

Charleston Shipping Lanes Risk Analysis

Christopher J. Fonnesebeck, Ph.D.

August 3, 2015

Research Question

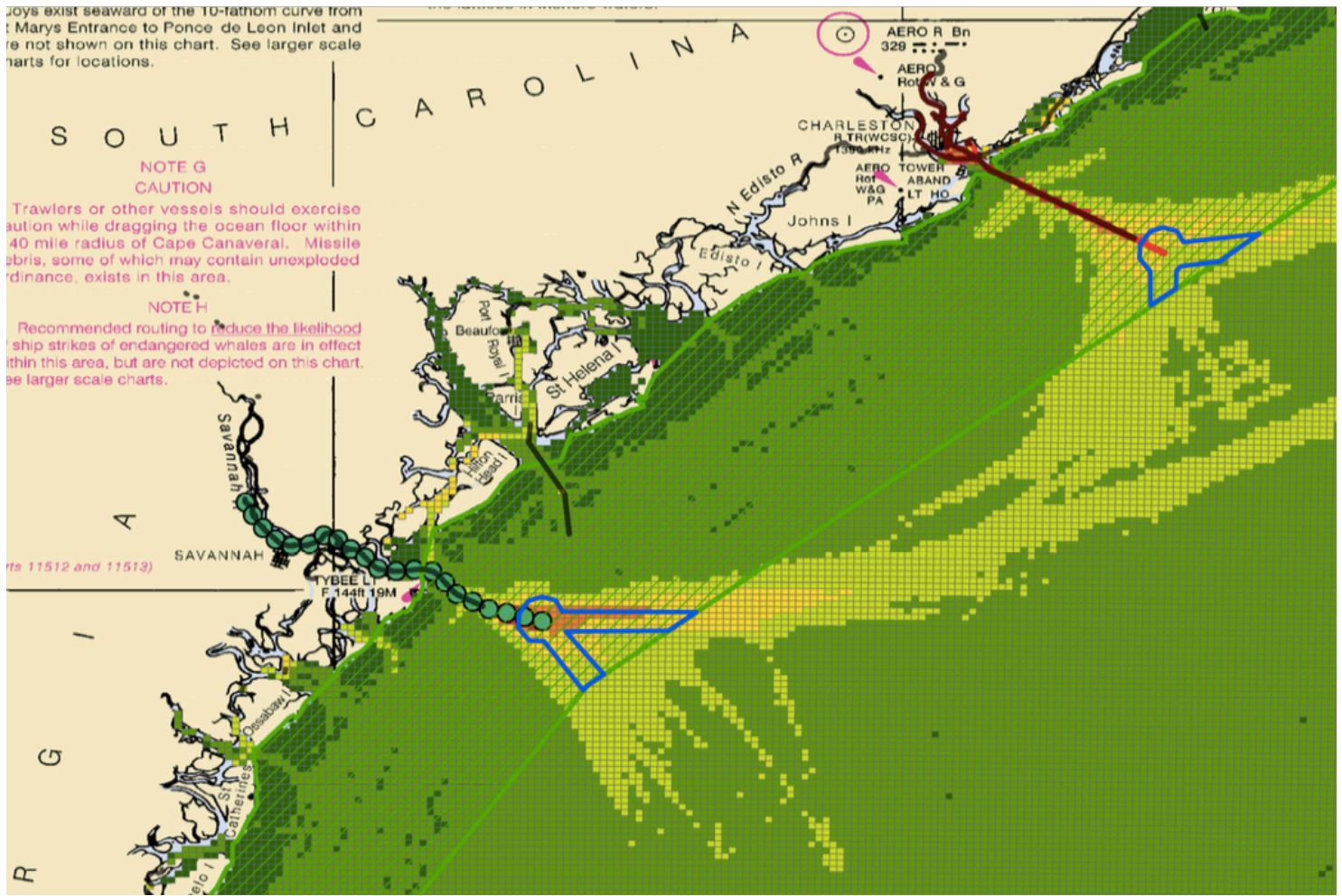
Presently, there are no recommended lanes off Charleston. Would recommended lane/s reduce the relative risk of interactions between commercial vessels and right whales? Can it be determined? If so, what is the expected reduction in risk of interactions and where should the lanes be placed?

