Development of a species distribution model for fin whales (*Balaenoptera physalus*) within a Bayesian hierarchical framework: Implications for uncertainty

Species distribution models (SDMs) have proven to be an integral tool in the conservation and management of cetaceans. Many applications have adopted a two-step approach where a detection function is estimated using conventional distance sampling in the first step and subsequently used as an offset to a density-habitat model in the second step. A drawback to this approach, hereafter referred to as the conventional species distribution model (CSDM), is the difficulty in propagating the uncertainty from the first step to the final density estimates. We describe a Bayesian hierarchical species distribution model (BHSDM) which has the advantage of simultaneously propagating multiple sources of uncertainty. Our framework includes 1) a mark-recapture distance sampling observation model that can accommodate two team line transect data, 2) an informed prior for surface availability 3) spatial smoothers using spline-like bases and 4) a compound Poisson-gamma likelihood which is a special case of the Tweedie distribution. We compare our approach to the CSDM method using a simulation study and a case study of fin whales (*Balaenoptera physalus*) off the East Coast of the USA. Simulations showed that the BHSDM method produced estimates with lower precision but with confidence interval coverage closer to the nominal 95% rate (94% for the BSHDM vs 85% for the CSDM). Results from the fin whale analysis showed that density estimates and predicted distribution patterns were largely similar among methods. Abundance estimates were also similar though modestly higher for the CSDM (4700, CV=0.13) than the BHSDM (4526, CV=0.26). Estimated sampling error differed substantially among the two methods where the average CV for density estimates from BHSDM method was approximately 3.5 times greater than estimates from the CSDM method. Successful wildlife management hinges on the ability to properly quantify uncertainty. Underestimates of uncertainty can result in ill-informed management decisions. Our results highlight the additional sampling uncertainty that is propagated in a hierarchical framework. Future applications of SDMs should consider techniques that allow all sources of error to be fully represented in final density predictions.

Development of a species distribution model for fin whales (*Balaenoptera physalus*) within a Bayesian hierarchical framework: Implications for uncertainty

Douglas B Sigourney Corresp., [1] , **Samuel Chavez-Rosales** [1] , **Paul Conn** [2] , **Lance Garrison** [3] , **Elizabeth Josephson** [1] , **Debra Palka** [4]

[1] Integrated Statistics, Woods Hole, MA, USA

[2] National Marine Mammal Laboratory, NOAA, National Marine Fisheries Service, Alaska Fisheries Science Center, Seattle, US

[3] NOAA Southeast Fisheries Science Center, Miami, FL, US

[4] NOAA Northeast Fisheries Science Center, Woods Hole, MA, US

Corresponding Author: Douglas B Sigourney
Email address: douglas.sigourney@noaa.gov

Species distribution models (SDMs) have proven to be an integral tool in the conservation and management of cetaceans. Many applications have adopted a two-step approach where a detection function is estimated using conventional distance sampling in the first step and subsequently used as an offset to a density-habitat model in the second step. A drawback to this approach, hereafter referred to as the conventional species distribution model (CSDM), is the difficulty in propagating the uncertainty from the first step to the final density estimates. We describe a Bayesian hierarchical species distribution model (BHSDM) which has the advantage of simultaneously propagating multiple sources of uncertainty. Our framework includes 1) a mark-recapture distance sampling observation model that can accommodate two team line transect data, 2) an informed prior for surface availability 3) spatial smoothers using spline-like bases and 4) a compound Poisson-gamma likelihood which is a special case of the Tweedie distribution. We compare our approach to the CSDM method using a simulation study and a case study of fin whales (*Balaenoptera physalus*) off the East Coast of the USA. Simulations showed that the BHSDM method produced estimates with lower precision but with confidence interval coverage closer to the nominal 95% rate (94% for the BSHDM vs 85% for the CSDM). Results from the fin whale analysis showed that density estimates and predicted distribution patterns were largely similar among methods. Abundance estimates were also similar though modestly higher for the CSDM (4700, CV=0.13) than the BHSDM (4526, CV=0.26). Estimated sampling error differed substantially among the two methods where the average CV for density estimates from BHSDM method was approximately 3.5 times greater than estimates from the CSDM method. Successful wildlife management hinges on the ability to properly quantify uncertainty. Underestimates of uncertainty can result in ill-

informed management decisions. Our results highlight the additional sampling uncertainty that is propagated in a hierarchical framework. Future applications of SDMs should consider techniques that allow all sources of error to be fully represented in final density predictions.

Development of a Species Distribution Model for Fin Whales (*Balaenoptera physalus*) within a Bayesian Hierarchical Framework: Implications for Uncertainty

Douglas B. Sigourney[1], Samuel Chavez-Rosales[1], Paul Conn[2], Lance Garrison[3], Elizabeth Josephson[1] and Debra Palka[4]

[1] Integrated Statistics, 16 Sumner Street, Woods Hole, MA 02543

[2]National Marine Mammal Laboratory, NOAA, National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115

[3]NOAA Southeast Fisheries Science Center, 75 Virginia Beach Drive, Miami, FL 33149

[4] NOAA Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543

Corresponding Author: Douglas B. Sigourney

Email address: douglas.sigourney@noaa.gov

Abstract

Species distribution models (SDMs) have proven to be an integral tool in the conservation and management of cetaceans. Many applications have adopted a two-step approach where a detection function is estimated using conventional distance sampling in the first step and subsequently used as an offset to a density-habitat model in the second step. A drawback to this approach, hereafter referred to as the conventional species distribution model (CSDM), is the difficulty in propagating the uncertainty from the first step to the final density estimates. We describe a Bayesian hierarchical species distribution model (BHSDM) which has the advantage of simultaneously propagating multiple sources of uncertainty. Our framework includes 1) a mark-recapture distance sampling observation model that can accommodate two team line transect data, 2) an informed prior for surface availability 3) spatial smoothers using spline-like bases and 4) a compound Poisson-gamma likelihood which is a special case of the Tweedie distribution. We compare our approach to the CSDM method using a simulation study and a case study of fin whales (*Balaenoptera physalus*) off the East Coast of the USA. Simulations showed that the BHSDM method produced estimates with lower precision but with confidence interval coverage closer to the nominal 95% rate (94% for the BSHDM vs 85% for the CSDM). Results from the fin whale analysis showed that density estimates and predicted distribution patterns were largely similar among methods. Abundance estimates were also similar though modestly higher for the CSDM (4700, CV=0.13) than the BHSDM (4526, CV=0.26). Estimated sampling error differed substantially among the two methods where the average CV for density estimates from BHSDM method was approximately 3.5 times greater than estimates from the CSDM method. Successful wildlife management hinges on the ability to

properly quantify uncertainty. Underestimates of uncertainty can result in ill-informed management decisions. Our results highlight the additional sampling uncertainty that is propagated in a hierarchical framework. Future applications of SDMs should consider techniques that allow all sources of error to be fully represented in final density predictions.

Introduction

Species distribution models (SDMs) have become valuable tools to help characterize the spatial distribution and abundance of many species (Elith & Leathwick, 2009) and have provided critical information to help guide management decisions of cetacean populations (Forney et al., 2012; Roberts et al., 2016). The rapid development of techniques for fitting SDMs to data has provided multiple options and the need to evaluate their advantages and limitations. Some attempts have been made to compare methods with a focus on prediction accuracy (Elith & Graham, 2009; Oppell et al., 2012). Uncertainty, however, is rarely addressed when evaluating SDMs (Robinson et al., 2017). Because proper assessment of uncertainty is critical to effective management of cetacean populations (Taylor et al., 2000) attempts to model their distribution and abundance should carefully consider the ability of the chosen method to quantify uncertainty.

