	+	51	705	4581	0	0.99	86.2	0.78	0.79	7.7
4	+	+	+	+		50	721	4593	13	0.00	85.8	0.76	0.78	8.0
3	+	+	+			44	750	4609	28	0	85.2	0.75	0.76	8.2
2	+	+				35	869	4709	129	0	82.9	0.62	0.63	10.2
1	+					24	1044	4860	279	0	79.4	0.56	0.57	11.0
SUMMER														
6	+	+	+		+	48	843	4057	0	0.91	88.6	0.62	0.64	15.1
5	+	+	+	+	+	45	855	4061	5	0.09	88.4	0.63	0.65	14.9
3	+	+	+			46	883	4090	34	0	88.0	0.61	0.63	15.3
4	+	+	+	+		47	883	4092	66	0	88.0	0.61	0.63	15.4
2	+	+				34	1150	4331	275	0	84.4	0.47	0.49	18.0
1	+					24	1268	4428	371	0	82.8	0.43	0.45	18.7
FALL														
5	+	+	+	+	+	52	753	3978	0	0.99	86.1	0.67	0.71	9.9
4	+	+	+	+		49	772	3990	12	0.00	85.8	0.67	0.71	9.9
3	+	+	+			45	801	4011	33	0	85.2	0.61	0.63	10.8
2	+	+				36	1081	4270	292	0	80.1	0.45	0.47	12.9
1	+					24	1240	4403	425	0	77.2	0.41	0.43	13.4

<https://doi.org/10.1371/journal.pone.0208126.t003>

Discussion

Our results show that the seasonal distribution of offshore recreational vessels can be explained mainly by geographical variables such as latitude and longitude, abundance of artificial reefs, and chlorophyll *a* concentrations. Neither SST nor water depth showed any effect on the seasonal distribution of recreational vessels. Other studies have used water depth [18] and SST [7] to describe boat traffic, but the temporal scale in these cases was shorter than our study (e.g., Wu et al., 2009 used hourly, weekly, and monthly estimates), which could explain the differences. Furthermore, SST and water depth were highly correlated with other variables and to avoid problems with multicollinearity they were not included in the final seasonal models. By analyzing the dataset using mixed model methods, we were able to measure the effects of variables related to boaters, their boats, and their self-mapped boating trips. In general, when used to describe presence/absence patterns of boating routes in the sampling grid cells, all the relationships between these variables were weak, ranging from -0.18 to 0.1 (with 1 being highly correlated). Nevertheless, the analyses of the dataset using this methodology allowed us to corroborate the results obtained by the GAM analyses.

The frequency of offshore recreational boating trips per spatial cell follows a negative binomial distribution. Although there are a great number of modeling approaches that can be used