Modeling Approach

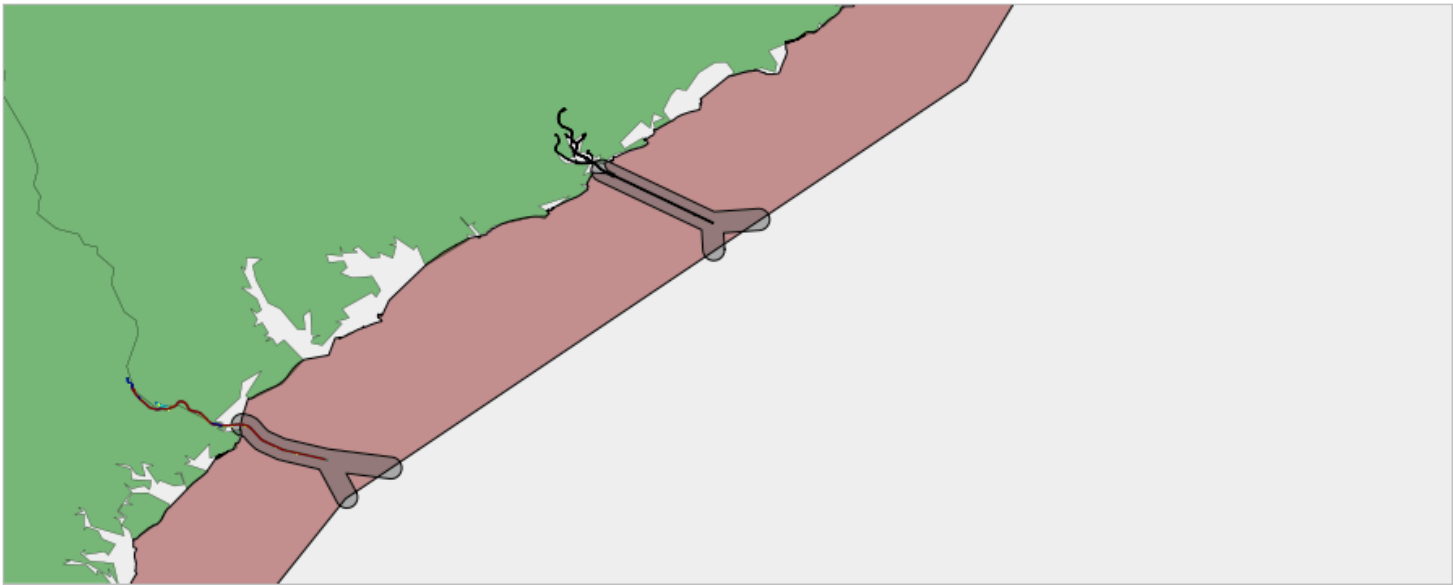
Following Fonnesebeck *et al.* (2008), we constructed a predictive model of right whale encounter probability that is based on existing aerial survey data and relevant environmental predictors of whale habitat use, such as bathymetry and water temperature. By integrating automatic identification system (AIS) data, the model was used to evaluate the use of shipping lanes to reduce the risk to the migratory northern right whale population relative to existing patterns of traffic. The metric for risk for a particular lane designation was the expected co-occurrence rate of whale groups and ships over the total seasonal management area (SMA) associated with the ports of Charleston and Savannah. To avoid having to account for small-scale factors related to the interaction of whales and ships, risk was estimated at a relatively coarse scale. Specifically, I estimated the expected occurrence probability of whale groups over the cells of a 3 x 3 nmi grid in the southern part of the study area, and a 4 x 4 nmi grid in the north.

A key component of the modeling approach was the estimation of the rate of encounter with whale groups in each grid cell. These estimates are informed by aerial survey data, which includes both sightings and on-watch effort. The key parameter of interest is the encounter rate, which was used to model the survey observations as a Poisson random variables, using the survey effort as an offset. To aid the predictive performance of the model, we used two covariates to estimate encounter rates; specifically, we incorporated the mean sea surface temperature and mean grid cell depth as non-parametric regression terms to predict whale encounters, using cubic splines. This allowed for arbitrary non-linear relationships between either covariate and whale habitat use. To account for spatial autocorrelation, we included a conditional auto-regressive (CAR) function for the cells of the habitat grid, allowing information from adjacent grid cells to help inform one another, thereby improving model estimates. Observations among nearby cells are likely not independent, with whales sighted in one cell being very likely to visit adjacent cells within any 2-week period characterized by the aggregated survey data.

The lanes that we are evaluating are illustrated in the figure below.



Map of candidate lanes



Lanes plotted with navigation channels and seasonal management area.

Semi-parametric conditional autoregressive model of whale occurrence

One of the key features of this analysis is the modeling of spatial autocorrelation via conditional autoregression methods. Since the number of surveys and sightings are limited, we need to borrow strength among the grid cells of the region, recognizing that an animal observed in one cell is likely to also be in neighboring cells within the same period. The CAR model accounts for the spatial covariance directly, here estimating a risk surface of whale occurrence. We implemented a first-order conditional autoregressive model, which expresses the effect in the current cell k as a Gaussian-distributed random variable, with a mean equal to the weighted average of the direct neighbors (typically 8 cells).

$$\phi_k | \phi_{-k} \sim N \left(\frac{\sum_{i=1}^n w_{ki} \phi_i}{\sum_{i=1}^n w_{ki}}, \frac{\tau^2}{\sum_{i=1}^n w_{ki}} \right)$$

Here, we equally weight all neighbors, $w_{ki} = 1 \forall i, k$.