For cetaceans, line transect data are often used to fit SDMs. Fitting SDMs to these data can be challenging because relationships between habitat variables and density are often nonlinear and subject to unexplained variance. In addition, not all animals are detected on the trackline and therefore probability of detection needs to be taken into account. Diving animals such as cetaceans offer particular challenges because detection can be influenced by two independent factors. Surface detectability refers to the probability observers detect animals that are at the surface. This probability can be estimated using conventional distance sampling techniques (Buckland et al., 2001). Surface availability refers to the probability that animals are at the surface and therefore available for detection. Surface availability is not as easily estimated from line transect data alone and requires additional information on diving behavior (Langrock et al., 2013).

One common approach is to use a two-step method where the detection function is estimated using conventional distance sampling techniques in the first step and used as an offset when relating observed animal counts to habitat covariates in the second step (Miller at al., 2013). Generalized additive models (GAMs) are commonly used in the second step due to their flexibility to capture non-linear density-habitat relationships and flexible distributions such as the negative binomial or Tweedie distribution can be adopted to model overdispersion. This method, hereafter referred to as the conventional species distribution model (CSDM) method, has proven quite robust (Forney et al., 2012; Roberts et al., 2016). However, it is difficult to propagate uncertainty from the detection function that is estimated in the first step to the final density predictions that are made in the second step. Bootstrap techniques can be used to estimate uncertainty (Hedley & Buckland, 2004), but this method requires resampling the data multiple times and coverage can be poor (Miller et al., 2013). Williams et al. (2011) used a random effects approach to propagate error from the detection function but their method is limited to cases where there is only a single team of observers. As an alternative, a number of studies have adopted a Bayesian hierarchical framework.

Hierarchical analysis of distance sampling data has also been developed in the literature (Royle, Dawson & Bates 2004; Royle & Dorazio, 2008). This approach integrates over the uncertainty in the detection function, effectively propagating the uncertainty into final density estimates (Miller et al., 2013). When estimated in a Bayesian framework, prior information about other sources of error (e.g. surface availability) can also be included. Although there are a number of examples of applying a Bayesian hierarchical approach to line transect data (Eguchi & Gerrodoette, 2009; Moore & Barlow, 2011; Conn, Johnson & Laake, 2012) this framework is still ripe for further development. For example, applications to line transect data of cetaceans have generally used single team shipboard data where detectability on the trackline (i.e. *g(0)*) cannot be estimated directly (Moore & Barlow, 2011; Pardo et al., 2015; Pavanato et al., 2017). There have been fewer attempts to develop a framework that can accommodate two team survey data (but see Conn, Laake & Johnson, 2012 for an example). In addition, recent examples of estimating SDMs in a Bayesian framework have used generalized linear models (GLMs) to parameterize the habitat function (Conn, Laake & Johnson, 2012; Pardo et al., 2015; Goyert et al., 2016). Thus, they lack the flexibility that GAMs provide in the two-step method.

In this paper we present a Bayesian hierarchical species distribution model (BHSDM). Our main goal was to develop a comprehensive framework that incorporates the multiple components that influence detection, flexibility in the habitat function and flexible distributions that can accommodate overdispersion and excessive zeros. A secondary goal was to verify that an appropriate level of uncertainty is propagated into final density estimates. We compare the BHSDM with a CSDM analysis using a simulation study and a case study with line transect data of fin whales (*Balaenoptera physalus*) off the east coast of the U.S.

Materials & Methods

Data Collection

Line transect data were collected as part of the Atlantic Marine Assessment Program for Protected Species (AMAPPS) conducted by the Northeast Fisheries Science Center (NEFSC) and the Southeast Fisheries Science Center (SEFSC). The study area ranged from Halifax, Nova Scotia, Canada to the southern tip of Florida and from the coastline to slightly beyond the US exclusive economic zone covering approximately 1,193,320 km^2 (Fig. 1). A total of 16 AMAPPS surveys were conducted using both shipboard and aerial platforms from July 2010 to August 2013 covering approximately 104,000 km of line transect survey effort (Table 1). Shipboard surveys were primarily conducted during summer months in offshore waters and aerial surveys were conducted throughout the year primarily in coastal waters. Each platform included two independent observer teams.

We divided the study site into 10 x10 km oblique Mercator grid cells and into 8-day temporal time periods. For each spatial-temporal cell we calculated the amount of on-effort trackline, number of sightings and obtained the corresponding values of a suite of static physiographic variables and dynamic environmental variables (Table S1). More details on the methods to collect and process the line transect and environmental data are found in Palka et al. (2017).

Model Overview

A general form of the SDM model for a given unit of a study area can be written as

$$E(n_i) = \hat{p}_i A_i \exp\left(f(x)\right), \qquad \text{eqn 1}$$

where $E(n_i)$ is the expected number of sightings in unit i, $\hat{p}_i$ is a distance integrated probability of detection within the search area of unit i, A_i is an offset term for the amount of search effort and $f(x)$ is a user-defined habitat function that relates habitat covariates to the true density of animals in unit *i*. This model can be fit within a frequentist framework using a two step process (Miller et al., 2013).

We take a hierarchical approach to modeling the spatial density of animals. Our modeling framework consists of a number of subcomponent models that include 1) a detection function model based on distance sampling, 2) a group size model to model the average group size (i.e. number of individuals within a group), 3) an informed prior for surface availability based on information on the diving behavior of tagged fin whales and 4) an underlying habitat model. Below we outline the development of each subcomponent and its implementation in a Bayesian framework.

Detection Function

To estimate surface detectability we used information from the double platform survey method. Information collected from this survey design allowed us to apply mark-recapture distance sampling (MRDS) methods (Laake & Borchers, 2004). To model the sightings data from the dual observers we adopt the formulation for point independence outlined by Laake and Borchers (2004). This estimator combines a mark-recapture analysis with conventional distance sampling to estimate detection probability such that detection on the trackline (i.e. *g(0)*) can be estimated directly, and therefore, is not assumed to be 1. The estimator is

$$\hat{p}_{it} = g(0,\boldsymbol{Z_{it}}) * \frac{\int_0^W g(y,\boldsymbol{Z_{it}})dy}{W} \quad , \qquad \text{eqn 2}$$

where $\hat{g}(0, \boldsymbol{Z_{it}})$ represents the estimate of detection probability on the trackline and is estimated from the mark-recapture data; $\hat{g}(y, \boldsymbol{Z_{it}})$ represents the detection function at distance *y* and is estimated from the distance data; $\boldsymbol{Z_{it}}$ is a matrix of detection covariates that influence surface detectability in grid cell *i* at time *t* and *W* is the truncation distance. To model the likelihood for the distance data, we considered half-normal and hazard rate detection functions. For the mark-recapture component of the data we adopted the approach outlined by Laake and Borchers (2004). Specifically, we modeled the binary outcome of whether or not an observer successfully detected an animal group that was present at distance *y* as the outcome of a Bernoulli trial. Further details of analyzing the double platform line transect data using the MRDS method are provided in Appendix S1.

For the aerial surveys, the secondary team was situated toward the back of the plane but had an obstructed view of the trackline complicating a direct implementation of the MRDS approach. Therefore, we estimated an average *g(0)* for both the NEFSC and SEFSC aerial surveys independently where we treated the front team as a single platform and estimated *g(0)* using a trial configuration (i.e., using detections by the rear observers as "trials" for the front observers). The estimated *g(0)* was 0.50 (CV=0.17) and 0.90 (CV=0.09) for NEFSC and SEFSC, respectively. We used these estimates to develop informative priors in the BHSDM. Information on estimating *g(0)* and applying it to the aerial data is provided in Appendix S1.

For each platform and survey, the best detection function was determined through a stand-alone MRDS analysis using the program Distance and fitting both half-normal and hazard rate likelihoods (see Palka et al., 2017). Because sample sizes were low for fin whales, we pooled data from several other large whale species to estimate survey specific detection functions. Models were compared using AIC and the top model for each survey was included in the Bayesian framework. A description of top model used for each platform and survey combination is provided in Appendix S2.