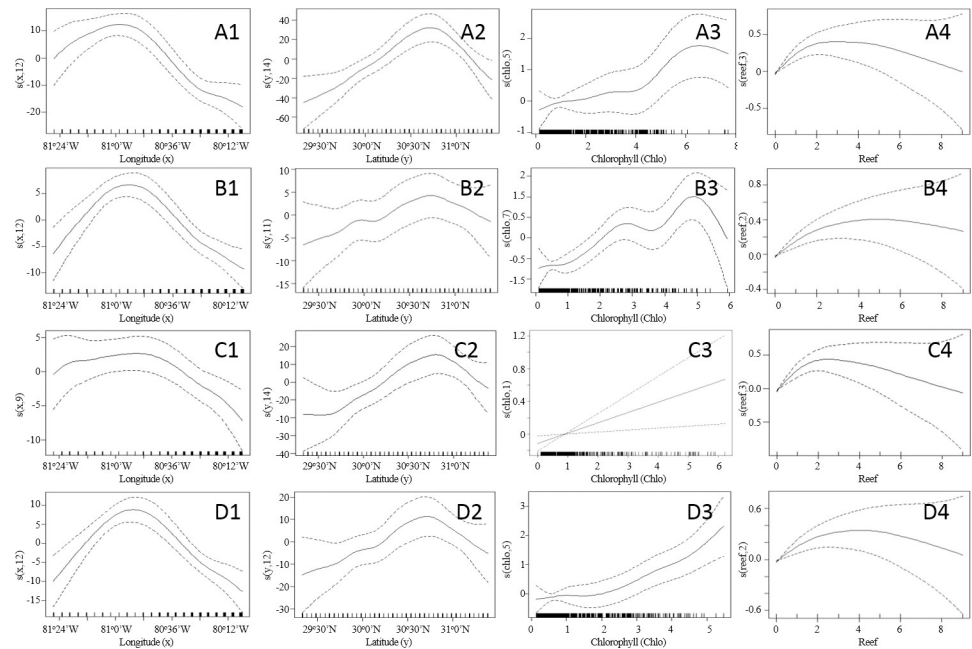


Fig 3. Smoothed curve of the additive effect of covariates modeling the abundance of offshore recreational boating trips in the southeastern United States. Winter (A), Spring (B), Summer (C), and Fall (D). Predictor variables include longitude (x), latitude (y), chlorophyll *a* concentrations (Chlo), and abundance of artificial reefs (Reef). Solid line depicts the estimate of the smooth function and dashed lines indicate 95% confidence bands. Estimated degrees of freedom for each smoothed variable are shown on the y-axis.

<https://doi.org/10.1371/journal.pone.0208126.g003>

to describe and predict offshore recreational boating trips, our results show that the GAM produced a superior model with better fit than did the other studied models. The use of spline functions to describe the relationship between the dependent and independent variables proved to be more effective.

The most parsimonious model for almost all seasons includes latitude and longitude, chlorophyll *a* concentrations, and abundance of artificial reefs. Because there are only a limited number of access points to offshore waters, we expected location (in the form of latitude and longitude) to have an important role in describing the distribution of offshore recreational vessels. Sidman and Fik [20] also highlighted the importance of geographic and network variables in modeling the distribution of recreational vessels and destination choices. Recreational vessel distribution has also been linked with the distribution of target fish species. Fishing is the leading activity of recreational vessels in Florida. Sidman and others [16] found that 73% of recreational boaters in Brevard County (adjacent to our study area) engaged in fishing activities. Several studies found a link between concentrations of chlorophyll *a* and sea surface temperature and fishery yields [49,50]. In the case of recreational fisheries, Wall and others [51] found that positive catch rates of king mackerel (*Scomberomorus cavalla*) were significantly influenced by chlorophyll *a* values and temporal variables (season and year). As a result, we included chlorophyll *a* concentration as a proxy variable for fish distribution (which is the main target of anglers). Overall, we found a positive effect of chlorophyll *a* concentration on the distribution of recreational boating trips. The only exception for this finding was the boating routes described over the summer months. Consequently, we included chlorophyll *a* concentration as a proxy variable for fish distribution (which are the main target of anglers). In general, we found a positive effect of chlorophyll *a* concentration on the distribution of recreational boating trips. The only exception for this finding was the boating routes described over