It is natural to model counts of surveyed groups y_{it} in location i during period t as a Poisson-distributed random variable, with some expected rate of occurrence θ_{it} that is multiplied by the exposure u_{it} (the number of surveys flown over the location during the period):

$$y_{it} \sim \text{Poisson}(\theta_{it} u_{it})$$

While I assume that the rate of encounter π_i is relatively constant within each month and grid cell, this assumption is likely not valid over longer periods of time and over space, due to seasonal movement by migratory whales. This leads to modeling unique encounter probabilities for each cell in each month (December to March) in each year. We characterize spatial variability as a function of physical variables, mean sea surface temperature s and mean water depth d . Hence, θ_{it} is expressed as a hierarchical function:

$$\log(\theta_{it}) = s_{sst}(x_{it}) + s_{depth}(d_i) + \phi_i.$$

where x_{it} and d_i are the mean sea surface temperature and mean depth for location i at time t , and the s functions are their associated spline models. The ϕ_i term is a conditional auto-regressive factor that is described below, which models the covariance among neighboring grid cells. The log transformation is applied to constrain estimates of encountered rate to positive values.

In order to avoid making assumptions about the functional form of the relationship between spatial covariates (water depth and sea surface temperature) and the encounter rate with whale groups. For each variable, we specify a 5-knot spline function over range of the variables. The result is a flexible, non-parametric function that is not constrained by assumptions of form.

Estimation of risk

Consistent with the goals of this research, risk may be defined as the co-occurrence (in time and space) of whales and vessels. Thus, we seek to identify routes through cells that predict relatively few whale occurrences throughout the calving season, according to the predicted encounter rates. A reasonable risk function, then, is the vessel track count per cell, weighted by the probability of whale occurrence and summed over all cells and months:

$$\hat{\rho} = \sum_t \sum_i \hat{\theta}_{it} \hat{v}_{it}$$

where \hat{v}_{it} is the estimated number of vessels traversing cell i during month t . Thus, risk is a monotonically increasing function of both vessel density and encounter rate, and is zero for any cells in which either quantity is zero. An obvious and important assumption of this measure is a direct relationship between the encounter rates estimated from survey data and actual encounter rates with ships. These rates are surely not identical, but I assume that they are directly and positively correlated.

To assess the relative risk of implementing a ship lane, we compared the configuration depicted above to the actual MSR traffic from seasons 2010–11 through 2012–13, using the estimated risk metric above. In order to do so, each transit in the AIS dataset was assigned to a particular lane, depending on whether it passed through the navigation channel for either port, and to a “branch” (north or south) within the lane, depending on whether it was coming from or leaving to that direction in relation to the port. This was used to represent how the traffic might have been had these lanes been implemented during that time period (assuming 100% adherence to the lanes).

For reference, the [full analysis](#)

(<https://github.com/fonnesbeck/CharlestonLanesAnalysis/blob/master/Charleston%20Shipping%20Lanes%20Analysis.ipynb>)

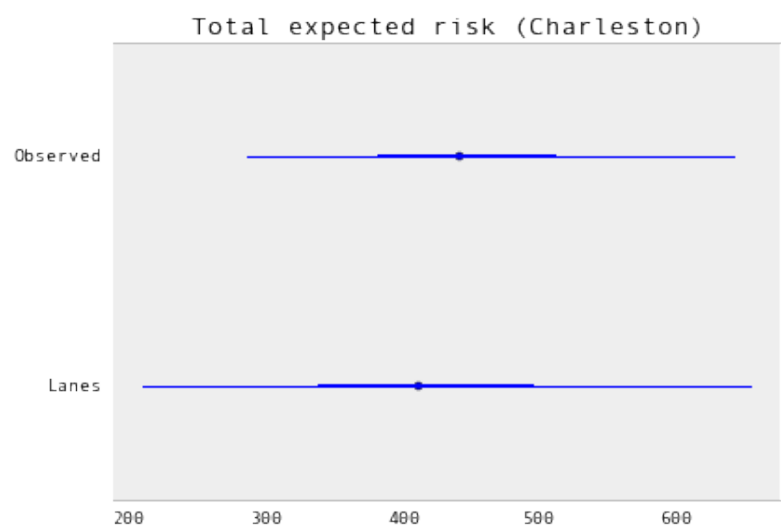
is available in a public [GitHub repository](#) (<https://github.com/fonnesbeck/CharlestonLanesAnalysis>). It includes all the code, data and output from the current model.