Surface Availability

Because most marine animals spend some amount of time below the surface there is a need to also correct for surface availability ($\hat{a}$) (Laake et al. 1997; Forcada et al. 2004). Species-specific surface availability and the corresponding standard error by platform, was taken from Palka et al. (2017) who adopted the method of Laake et al. (1997). This method was based on the probability of an animal being detectable at the surface during a survey, and took into consideration the species diving and aggregation behaviors, in addition to the amount of time the observer had to analyze any spot of water from each of the survey platforms. This correction tended to be larger for aerial surveys than for shipboard surveys, and larger for long diving species than for short diving species. The estimate for fin whales for aerial surveys was 0.37 (CV=0.34). This information was then used to develop an informative prior. Combining surface detectability and surface availability, our final correction for detection probability in each grid cell i at time t can be written as:

$$\hat{P}_{it} = \hat{p}_{it} * \hat{a}, \quad \text{eqn 3}$$

Habitat Function

We take a generalized additive modeling (GAM) approach to parameterize the habitat function. Because the smooth terms of a GAM have a common multivariate Gaussian (MVN) form they can be estimated relatively easily with standard Markov Chain Monte Carlo (MCMC) techniques (Hastie & Tibshirani, 1990; Wood, 2016). The basic GAM formula can be written as

$$f(x) = \sum_{j=1}^{K} \beta_j b_j(x), \quad \text{eqn 4}$$

where $b_j(x)$ are spline like basis functions and β_j are parameters to be estimated. The scalar K is usually chosen by the user to be large enough to allow appropriate amount of flexibility in $f(x)$. To avoid overfitting quadratic penalty terms are included which take the form

$$\sum_j \gamma_j \beta^T S_j \beta,$$

where S_j are matrices of known coefficients and γ_j are smoothing parameters to be estimated. The precision matrix of these distributions can be derived such that

$$\beta \sim \text{MVN}(0, \sum_j \gamma_j S_j),$$

where the penalty terms are given a vague, gamma prior such as

$$\gamma_j \sim \text{Gamma}(0.05, 0.005),$$

The terms can be estimated efficiently using Gibbs sampling with conjugate priors.

To calculate the precision matrices we used the jagam function in the R package mgcv (Wood, 2016). This function allows the user to specify a number of different smooths (cubic splines, tensor products, etc.) and provides the basic code and input of a JAGS model. In addition, it centers the smooths to facilitate faster convergence.

We can now combine the habitat function with detection probability and an offset from the amount of search effort in each grid cell to estimate the total number of animal groups per grid cell using eqn (1). The offset term was calculated by dividing the area searched within grid cell I by the total area of the grid cell i ($Area_i$) such that $A_i = 2*W*L_i/Area_i$, where L_i is the length of on-effort line transect.

Likelihood

Line transect data tend to be noisy so it is common when constructing SDMs to use a likelihood model that can accommodate outliers. We implemented a Tweedie distribution which has been shown to provide a good fit to cetacean data (Miller at al. 2013; Roberts et al., 2016). The Tweedie distribution is a three parameter family of distributions that can take the form of more commonly used distributions such as the normal, Poisson and gamma. If the power parameter is in the range $1<p<2$ than the distribution can also be referred to as the compound Poisson-gamma (CPG). Because the Tweedie random variables are a sum of G gamma variables where G is Poisson distributed (Jørgensen, 1987), it can be expressed in terms of a Poisson and a gamma distribution such that

$$G \sim \text{Poisson}(\lambda_p)$$

$$M \sim \text{Gamma}(\alpha, \beta)$$

where

$$\lambda_g = \frac{\alpha}{\beta} \qquad \text{eqn 5}$$

and

$$x = \begin{cases} \sum_{i=1}^{G} M, & G > 0 \\ 0 & G = 0 \end{cases} \qquad \text{eqn 6}$$

the expectation is then $E(x)=\lambda_p\lambda_g$. Lauderdale (2012) shows that under a specific parameterization, the coefficients of the regression model can be estimated by estimating both the Poisson and gamma components separately. Specifically, this parameterization can be written as

$$\lambda_P = e^{\frac{X(\beta - \phi)}{2}} \qquad \text{eqn 7}$$

$$\lambda_g = e^{\frac{X(\beta + \phi)}{2}} \qquad \text{eqn 8}$$

where **X** is a matrix of covariate values, **β** is a vector of regression coefficients and **ϕ** is a vector of coefficients that control the extent to which the regression coefficients vary between the Poisson component and gamma component of the compound distribution (Lauderdale, 2012).

Group Size

To model group size we use a zero truncated Poisson such that the group size of each sighting is modeled as

$$(s_k\text{-}1)\sim\text{Pois}(\lambda_s),$$

where s_k is the k_{th} observation of group size and λ_s+1 represents the average group size. This approach assumes that group size is unrelated to detection probability. This assumption is supported by our analysis of fin whale sightings data which did not indicate a strong influence of group size on detection probability.

Density and abundance estimation

To estimate the density within a grid cell we multiply the estimate of group size with the estimate of group density such that

$$N_{it} = D_{it} * (\lambda_s + 1), \qquad \text{eqn 9}$$

where D_{it} is the predicted number of groups in grid cell i at time t estimated from the GAM model, $\lambda_s + 1$ represent the average group size and N_{it} is the predicted number of individuals in grid cell i at time t. To estimate total abundance of individuals within the study area we sum N_{it} over all grid cells within the study area.

Model Fitting

We fit the BHSDM outlined above using MCMC sampling implemented with the JAGS software (Plummer 2003). Vague prior distributions were used for all parameters with the exception of *g(0)* for the aerial surveys and $\hat{a}$ where we used estimates and associated CVs to develop informative beta prior distributions using the approach of Pardo et al. (2015). We included a burnin of 20000 samples and two chains of 50000 with a thinning rate of 50. Convergence was assessed by examining traceplots and calculating Gelman-Rubin diagnostics.

Simulation study

To quantify differences in precision and statistical coverage probability between the BHSDM and CSDM method we used a simulation study. We simulated spatial variation in abundance over 300 hypothetical grid cells that were all 100 km^2 in area. Each grid cell was assigned a covariate value and the associated density of animals was generated as a quadratic function of its covariate value. We simulated variation in search effort and detection including both surface detectability and surface availability. We applied both the CSDM method and the BHSDM method to each of 1000 independently simulated datasets. For each simulation, we estimated population size by summing up the estimated number of animals in each grid cell. We estimated statistical interval coverage by determining whether or not the true population size fell within the 95% confidence intervals of the CSDM method or the 95% credible region of the posterior estimate of the BHSDM method. In addition, we calculated percent relative bias for each simulation and method as $(\hat{N}_{ij}\text{-}N_i)/ N_i\text{*}100$ where N_i is true abundance for simulation i and

$\hat{N}_{ij}$ is the estimate of population size for simulation i and method j. A more detailed explanation of the simulations is provided in Appendix S2.

Case Study with Fin Whales

We tested our model on a four year dataset of fin whale sightings collected during the AMAPPS surveys. Similar to our simulation study we also compared model predictions to predictions from a CSDM using a version of the two-step method. Details for fitting the CSDM to the fin whale dataset are explained elsewhere (see Palka et al., 2017) but we briefly review the process here. In the first step, we calculated densities of fin whales from the stand-alone MRDS analysis in program Distance. We adjusted estimates from the aerial surveys by dividing by an estimate of surface availability. In the second step, we fit GAMs to these effort-corrected estimates of density using the mgcv software (Wood, 2011) using thin plate regression splines and restricted error maximum likelihood (REML) to estimate parameters. To account for overdispersion, a Tweedie distribution was assumed. A multi-stage process was used to determine the best set of covariates and the best structure of the smooth terms (see Palka et al. 2017 Appendix I, chapter 3 for results). We then fit the same model using the BHSDM. Our interest was in keeping as many components of the model structure consistent so we could focus on comparing output across modeling frameworks.