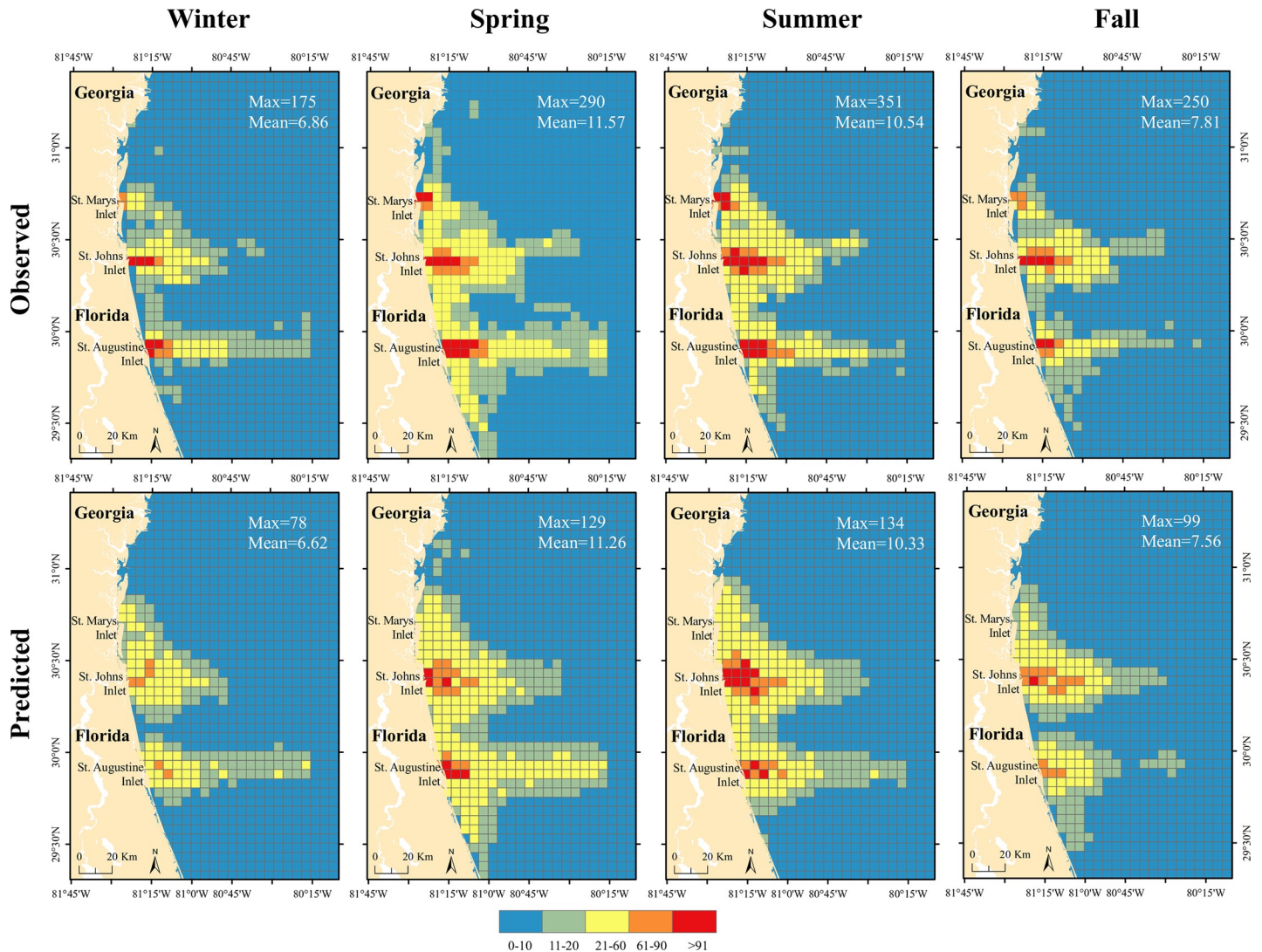


Fig 4. Observed and predicted offshore recreational boating trip abundance per grid cell for different seasons using GAMs.

<https://doi.org/10.1371/journal.pone.0208126.g004>

the summer months. During the summer, chlorophyll *a* concentrations were more homogeneously distributed throughout the study area, resulting in non-significant differences between grid cells with and without boating routes and these variables.

We also included abundance of artificial reefs in our analysis due to their relationship with fish abundance [52,53,54]. Based on our results, there seems to be a positive relationship with abundance of artificial reefs and offshore recreational boating routes for all seasons. Nevertheless, we need to be cautious when interpreting this observed association given that the placement of artificial reefs is likely predicated on several factors (e.g., convenience, near inlets, or close to specific habitats). More research is needed to clarify whether recreational boaters are really targeting areas near artificial reefs or if the observed relationship is an effect of the placement of these artificial structures at vessel transiting areas.

