Using GIS to Model Probability of Shark Attacks on Florida's Coast

Laurel Whitney '16, Emily Walker '16
Environmental Studies Program, Colby College, Waterville, Maine

Abstract

Shark attacks represent a major human wildlife conflict in Florida. Florida consistently leads the United States in shark attacks annually, yet these shark attacks are not evenly distributed along the coastline. This exploratory analysis investigated various attributes to determine if some characteristics of a particular area off the coast correlated with higher shark attack incidents. Using geographic information system analysis and statistical tests, this project found proximity to beach, shorelines, and flood areas to be significant predictors of shark attack locations. This exploratory analysis, suggests that more rigorous spatial models may be able to predict shark human conflict in the future.

Introduction

Globally, shark attacks represent the most dispersed human wildlife conflict (Neff 2012). Shark attack incidents have been steadily increasing worldwide, which many scientists attribute to human population growth (Fleshler 2015). Mitigation strategies require decision-making from a number of parties including coastal managers, scientists, policymakers and conservationists, who have to balance public interest and ecological sustainability (Neff 2012). Florida represents a site of particular interest as it led the United States in shark attacks in 2014, with 28 incidents and no fatalities (Fleshler 2015). We utilized geographic information system (GIS) analysis and statistical tests to determine significant factors influencing shark attacks to model the probability of shark attacks along Florida's coast.

Methods

ArcGIS 10.2.2 by ESRI was the software used to conduct the spatial analysis on shark attacks. All data were projected in Albers Conical Equal Area (the projection used by the Florida Geographic Data Library).

Data of shark attack coordinates from 1984 to 2012 were obtained from ESRI database (Boehme 2012). Mangroves, Flood Areas, Beaches, FL County Boundaries, bathymetry and marine sanctuary data were downloaded from Florida's Geographic Data Library (www.fgdl.org). Temperature data was collected from NOAA.

First, shark data was reprojected into Albers Conical Equal area to fit the Florida standard. A 25 mile buffer was created around the coast of Florida, which was used to clip out the shark attack data points. The shark points that remained were manually edited removing points that were off the coast of Georgia or Alabama and attacks that occurred at aquariums. To select some points to create the model and save some attacks to test the model, we sorted the shark attack points by name of person attacked in the attribute table and deleted the first 70 of 139 points (approximately 50%). Sixty-nine points were used to create the model and the other 67 were saved to test the model. The city data was also edited by selecting only cities with populations >50,000 and defining them as population "hot spots." The nearest distance from each shark attack point to shore was used to determine the furthest shark attack from shore. The furthest point was 10.285 km. This information was used to create a 10.286 km buffer on the outside of the Florida polygon (created by dissolving FL County Boundaries). The buffer was cut to remove the buffer around the Florida/Georgia and Florida/ Alabama border. Sixty-nine random points were created within this buffer. The distance from shark points and random points to coast, beach, flood area, marine protected area, mangroves, and cities was calculated using the near tool. To analyze the relationship between shark attacks and bathymetry data, the contour lines of the bathymetry were first converted to points and these points were interpolated to create a raster layer with 30 m cell size. The values were extracted to shark attack points and random points. Temperature raster values were also extracted from the these points. For each variable a Mann-Whitney U test was performed using SPSS to determine if the shark attack points were significantly different from the random points. All distance values were converted from meters to kilometers and then SPSS was used to run a binary logistic regression to create a formula for the probability of a shark attack. The logistic function was used to create a raster layer with each cell value equal to the probability of shark attack based on distance to significant variables. The values of this raster layer were extracted to shark attack points that were not utilized to make the model and a new set of randomly generated points. Finally, a Mann-Whitney U test was performed to evaluate if the model was a good predictor of these shark attacks. To create a visual comparison to our model, we created a kernel density map of every shark attack between 1984 and 2012.

Results

The attributes that were found to be significant predictors of shark attacks from the Mann-Whitney U tests were proximity to shoreline, proximity to a major city, proximity to a beach, proximity to a flood area, and proximity to mangroves ($p<0.05$) (Table 1 and Figure 1). Next, a logistic regression was run with these attributes as the independent variables and the dependent variable was whether or not there was a shark attack. This analysis determined that only proximity to beach, proximity to shoreline, and proximity to flood area were significant factors in determining the probability of a shark attack. A second logistic regression run with just three variables confirmed that they were significant ($p<0.053$) (Table 2). A Hosmer and Lemeshow test demonstrated that the there was a significant difference between observed shark attack incidents and expected shark attacks (Chi-square=43.619, $p<0.0001$). An Omnibus test of Model Coefficients determined model improvement over base was significant (Chi-square=71.067, $p<0.0001$). The R^2 value was 0.59, meaning the model can account for 59% of the variance. The logistic function was modeled as:

$$P(\text{sharks})= 1/\ (1+ e^{-(1.569-0.093[\text{proximity to beach}]-0.771[\text{proximity to coast}]+0.208[\text{proximity to flood area}]})$$

Finally, a Mann-Whitney U test was utilized to test the logistic regression model. The value of the regression model was significantly higher at points of shark attacks (not included in making the regression) compared to randomly generated points ($z=-8.064$, $p<0.0001$) (Figure 3).