Results

Simulation Study

Simulations demonstrated higher uncertainty and higher statistical interval coverage using the BHSDM method as compared to the CSDM method. We found approximately 94% coverage probability for the BHSDM compared to 84% coverage probability for the CSDM (Fig. 2). The average CV of $\hat{N}$ calculated from the BHSDM method was approximately 20% higher than the average CV calculated from the CSDM method demonstrating the additional uncertainty that is propagated in a BHSDM framework. Overall bias was low and positive for both methods but slightly more positive for the CSDM method (Fig. 2).

Case Study

A comparison of the resulting detection functions between the stand alone MRDS and the BHSDM showed detection probabilities were similar (Table 2). Estimates from the distance sampling component were similar among all framework although estimates from the BHSDM were consistently higher. Similarly, estimates of *g(0)* for the shipboard surveys were also similar but consistently higher in the BHSDM framework. The posterior estimate of mean group size was 1.4 (CV=0.14) indicating relatively small group sizes. Most observed group sizes were less than 2 animals with 4 % greater than 3 animals.

The top model for the habitat function included latitude, chlorophyll, sea surface temperature and distance to 125 m isobath as covariates (see Palka et al,. 2017 Appendix I, chapter 3**)** and were subsequently used in the BHSDM. Results from fitting the BHSDM to the observed sightings data showed good agreement between predicted and observed number of groups per grid cell although there was some tendency of the model to under predict as the number of sightings increased (Fig. S1). In comparison to the CSDM, density estimates for the grid cells during the summer time period were similar between the two frameworks (Fig. 3).

Overall abundance estimate for the entire study area was 4 % lower with lower precision for the BHSDM (4526,CV=0.26) compared to the CSDM (4718,CV=0.13) .

In contrast to density estimates, sampling uncertainty of the grid cell density estimates varied substantially between the two frameworks (Fig. 4). The median CV's were 0.28 and 0.98 for the CSDM and the BHSDM, respectively. The distribution of CV's from the BHSDM were highly skewed. However, the highest CV's from BHSDM were associated with low density estimates (Fig. 4). Ignoring grid cells with low density estimates (density <0.0001 animals/km^2), the median CV's were 0.45 and 0.15 for the BHSDM and CSDM, respectively.

The predicted spatial average seasonal density distribution patterns were largely similar for the two frameworks (Fig. 5a and Fig. 5b). Both modeling frameworks indicated that the Gulf of Maine and the shelf break where areas of relatively high density. There were some minor differences south of the Gulf of Maine where the CSDM indicated slightly higher densities than the BHSDM. The overall pattern of uncertainty was similar for the two frameworks where uncertainty was lowest in areas with the highest density and vice versa for the areas of highest uncertainty (Fig. 5c and Fig. 5d).

Discussion

The use of SDMs to predict the abundance and distribution of animals in time and space is increasingly becoming a cornerstone in the conservation and management of cetacean populations (Gregr et al., 2012). Techniques for fitting SDMs have grown rapidly and although there have been a number of attempts to compare methods in terms of model fit and prediction (Elith & Graham, 2009; Oppel et al., 2012), few studies have focused specifically on uncertainty. Yet, properly accounting for uncertainty is crucial to informing good management decisions

(Ludwig, Hilborn, & Walters 1993). Our focus was on further developing the one-step BHSDM and comparing its performance to the more standard CSDM with the explicit goal of highlighting differences in uncertainty between the two methods.

Two-step approaches to SDMs have the advantage of being able to use all the built-in options available in different software packages such as Distance to model detection and mgcv to model habitat relationships (Miller et al. 2013). In contrast, Bayesian methods have been more limited in these options. We have taken steps to expand the BHSDM framework such that it is more compatible with the two-step approach and similar models can be compared. For example, we adopted the MRDS approach available in Distance to model two team data while also including a hazard rate option. Previous Bayesian applications to line transect data of cetaceans have generally been applied to single shipboard team data using a half-normal detection function (Moore & Barlow, 2011; Pavanato et al., 2017). We have also included nonparametric GAMs which allow for flexible, data-driven relationships between habitat and density. Other approaches to BHSDM have included quadratic terms in a GLM framework (Pardo et al., 2015; Goyert et al., 2016) to capture nonlinear relationships, but this approach is still parametric in form and limited in flexibility. Finally, we implemented a Tweedie distribution within this framework. Because the Tweedie is not a built in distribution in most Bayesian software packages, we adopted the CPG approach of Lecomte et al. (2013). This approach is a limited version of the Tweedie where the power parameter is constrained between 1 and 2; however, it is not uncommon in studies using the two-step approach with a Tweedie distribution to restrict the power parameter to be within this range (Williams et al., 2011; Cañadasa et al., 2018). Together these features provide more options for users when applying a BHSDM to line transect data.

In a review of marine SDMs, Robinson et al. (2017) notes that uncertainty is rarely assessed rigorously. We explicitly addressed this issue with simulations. We found the BHSDM method not only results in higher uncertainty but greater statistical coverage than the CSDM method. When interpreting the magnitude of the difference between the two methods it is important to take into account the simplicity of the simulations. For instance, we assumed a low amount of scientific uncertainty around the estimate of surface availability (as represented by a relatively narrow distribution). In addition, our estimates of surface detectability were relatively precise as we used simulated data with no detection covariates. Finally, we did not include uncertainty in group size. Whether using the two-step method or one-step method it is important to consider all the sources of uncertainty and their influence on final predictions. For example, estimates of surface availability from tag data are generally limited to a few individuals and may include high uncertainty. As a corollary, these results also point to the value of trying to derive more precise estimates of detectability. Recent advances in sampling technology such as passive acoustic technology (Marques et al., 2012) and aerial drones (Brack et al., 2018) may be greatly beneficial in estimating both more accurate and more precise measurements of surface availability and in turn could greatly reduce uncertainty in final estimates.

Our simulations also demonstrated that overall bias of the BHSDM method is relatively small and slightly positive. When using Bayesian methods, bias will partly be a function of how one calculates point estimates using posterior samples. For our simulations we used the posterior mean, but the posterior median or mode could also be used. In a study of harbor seals, Ver Hoef and Jansen (2007) showed that in cases where the posterior distributions are skewed, the posterior median estimate can be biased low and the posterior mean estimate can be biased high. They suggest use of the linex loss function to achieve more accurate estimates (Varian, 1975).

We did not take that approach here, but it is worth considering if small amounts of bias are a concern.

When applied to field data, the BHSDM produced density estimates and distribution patterns that were similar to the CSDM albeit not exact. Although we attempted to keep the two frameworks as similar as possible, there were some structural differences that may have influenced density estimates. For example, the CSDM used a Horvitz-Thompson like estimator to estimate observed densities whereas the BHSDM models used the observed sighting of groups in each grid cell directly and used a mean estimate of group size to calculate total density of animals. In addition, the estimates of detection from the BHSDM tended to be higher than estimates from a stand-alone MRDS analysis which might have translated into slightly lower density estimates. Overall, the spatial distribution and abundance from both methods were comparable to a previous study by Roberts et al. (2016) for the same general area.

Estimates of precision differed substantially between the two methods. This result was anticipated as the hierarchical structure of the BHSDM framework propagates more uncertainty from the other components of the model. Several factors contributed to uncertainty including uncertainty in average group size, surface detectability and surface availability. Estimates of precision from most of these components were relatively high with CVs ranging from 0.06 to 0.26. Our estimate of surface availability had the lowest precision with a CV of 0.34 and likely contributed the most to differences in CVs among density estimates.