As expected, there were slight differences in the seasonal distribution of recreational vessels, with the spring season showing a more dispersed distribution when compared with other seasons. Adverse climatic conditions (e.g., decreased temperatures, rough sea states) over the

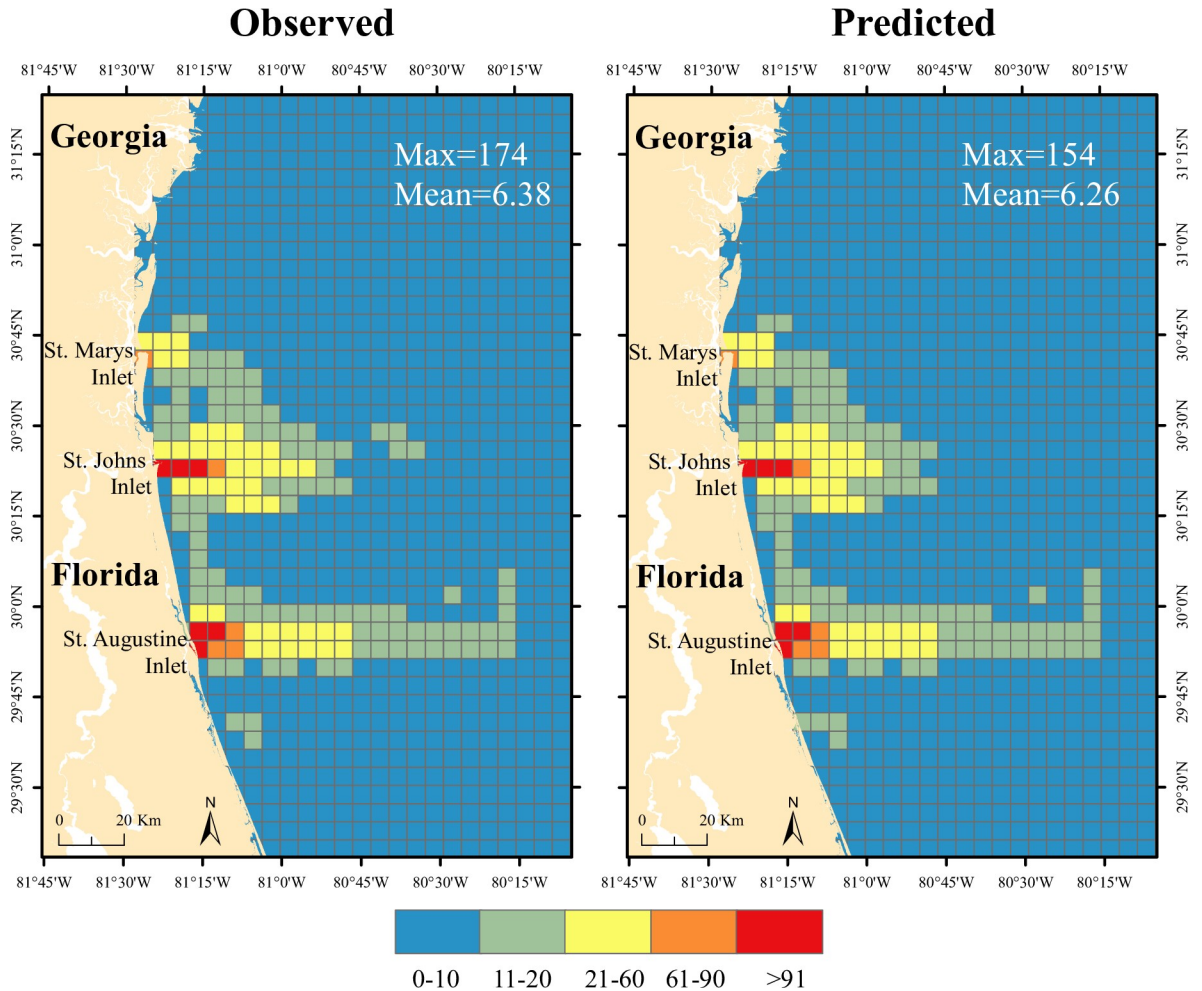


Fig 5. Presence/absence and predicted offshore recreational boating trip abundance for the winter season from the GLMM.