Model fitting

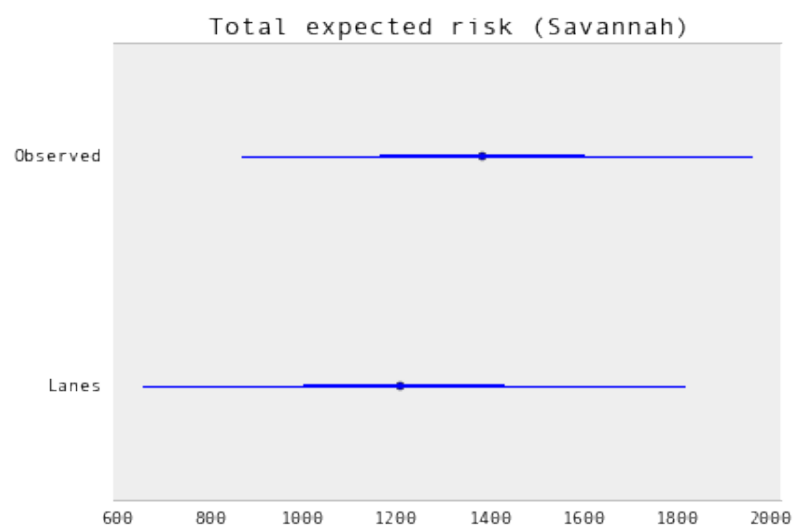
The model was fit using Markov chain Monte Carlo (MCMC, Brooks et al. 2011) methods, implemented in the PyMC 2.3 software package (Patil et al. 2010). The model was run for 500,000 iterations, with the first 400,000 conservatively discarded as burn-in.

Results and Conclusions

The estimates of total risk for both ports under observed traffic patterns and the candidate lanes are shown in the two figures below.

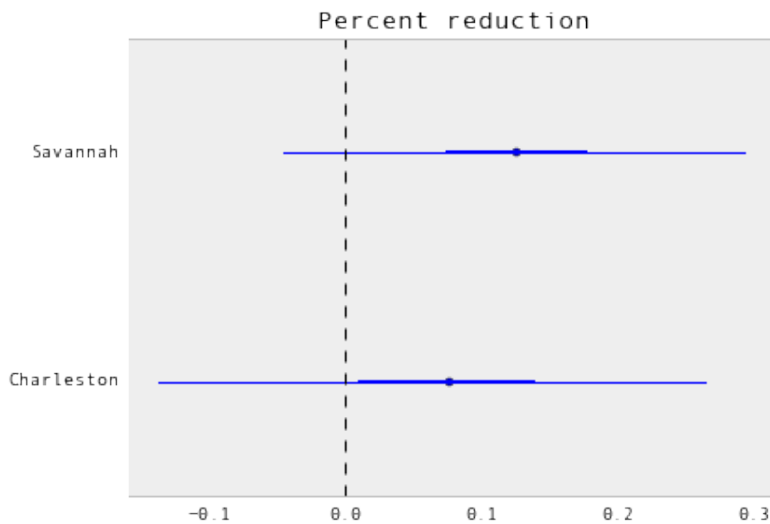


Total expected risk for Charleston, with and without lanes, with 95% posterior credible intervals.



Total expected risk for Savannah, with and without lanes, with 95% posterior credible intervals.

The reduction of risk resulting from the implementation of lanes is the difference between these estimates. The differences are plotted as the proportional reduction for both ports:



Estimated reduction in risk by implementing lanes, for both ports

Implementing the lanes specified here reduces the risk of whale/vessel co-occurrence by a mean estimated value of 12 percent in Savannah and by 6 percent in Charleston, relative to observed traffic during the three calving seasons 2010–11, 2011–12 and 2012–13. There is, however, a substantial amount of residual uncertainty; the 95% posterior credible interval of both of these estimates include zero (*i.e.* no difference) and include values as high as 25–30% reduction in risk.

The relatively modest effect of the lanes may be due to the fact that the lanes consist of the navigation channel, plus a short extension to the eastern edge of the SMA. As a result, much of the traffic ends up occupying the navigation channels in both scenarios, save for a few kilometers at the eastern end of the routes. Hence, much of the risk (as defined by our formula) is in the navigation channels, and the marginal effects of determining how vessels reach the channels are modest. Nevertheless, a 10–15% reduction in risk may be sufficient to warrant their use.

References

1. Fonnesbeck CJ, Garrison LP, Ward-Geiger LI, Baumstark RD. Bayesian hierarchical model for evaluating the risk of vessel strikes on North Atlantic right whales in the SE United States. *Endangered Species Research*. 2008;6:87–94. doi:10.3354/esr00134.
2. Patil A, Huard D, Fonnesbeck CJ. PyMC: Bayesian Stochastic Modelling in Python. *J Stat Softw*. 2010;35(4):1–81.
3. Brooks S, Gelman A, Jones G, Meng X-L. *Handbook of Markov Chain Monte Carlo*. CRC Press; 2011.