Proper consideration of uncertainty is crucial to effective management of natural resources (Ludwig, Hilborn & Walters, 1993). A number of studies have shown how failure to consider uncertainty can result in poor management decisions (Regan et al., 2005; Artelle et al., 2013). For example, in population viability analysis, ignoring error in initial population size may result in misleading estimates of population persistence (McLoughlin & Messier, 2004). In the management of cetacean populations, overly precise estimates of abundance can have direct consequences on the determination of potential biological removal and may result in a lack of management action when action should be taken (Taylor et al., 2000). Using a BHSDM, Gerrodette & Eguchi (2012) demonstrated how a more complete consideration of uncertainty of spatial distribution can result in a more cautionary approach to the design of a marine reserve that may ultimately be more effective for conservation. Taken together, these studies suggest that modeling tools used to inform management decisions must prioritize a full assessment of uncertainty to avoid undesirable outcomes.

Conclusions

Rigorously quantifying uncertainty is a challenging but important goal. Recently, Bravington, Miller & Hedley (2018) developed alternative methods for propagating uncertainty from estimation of detectability into final density estimates within two-stage line transect SDMs. Their approach appears promising, and we expect it will likely become common practice for those conducting two-stage SDM modeling with line transect data. Nevertheless, one stage hierarchical models may be the only way to resolve certain detection processes – for instance, in cases where detectability is a function of individual covariates such as group size, or when species misclassification occurs (e.g. Conn et al. 2012, Conn et al. 2013). Thus, we expect to see continued, parallel development of hierarchical models for line transect data together with two-stage SDMs.

Acknowledgements

The authors would like to thank the crews of the NOAA ships Henry B. Bigelow and Gordon Gunter and the NOAA Twin Otter aircrafts, all of the shipboard and aerial observers and all the people involved in the data collection and analyses.

References

Artelle, K. A., Anderson, S. C., Cooper, A. B., Paquet, P. C., Reynolds, J. D., & Darimont, C. T. (2013). Confronting uncertainty in wildlife management: Performance of grizzly bear management *PLoS ONE*, **8**, e78041.

Bravington, M. V., Miller, D. L. & Hedley, S. L. (2018). Reliable variance propagation for spatial density surface models. *Biometrics*.

Cañadasa, A., Aguilar de Sotoc, N., Aissie, M., Arcangelif, A., Azzoling, M., B-Nagyh, A., Bearzii, G., Campanaj, I., Chicotek, C., Cottel, C., Crostin, R., Davidt, L., Di Nataleo, A., Fortunaf,C., Frantzisq, A., Garciar, P., Gazok, M., Gutierrez-Xarxas, R., Holcerp, D., Larant, S., Laurianof,G., Lewisv, T. Moulinsw, A., Mussix, B., Notarbartolo di Sciarau, G., Panigadau, S., Pastory, X., Politiu, E., Pulcinif,M., Raga, J. A., Rendell, L., Rossow, M., Tepsich, P., Tomás, J., Tringali, M., and Roger, Th. (2018). The challenge of habitat modelling for threatened low density species using heterogeneous data: The case of Cuvier's beaked whales in the Mediterranean. *Ecological Indicators*, **85**, 128-136.

Conn, P.B., Laake, J.L. & Johnson, D.S. (2012) A hierarchical modeling framework for multiple observer transect surveys. *PLoS One*, **7**, e42294.

Conn, P.B., McClintock, B.T., Cameron, M., Johnson, D.S., Moreland, E.E. & Bovent, P.L. 2013. Accommodating species identification errors in transect surveys. *Ecology* 94:2607-2618.

Eguchi, T. & Gerrodette, T. (2009). A Bayesian approach to line-transect analysis for estimating abundance. *Ecological Modelling*, **220**, 1620-1630.

Elith, J., & Graham, C. H. (2009). Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, **32**, 66-77.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677-697.

Forcada, J., Gazo, M., Aguilar, A., Gonzalvo, J. & Fernández-Contreras, M. (2004) Bottlenose dolphin abundance in the NW Mediterranean: addressing heterogeneity in distribution. *Marine Ecology Progress Series*, **275**, 275-287.

Forney, K.A., Ferguson, M.C., Becker, E.A., Fiedler, P.C., Redfern, J.V., Barlow, J., Vilchis, I.L. & Balance, L.T. (2012) Habitat-based spatial models of cetacean density in the eastern Pacific Ocean. *Endangered Species Research*, **16**,113-133.

Gerrodette, T. and Eguchi, T. (2011). Precautionary design of a marine protected area based on a habitat model. *Endangered Species Research*, **15**, 159–166.

Goyert, H.F., Gardner, B., Sollmann, R.,Veit, R.R., Gilbert, A.T., Connelly, E.E. & Williams, K.A. (2016) Predicting the offshore distribution and abundance of marine birds with a hierarchical community distance sampling model. *Ecological Applications*, **26**, 1797-1815.

Gregr, E. J., Baumgartner, M. F., Laidre, K. L., & Palacios, D. M. (2013). Marine mammal habitat models come of age: the emergence of ecological and management relevance. *Endangered Species Research*, **22**, 205–212.

Hastie, T. & Tibshirani, R. (1990) Generalized Additive Models. Chapman & Hall.

Hedley, S.L. & Buckland, S.T. (2004) Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics*, **9**, 181–199.

Jorgensen, B. (1987) Exponential dispersion models. *Journal of the Royal Statistical Society: Series B, Statistical Methodology*, **49**, 127–162.

Laake, J.L., Calambokidis, J., Osmek, S.D. & Rugh, D.J. (1997) Probability of detecting harbor porpoise from aerial surveys: Estimating g(0). *Journal of Wildlife Management*, **61**, 63-75.

Laake, J.L. & Borchers, D.L. (2004) Methods for incomplete detection at distance zero. In: Buckland, S., Anderson, D., Burnham, K., Laake, J. & Borchers, D. et al. editors, Advanced Distance Sampling, Oxford, U.K. Oxford University press 108-189.

Langrock, R., Borchers, D.L. & Skaug, H.S. (2013) Markov-modulated nonhomogeneous Poisson processes for modeling detections in surveys of marine mammal abundance. *Journal of the American Statistical Society* **108**, 840-851.

Lauderdale, B. (2012) Compound Poisson–gamma regression models for dollar outcomes that are sometimes zero. *Political Analysis*, **20**, 387–399

Lecomte, J.B., Benoıˆt, H.P., Ancelet, S., Etienne, M.P., Bel, L. & Parent, E. (2013) Compound Poisson-gamma vs. delta-gamma to handle zero-inflated continuous data under a variable sampling volume. *Methods in Ecology and Evolution*, **4**, 1159-1166.

Ludwig, D., Hilborn, R., & Walters, C. (1993). Uncertainty, resource exploitation, and conservation: lessons from history. *Science*, **260**, 17−36.

Mcloughlin, P. D., and Messier, F. (2004) Relative contributions of sampling error in initial population size and vital rates to outcomes of population viability analysis. *Conservation Biology*, **18**, 1665–1669.

Miller, D. L., Burt, L. M., Rexstad, E. A., & Thomas, L. (2013) Spatial models for distance sampling data: recent developments and future directions. *Methods in Ecology and Evolution*, **4**, 1001-1010.

Moore, J.E. & Barlow, J.P. (2011) Bayesian state-space model of fin whale abundance trends from a 1991-2008 time series of line-transect surveys in the California Current. *Journal of Applied Ecology*, **48**,1195-1205.

Oppel, S., Meirinho, A., Ramírez, I., Gardner, B., O'Connell, A. F., Miller, P. I., & Louzao, M. (2012). Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. *Biological Conservation*, **156**, 94-104.

Palka, D.L., Chavez-Rosales, S., Josephson, E., Cholewiak, D., Haas, H.L., Garrison, L., Jones, M., Sigourney, D., Waring, G.(retired), Jech, M., Broughton, E., Soldevilla, M., Davis, G., DeAngelis, A., Sasso, C.R., Winton, M.V., Smolowitz, R.J., Fay, G., LaBrecque, E., Leiness, J.B., Dettloff, Warden, M., Murray, K. & Orphanides, C. (2017) Atlantic Marine Assessment Program for Protected Species: 2010- 2014. US Dept. of the Interior, Bureau of Ocean Energy Management, Atlantic OCS Region, Washington, DC. OCS Study BOEM 2017-071. 211 pp

Pardo, M.A., Gerrodette, T., Beier, E., Gendron, D., Forney, K.A., Chivers, S., Barlow, J. & Palacios, D.M. (2015) Inferring cetacean population densities from absolute dynamic topography of the ocean in a hierarchical Bayesian framework. *PLoS ONE*, **10**,UNSP e0120727.