<https://doi.org/10.1371/journal.pone.0208126.g005>

winter and fall seasons, and nearshore locations of target fish species may influence how far east (longitudinal movements) recreational boating trips reach; which would affect not only the number of boating trips but also their distribution. With respect to the north-south movements (latitudinal), two areas were shown to be of great importance: St. Johns inlet and St. Augustine inlet. Year-round, most of the boating trips departed from latitudes that correspond with these two inlets. However, we did observe seasonal changes with respect to the abundance of boating trips originating from these inlets. For the winter and spring seasons we observed a greater number and a wider distribution of boating trips at St. Augustine inlet. This inlet is closest (shortest distance) to the Florida Hatteras slope and to some important deep-water fishing areas, such as “The Ledge” (Waterproof Chart #125F). Over the summer and fall seasons, even though the area experiences an increase in temperature and calmer waters, it also experiences an increase in weather events such as hurricanes and tropical storms. During the study period (2011–2012) there were 39 tropical storms and 17 hurricanes [55,56], some of which could have affected the observed distribution of boating routes and ultimately affected predicted values generated from those observations. Our dataset has limitations due to the data collection timeframe and the temporal scale of our study. More on-the-ground data is

needed at different temporal and spatial scales to clarify the effect of major weather events on the distribution of offshore recreational boating trips in the study area.

Although our dataset conforms to a negative binomial distribution, there are some high values located near the inlets that affected the predicted values generated for each of the best seasonal models. In most of the instances, those high values could be considered outliers. However, because it was necessary to consider them due to the nature of our study, they were not discarded from our analyses. Nonetheless, when we compared the predictions generated from the GAM models with the presence/absence approach (GLMM), there is a good level of agreement with respect to the predicted values. Therefore, it is important to point out that although our models showed a moderate fit, they present limitations that should be taken into account when using this information for spatial planning and management purposes.

It is important to acknowledge that the findings of our study may be limited to the surveyed sample because the return rate of our mail survey (19%) might be considered low. The main issue with low response rates is that non-respondents may be underrepresented in this sample, limiting the generalizability of the results to the population of interest [57]. However, this is typical of recreational boater surveys, with similar studies reporting return rates between 19–23% [13,14,15,16].

Management implications

The information provided by this study can be used by managers and law enforcement to determine areas with high/low probability of recreational vessel occurrence and abundance. The study area is large, and this information could help prioritize areas for spatial planning and management purposes. The study area is also very complex, with a number of maritime activities and several marine species co-occurring. Therefore, these models could provide information to analyze the distribution and overlap of recreational boating trips with other maritime activities (e.g., cargo ships, commercial vessels) and species (e.g., right whales, sea turtles, sharks) that utilize this area. These analyses could be used to decrease harmful interactions among these groups and activities.

Supporting information

S1 File. Map-Based_Survey. Questionnaire used to survey offshore recreational boaters in the southeastern United States.
(PDF)

S2 File. Offshore recreational vessels. Abundance of offshore recreational vessels (vessels) and co-variables used to model offshore recreational trips in the southeastern United States. Co-variables are: longitude (x), latitude (y), sea surface temperature (sst), chlorophyll a concentration (chlo), water depth (bathy), abundance of artificial reefs (reef), distance to the nearest inlet (inlet), and season.
(CSV)

Author Contributions

Conceptualization: Nancy Montes, Robert Swett.

Formal analysis: Nancy Montes.

Funding acquisition: Robert Swett.

Investigation: Nancy Montes, Robert Swett.

Methodology: Nancy Montes, Robert Ahrens.