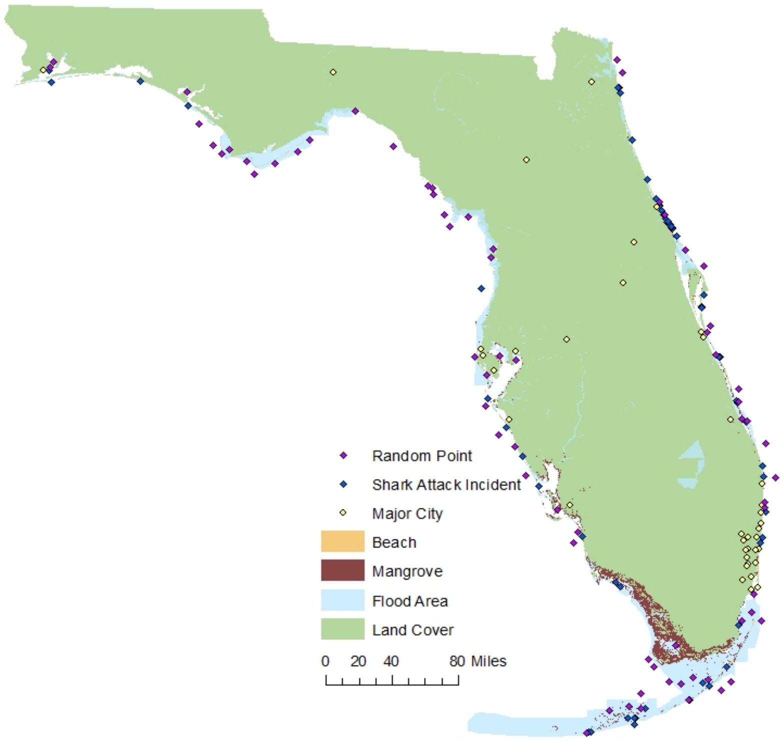

Figure 1. The attributes that were found to be significantly different in distance from shark attack incidents in comparison to random points by Mann-Whitney U tests.

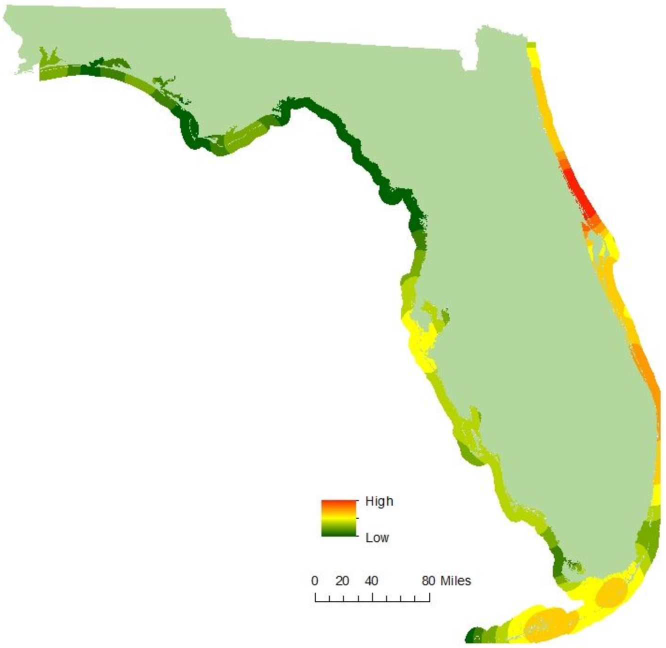

Figure 2. The density of shark attacks between 1984 and 2012 that occurred within 10.286 kilometers of Florida's coast.