Pavanato, H.J., Wedekin, L.L., Guilherme-Silveira, F.R., Engel, M.H. & Kinas, P.G. (2017). Estimating humpback whale abundance using hierarchical distance sampling. *Ecological Modelling*, **358**, 10-18.

Plummer, M. (2003) JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. Page 125 in Proceedings of the 3rd international workshop on distributed statistical computing. Technische Universit at Wien, Wien, Austria.

Regan, H. M., Ben-Haim, Y., Langford, B., Wilson, W. G., Lundberg, P., Andelman, S. J., & Burgman, M. A. (2005). Robust decision-making under severe uncertainty for conservation management. *Ecological Applications,* **15**, 1471–1477.

Roberts, J.J., Best, B.D., Mannocci, L., Fujioka, E., Halpin, P.N., Palka, D.L., Garrison, L.P., Mullin, K.D., Cole, T.V.N., Khan, C.B., McLellan, W.A., Pabst, D.A. & Lockhart, G.G. (2016) Habitat-based cetacean density models for the US Atlantic and Gulf of Mexico. *Scientific Reports*, **6**, 22615.

Robinson, N.M., Nelson, W.A., Costello, M.J., Sutherland, J.E. & Lundquist, C.J. (2017) A systematic review of marine-based species distribution models (SDMs) with recommendations for best practice. *Frontiers in Marine Science*, **4**, 421.

Royle J. A., & Dorazio, R. M. (2008). Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities. Elsevier, London

Royle, J. A., Dawson, D. K., & Bates S. (2004). Modeling abundance effects in distance sampling. *Ecology*, **85**, 1591−1597.

Taylor, B. L., Wade, P. R., De Master, D. P., & Barlow, J. (2000). Incorporating uncertainty into management models for marine mammals. *Conservation Biology*, **14**, 1243-1252.

Ver Hoef, J.M. & Jansen, J.K. (2007) Space–time zero-inflated count models of harbor seals. *Environmetrics*, **18**, 697-712.

Williams, R., Hedley, S.L., Branch, T.A., Bravington, M.V., Zerbini, A.N. & Findlay, K.P. (2011) Chilean blue whales as a case study to illustrate methods to estimate abundance and evaluate conservation status of rare species. *Conservation Biology*, **25**, 526–535.

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric Generalized Linear Models. *Journal of the Royal Statistical Society: Series B, Statistical Methodology*, **73**, 3–36.

Wood, S.N. (2016) Just Another Gibbs Additive Modeller: Interfacing JAGS and mgcv. *Journal of Statistical Software*, **75**, 1-15.

Table 1(on next page)

Summary of effort by season and platform

Table 1: Summary of effort by season and platform.

Platform	Effort (km)			
	Spring	Summer	Fall	Winter
NE Shipboard	0	8,146	0	0
NE Aerial	7,502	10,468	11,038	3,573
SE Shipboard	0	8,537	2,093	0
SE Aerial	17,978	16,835	11,818	6,007

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27424v1 | CC BY 4.0 Open Access | rec: 11 Dec 2018, publ: 11 Dec 2018

Table 2 (on next page)

Comparison of posterior estimates for detection functions

Comparison of posterior estimates for the detection function estimated from the Bayesian hierarchical species distribution model (BHSDM) to estimates of the detection function from the mark-recapture distance sampling (MRDS) analysis using program Distance. Results are shown for each survey along with the truncation distance (W) used in each analysis. Estimates of detection from the distance sampling component (P_D) and g(0) from the mark-recapture component are shown with coefficients of variation in parentheses.

Table 2: Comparison of posterior estimates for detection functions

Comparison of posterior estimates for the detection function estimated from the Bayesian hierarchical species distribution model (BHSDM) to estimates of the detection function from the mark-recapture distance sampling (MRDS) analysis using program Distance. Results are shown for each survey along with the truncation distance (W) used in each analysis. Estimates of detection from the distance sampling component (P_D) and g(0) from the mark-recapture component are shown with coefficients of variation in parentheses.

Model	Survey	W (km)	P_D	g(0)
BHSDM	NE Ship	4	0.31 (0.26)	0.79 (0.06)
MRDS	NE Ship	4	0.28 (0.26)	0.78 (0.08)
BHSDM	SE Ship	8.84	0.53 (0.11)	0.82 (0.08)
MRDS	SE Ship	8.84	0.51 (0.10)	0.77 (0.13)
BHSDM	NE Air	5.24	0.13(0.11)	0.50 (0.17)*
MRDS	NE Air	5.24	0.11 (0.10)	0.50 (0.17)*
BHSDM	SE Air	0.56	0.59 (0.14)	0.90 (0.10)*
MRDS	SE Air	0.56	0.56 (0.12)	0.90 (0.10)*

**Estimates of g(0) for aerial surveys were taken from Palka et al. (2017)*

Figure 1(on next page)

AMAPPS study area

Map of the AMAPPS study area with the shipboard survey track lines in blue and aerial survey track lines in orange for surveys conducted from 2010-2013.