Project administration: Nancy Montes, Robert Swett.

Resources: Nancy Montes, Robert Swett.

Validation: Nancy Montes.

Visualization: Nancy Montes.

Writing – original draft: Nancy Montes.

Writing – review & editing: Robert Swett, Robert Ahrens.

References

1. US Census Bureau. Database; 2013 [cited 2013 Feb 1]. Available from: <http://www.census.gov/2010census/data>.
2. FLHSMV. Florida vessel owners: Statistics. Florida Department of Highway Safety and Motor Vehicles. Database; 2013 [cited 2013 Feb 1]. Available from: <http://www.flhsmv.gov/dmv/vsifacts.html>.
3. GDNR (Georgia Department of Natural Resources). Wildlife resources information system. Department of Natural Resources. Georgia. Database; 2013 [cited 2013 Feb 1]. Available from: https://jca.activeoutdoorsolutions.com/gdnr_vrs/downloads/boatData.do.
4. Montes N, Swett R, Sidman C, Fik T. Offshore recreational boating characterization in the Southeast U. S. Florida Sea Grant, University of Florida; 2016. Report No. TP-226. Available from: https://www.flseagrant.org/wp-content/uploads/Tp_226_web.pdf.
5. Hain J, Ellis S, Seward P. Characterization of vessel traffic at the St John's and St. Mary's channel entrances, northeast Florida, January 1993. Final Report for MMC; 1994. Contract No. T94070460.
6. Gray DL, Canessa R, Keller P, Dearden P, Rollins R. Spatial characterization of marine recreational boating: Exploring the use of an on-the-water questionnaire for a case study in the Pacific Northwest. *Mar Policy*. 2011; 35(3):286–298.
7. Wu Y, Pelot R, Hillard C. The influence of weather conditions on the relative incident rate of fishing vessels. *Risk Anal*. 2009; 9(7):985–999.
8. Swett R, Sidman C, Fik T, Sargent B. Developing a spatial enabled inventory of recreational boats using vessel title registration data. *Coast Manag*. 2009; 37:405–420.
9. Gorzelany JF. A characterization of recreational boating activity and boater compliance with posted speed zones in Palm Beach County. Sea to Shore Alliance. 2013. Available from: http://discover.pbcgov.org/erm/Publications/BoaterSpeedComplianceStudy_DRAFT.pdf.
10. Bauduin S, Martin J, Edwards HH, Gimenez O, Koslovsky SM, Fagan DE. An index of risk of co-occurrence between marine mammals and watercraft: example of the Florida manatee. *Biol Conserv*. 2013; 159:127–136.
11. Swett R, Fann D, Staugler B. Regional waterway management system for Charlotte County, Florida. Florida Sea Grant, University of Florida; 2012. Report No. TP-181. Available from: https://www.flseagrant.org/wp-content/uploads/tp181_rwms_charlotte_county1.pdf.
12. Suitor D, Staiger J. Boating study and facility inventory. Collier County Environmental Service Division. Naples, Florida. 1994. Available from: <https://www.colliercountyfl.gov/home/showdocument?id=1019>.
13. Sidman C, Fik T, Sargent W. A Recreational Boating Characterization for Tampa and Sarasota Bays. Florida Sea Grant, University of Florida; 2004. Report No. TP-130. Available from: http://nsgl.gso.uri.edu/flsgp/flsgps04001/flsgps04001_full.pdf.
14. Sidman C, Swett R, Fik T, Fann S, Fann D, Sargent B. A Recreational Boating Characterization for the Greater Charlotte Harbor. Florida Sea Grant, University of Florida; 2005. Report No. TP-150. Available from: http://chnep.wateratlas.usf.edu/upload/documents/913_199_TP150_Charlotte_Harbor_Rec_Boating_web_revised.pdf.
15. Sidman C, Swett R, Fik T, Fann S, Sargent B. A Recreational Boating Characterization of Sarasota County. Florida Sea Grant, University of Florida; 2006. Report No. TP-152. 2006. Available from: <https://www.flseagrant.org/wp-content/uploads/flsgps06001.pdf>.
16. Sidman C, Fik T, Swett R, Sargent B, Fletcher J, Fann S, et al. A Recreational Boating Characterization of Brevard County. Florida Sea Grant, University of Florida; 2007. Report No. TP-160. Available from: <http://nsgl.gso.uri.edu/flsgp/flsgps07002.pdf>.