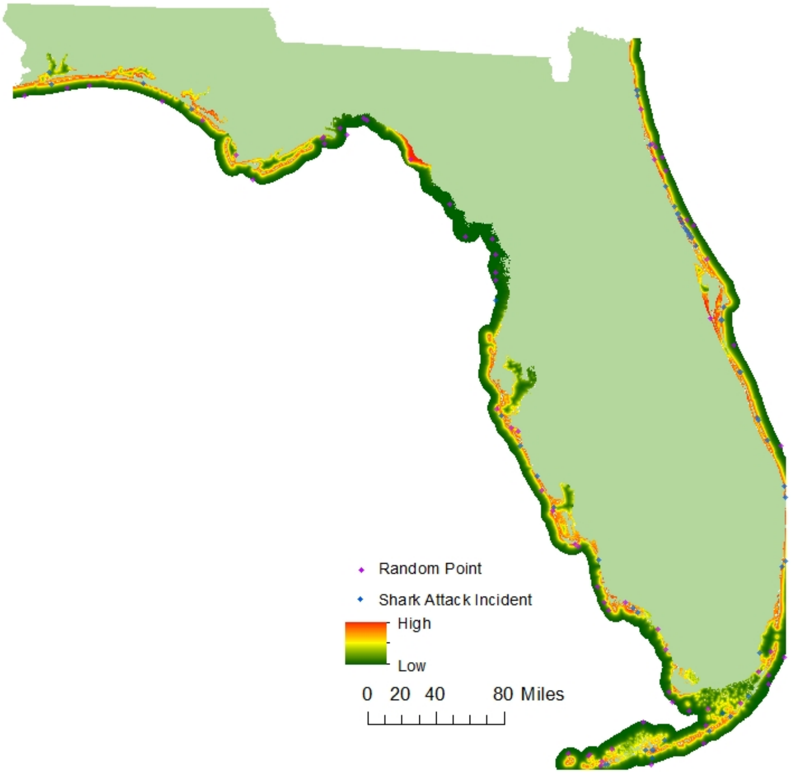

Figure 3. Probability of shark attack based on logistic function model. The model was tested using shark attacks between 1984 and 2012 that were not used to create the model.

Table 1. The results of Mann-Whitney U tests comparing various attributes in relation to shark attack incidents and random points. *significant difference

Attribute	Z-score	P-value
Proximity to Shoreline	-7.033	<0.0001*
Proximity to Major City (>50,000 people)	-2.127	0.033*
Proximity to Beach	-6.781	<0.0001*
Proximity to Flood Area	-3.540	<0.0001*
Proximity to Mangroves	-4.840	<0.0001*
Proximity to Marine Protected Area	-0.358	0.720
Gulf Coast vs. Atlantic Coast	<0.0001	1.000
Ocean Temperature	-0.331	0.741
Ocean Depth	-0.813	0.416

Table 2. Variables in logistic function for probability of shark attack

Attribute	Coefficient	P-Value
Proximity to Beach	-0.093	0.006
Proximity to Shoreline	-0.771	<0.0001
Proximity to Flood Area	0.208	0.052
Constant	1.569	<0.0001

Discussion

First, our model shows that there is a higher probability of a shark attack occurring closer to shores, beaches and flood areas (Figure 3). The increasing probability of a shark attack occurring closer to beaches and shores can potentially be attributed to more people being present in these areas, leading to a higher chance of conflict. It is also possible that proximity to flood areas increase the probability of a shark attack because they are close to the shoreline. Alternatively, the heightened shark attacks near flood areas could be due to increased turbidity in these areas, leading sharks to mistake humans as prey items.

Ideally, our model (Figure 3) would resemble the map of the concentration of observed shark attacks (Figure 2). Generally, the model does show similar trends to the observed shark attacks, such as the low probability (green) areas in the elbow of Florida. However, on the northeast coastline of Florida, there have been a lot of shark attacks (Figure 2), which is not predicted by our model (Figure 1), suggesting that our model is not extremely accurate. Additionally, The Hosmer and Lemeshow test suggests that our model did not fit a logistic curve well, which may contribute to the differences between the two maps.

Our model could be improved my determining more attributes that were significant in predicating shark attacks, but not related to proximity to shore. The only attributes that we were able to find as significant predictors were all essentially a measure of distance to shore, decreasing the precision of our model. If attributes such as shark migration routes, shark breeding areas, and ideal habitat were included, they could reveal a more useful model of shark attacks.

Our model demonstrates that human prevalence is a large factor in shark attack incidents, which has been supported by past research. George Burgess, director of the Florida Program for Shark Research attributes the growth in human population to the rise in shark attacks (Fox News 2013). He believes that "[t]he human population is getting so much bigger that we're literally swamping sharks out of the water and dictating the number of encounters" (Fox News 2013). For example, New Smyrna Beach, in Volusia county, Florida leads the U.S. in annual shark attacks and also is a hotspot for swimmers, surfers and tourists (Eilperin 2013). As human population growth continues, shark attack incidents will likely continue to escalate. This model is purely an exploratory analysis, but it is possible to develop rigorous spatial models to predict shark-human conflict in the future.

References

Boehme, C. 2012. ArcGIS. Shark Attacks: All Documented Shark-Person Interactions Since 1982 Categorized by Injury. http://www.arcgis.com/home/item.html?id=cf561a64e4f348418376b4cb516f026b

Eilperin, J. 2013. Once bitten, twice shy: our exaggerated fear of shark attacks. The Guardian.

Fleshler, D. 2015. Florida again tops U.S. with 28 shark bites in 2014. Sun-sentinel.

Fox News. 2013. Shark attacks – and attack fatalities – up worldwide, expert says.

Neff, C. 2012. Australian Beach Safety and the Politics of Shark Attacks. Coastal Management 40:98-116

Acknowledgements

We would like to acknowledge Philip Nyhus and Manny Gimond for their support and information throughout the semester. Without their help this project would not have been possible.