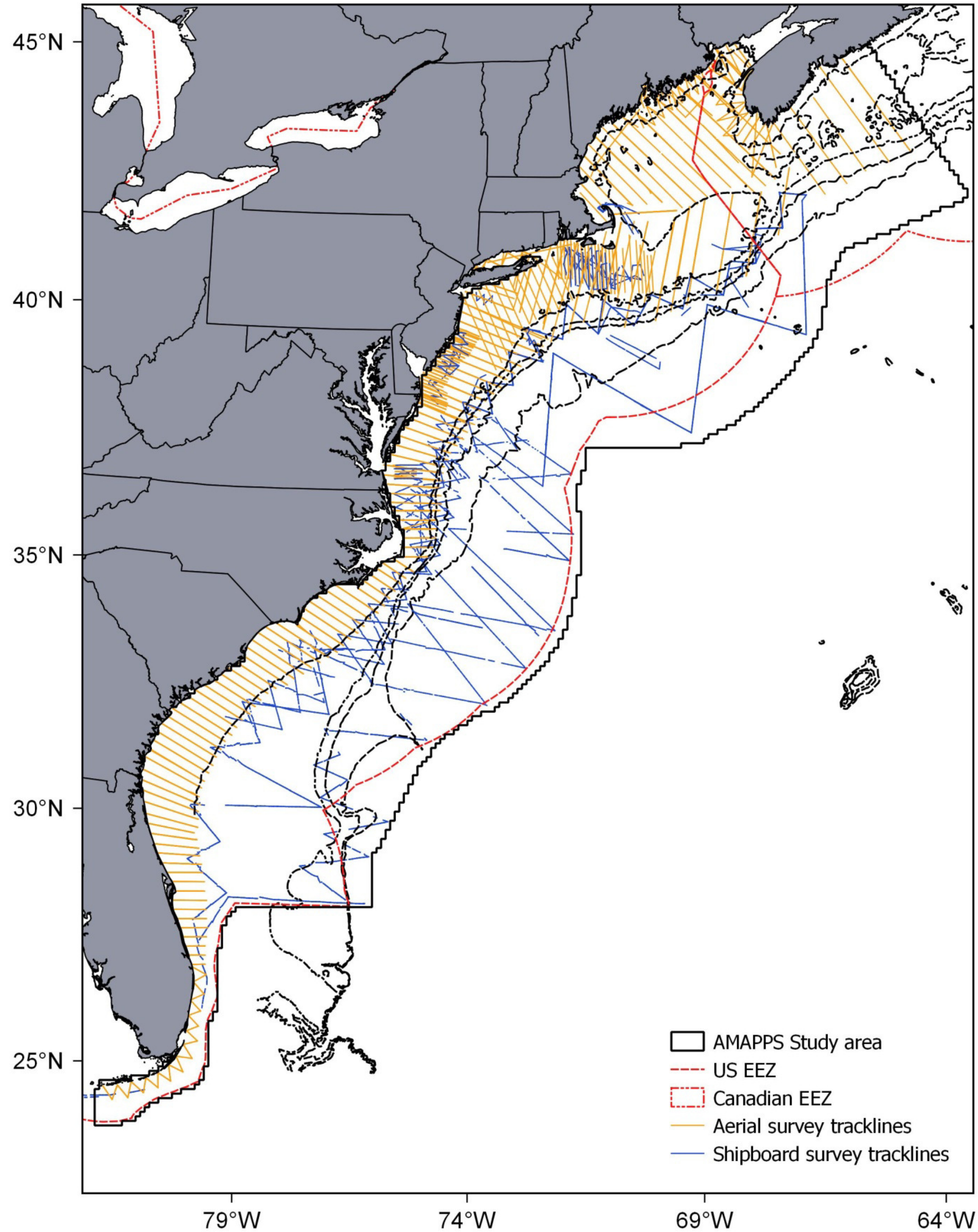

Figure 1: AMAPPS study area

Map of the AMAPPS study area with the shipboard survey track lines in blue and aerial survey track lines in orange for surveys conducted from 2010-2013.

Figure 2(on next page)

Summary of simulations

Summary results comparing coverage probability after applying the BHSDM and CSDM methods to 1000 simulated datasets. Black lines indicate estimates that covered the true value and red lines indicate estimates that did not cover the true value (a & b). Histograms of coefficients of variation (CV) (c & d) and bias in estimates of population size (e & f) are also shown.

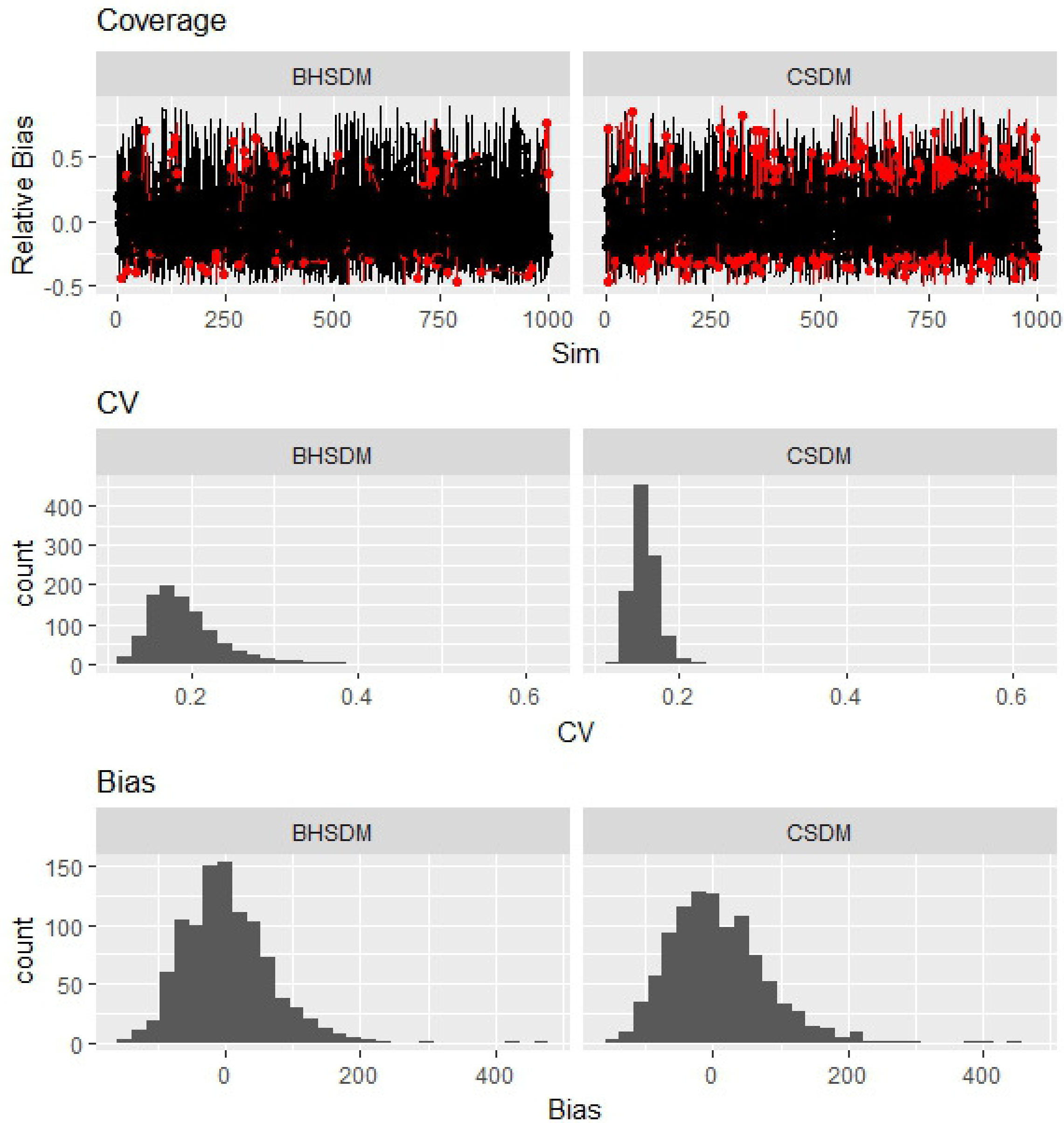

Figure 2: Summary of simulations

Summary results comparing coverage probability after applying the BHSDM and CSDM methods to 1000 simulated datasets. Black lines indicate estimates that covered the true value and red lines indicate estimates that did not cover the true value (a & b). Histograms of coefficients of variation (CV) (c & d) and bias in estimates of population size (e & f) are also shown.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27424v1 | CC BY 4.0 Open Access | rec: 11 Dec 2018, publ: 11 Dec 2018

Figure 3 (on next page)

Comparison of density estimates.

Comparison of density estimates from the Bayesian hierarchical species distribution model (BHSDM) vs density estimates from the Conventional species distribution model (CSDM) framework.