17. Box PW. Bottom-up simulation of recreational boat traffic monitoring [Dissertation]. Gainesville (FL): University of Florida; 1997.
18. Parrott L, Chion C, Martins CCA, Lamontagne P, Turgeon S, Landry JA, et al. A decision support system to assist the sustainable management of navigation activities in the St. Lawrence River Estuary, Canada. *Environ Model Softw*. 2011; 26:1403–1418.
19. Tixerant ML, Gourmelon F, Tissot C, Brosset D. Modeling of human activity development in coastal sea areas. *Coastal Conservation*. 2011; 5(4):407–416.
20. Sidman CF, Fik TJ. Modeling spatial patterns of recreational boaters: vessel, behavioral, and geographic considerations. *Leis Sci*. 2005; 27:1–15.
21. Dalton T, Thompson R, Jin D. Mapping human dimensions in marine spatial planning and management: an example from Narragansett Bay, Rhode Island. *Mar Policy*. 2010; 34:309–319.
22. Widmer WM, Underwood AJ. Factors affecting traffic and anchoring patterns of recreational boats in Sydney Harbour, Australia. *Lands Urban Plan*. 2004; 66:173–183.
23. National Vessel Documentation Center database [Internet]. NOAA. U.S. Coast Guard. 2014 [cited 2014 Feb 1]. Available from <https://www.st.nmfs.noaa.gov/st1/CoastGuard/VesselByName.html>.
24. ESRI. ArcGIS Desktop: Release 10.3. Redlands, CA: Environmental Systems Research Institute; 2015.
25. Etter JF, Perneger TV. Analysis of non-response bias in a mailed health survey. *Journal of clinical epidemiology*. 1997; 50(10):1123–8. PMID: 9368520
26. NASA Ocean Biology (OB). 2014 Moderate Resolution Imaging Spectroradiometer (MODIS) Ocean Color Data, 20154.0 Reprocessing. NASA OB.DAAC, Greenbelt, MD, USA. http://oceancolor.gsfc.nasa.gov/DOCS/modis_sst/. Maintained by NASA Ocean Biology Distributed Active Archive Center (OB.DAAC), Goddard Space Flight Center, Greenbelt MD.
27. National Centers for Environmental Information [Internet]. NOAA. 2013 [cited 2013 Apr 15]. Available from: <http://www.ngdc.noaa.gov/mgg/mggd.html>.
28. Data Library [Internet]. Florida Geographic Data Library. 2014 [cited 2014 Apr 1]. Available from <http://www.fgdl.org/metadataexplorer/explorer.jsp>.
29. Goslee SC, Urban DL. The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Softw*. 2007; 22:1–19.
30. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013. Available from: <http://www.R-project.org/>.
31. Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer Research*. 1967; 27:209–220. PMID: 6018555
32. Goslee S. Correlation analysis of dissimilarity matrices. *Plant Ecol*. 2010; 206:279–286.
33. Cameron P. Permutation group. London Mathematical Society. Cambridge University Press; 1999.
34. Wood SN. Generalized additive models: an introduction with R. Chapman and Hall/CRC; 2006.
35. Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. Mixed effects models and extensions in ecology with R. Springer; 2009.
36. Venables WN, Ripley BD. Modern applied statistics with S. Fourth Edition. Springer, New York; 2002.
37. Fox J, Weisberg S. An {R} companion to applied regression, Second Edition. Thousand Oaks CA: Sage. URL: <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>. 2011.
38. Zeileis A, Kleiber C, Jackman S. Regression Models for Count Data in R. *J Stat Softw*. 2008; 27(8).
39. Wood SN. mgcv: GAMs and generalized ridge regression for R. *R News*. 2001; 1:20–25. Available from from <http://cran.r-project.org/web/packages/mgcv/index.html>.
40. Pinheiro J, Bates D, DebRoy S, Sarkar D and R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1–121. 2015. Available from: from <http://cran.r-project.org/web/packages/nlme>.
41. Barto K. MuMin: Model selection and model averaging based on information criteria (AICc and alike). R package version 1.42.1–2015. Available from: from <http://cran.r-project.org/web/packages/MuMin>.
42. Smith EP, Rose KA. Model goodness-of-fit analysis using regression and related techniques. *Ecol Modell*. 1995; 77:49–64.
43. Piñeiro G, Perelman S, Guerschman JP, Paruelo J M. How to evaluate models: observed vs. predicted or predicted vs. observed?. *Ecol Modell*. 2008; 216(3–4), 316–322.
44. Hollowed AB, Barbeaus SJ, Cokelet ED, Farley E, Kotwicki S, Ressler PH, et al. Effects of climate variations on pelagic ocean habitats and their role in structuring forage fish distribution in the Bering Sea. *Deep Sea Res Part II Top. Stud. Oceanogr*. 2012; 65–70:230–250.
45. Pinheiro JC, Bates DM. Mixed-effects models in S and S-Plus. Springer; 2000.