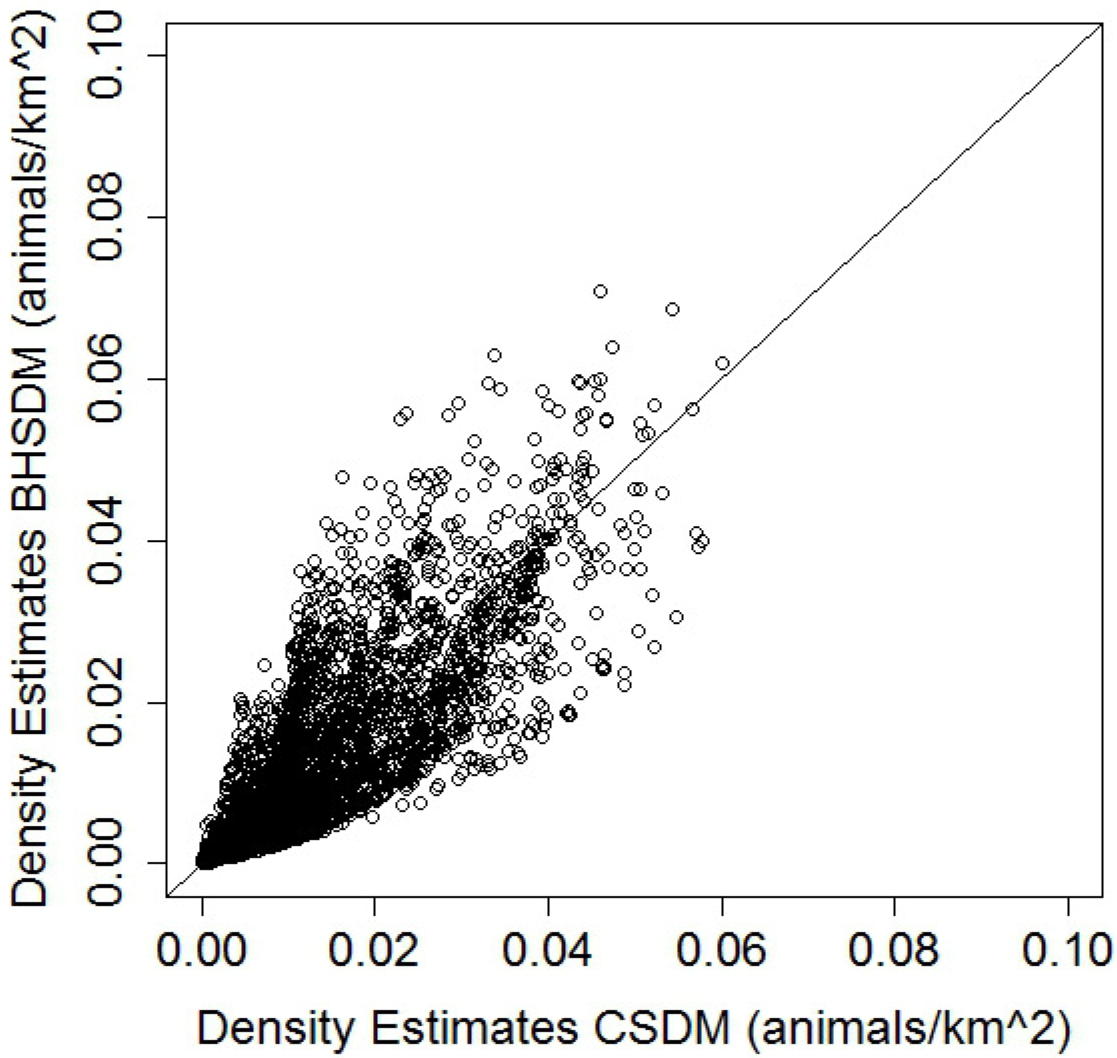

Figure 3: Comparison of density estimates.

Comparison of density estimates from the Bayesian hierarchical species distribution model (BHSDM) vs density estimates from the Conventional species distribution model (CSDM) framework.

Figure 4(on next page)

Comparison of coefficients of variation (CVs) for density estimates

Comparison of coefficients of variation (CVs) for density estimates from the Bayesian hierarchical species distribution model (BHSDM) and conventional species distribution model (CSDM). Inner panel shows a histogram of CVs for all density estimates for both the BHSDM and CSDM with median CVs represented by dashed lines. Outer panel shows a plot of CVs vs density estimates for both the BHSDM and CSDM. Only density estimates greater than 0.00001 animal/km^2 are shown.

Figure 4: Comparison of coefficients of variation (CVs) for density estimates

Comparison of coefficients of variation (CVs) for density estimates from the Bayesian hierarchical species distribution model (BHSDM) and conventional species distribution model (CSDM). Inner panel shows a histogram of CVs for all density estimates for both the BHSDM and CSDM with median CVs represented by dashed lines. Outer panel shows a plot of CVs vs density estimates for both the BHSDM and CSDM. Only density estimates greater than 0.00001 animal/km^2 are shown.

Figure 5 (on next page)

Predicted densities of fin whales in summer.

Predicted densities and abundance estimates with corresponding coefficients of variation (CV) of fin whales in summer from a species distribution model using a) the CSDM framework and b) the BHSDM framework. Coefficients of variation for the density estimates from c) the CSDM framework and d) BHSDM framework are also provided.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27424v1 | CC BY 4.0 Open Access | rec: 11 Dec 2018, publ: 11 Dec 2018

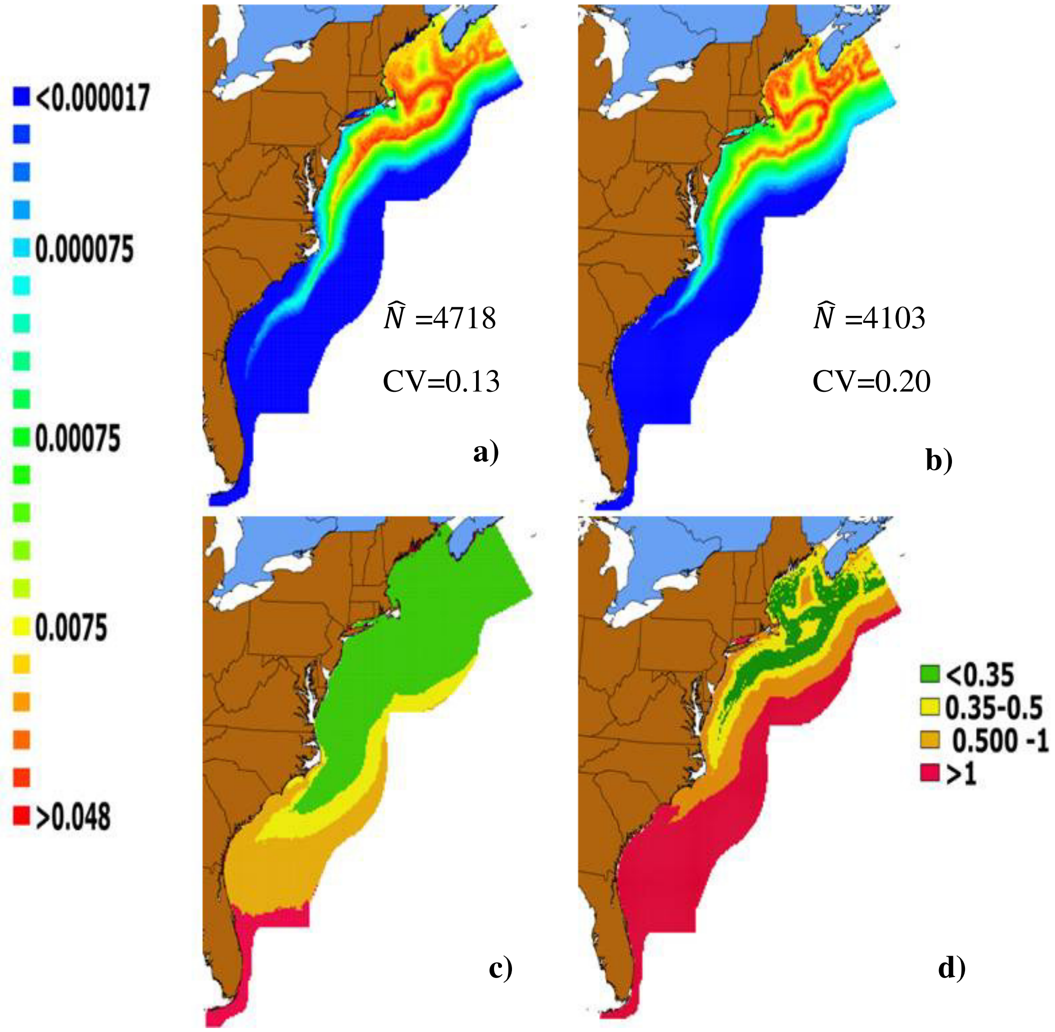

Figure 5: Predicted densities of fin whales in summer.

Predicted densities and abundance estimates ($\widehat{N}$) with corresponding coefficients of variation (CV) of fin whales in summer from a species distribution model using a) the CSDM framework and b) the BHSDM framework. Coefficients of variation for the density estimates from c) the CSDM framework and d) BHSDM framework are also provided.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27424v1 | CC BY 4.0 Open Access | rec: 11 Dec 2018, publ: 11 Dec 2018