46. Faraway JJ. Extending the linear model with R. Chapman and Hall/CRC;2006.
47. Bates D, Maechler M, Bolker BM and Walker S. Fitting Linear Mixed-Effects Models using lme4. J Stat Softw. 2015; 67(1),1–48.
48. Dutilleul P, Stockwell JD, Frigon D, Legendre P. The mantel test versus Pearson's correlation analysis: assessment of the differences for biological and environmental studies. J Agric Biol Environ Stat. 2000; 5(2):131–150.
49. Gomes LC, Miranda LE, Agostinho AA. Fishery yield relative to chlorophyll *a* in reservoirs of the Upper Parana River, Brazil. Fish Res. 2002; 55:335–340.
50. Choudhury SB, Jena B, Rao MV, Rao KH, Somvanshi VS, Gulati DK, et al. Validation of integrated potential fishing zone (IPFZ) forecast using satellite based chlorophyll and sea surface temperature along the east coast of India. Int J Remote Sens. 2007; 28(12):2683–2693.
51. Wall CC, Muller FE, Roffer MA. Linkages between environmental conditions and recreational king mackerel (*Scomberomorus cavalla*) catch off west-central Florida. Fish Oceanogr. 2009; 18(3):185–199.
52. Buchanan CC. Effects of an artificial habitat on the marine sport fishery and economy of Murrells Inlet, South Carolina. Marine Fisheries Review. 1973; 35(9):15–22.
53. Bohnsack JA, Sutherland DL. Artificial reef research: a review with recommendations for future priorities. Bull. Mar. Sci. 1985; 37(1):11–39.
54. Bohnsack JA. Are high densities of fishes at artificial reefs the result of habitat limitation or behavioral preferences? Bull. Mar. Sci. 1989; 44(2):631–645.
55. National Centers for Environmental Information, State of the Climate: Hurricanes & Tropical Storms for Annual 2011. NOAA; 2012 [cited 2015 Jun 18]. Available from <http://www.ncdc.noaa.gov/sotc/tropical-cyclones/201213>.
56. National Centers for Environmental Information, State of the Climate: Hurricanes & Tropical Storms for Annual 2012. NOAA; 2013 [cited 2015 Jun 18]. Available from <http://www.ncdc.noaa.gov/sotc/tropical-cyclones/201213>.
57. Vaske JJ. Survey research and analysis: Applications in parks, recreation and human dimensions. Venture Publ.; 2008.