

RESEARCH ARTICLE

A GLMM approach for combining multiple relative abundance surfaces

Paul B. Conn¹  | Jay M. Ver Hoef¹  | Brett T. McClintock¹  | Devin S. Johnson²  | Brian Brost¹ 

¹Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA, National Marine Fisheries Service, Seattle, WA, USA

²Pacific Islands Fisheries Science Center, NOAA, National Marine Fisheries Service, Honolulu, HI, USA

Correspondence

Paul B. Conn

Email: paul.conn@noaa.gov

Handling Editor: Sydne Record

Abstract

1. Spatio-temporal maps of organismal density or relative abundance are fundamental to many applications in conservation and ecology. There are often a number of data sources to inform species maps, including citizen science monitoring programs, remote sensing, geolocation data from satellite-tagged animals and formal scientific surveys. In these cases, it may be desirable to come up with a single map integrating all data sources.
2. We introduce a novel two-step method for combining inference about relative abundance maps using multiple data sources. Log-scale relative abundance surfaces are first estimated from individual data sets, and resultant surfaces are then treated as 'data' within a generalized linear mixed model framework with a spatially autocorrelated mean process. We also introduce generalizations allowing data misalignment and predictive processes to decrease computational burden.
3. Using simulation, we show that our approach frequently outperforms other approaches, including basing inference on a single surface, or taking a simple arithmetic mean (although the arithmetic mean performs well when there are few surfaces). We then demonstrate our method using citizen science and satellite-tracking data of Steller sea lions in Alaska. In this case, relative abundance surfaces consisted of an effort-adjusted map developed from platform-of-opportunity sightings and a utilization distribution developed from geolocation records. The resulting combined surface represented a compromise between single-data source predictions.
4. Our approach should be useful for ecologists seeking to reconcile alternative species distribution maps, particularly when individual surfaces are prone to bias, when there is no obvious common currency (e.g. point process), or when computational demands preclude a fully integrated analysis.

KEYWORDS

density surface model, *Eumetopias jubatus*, resource selection function, spatial autocorrelation, species distribution model, Steller sea lion, telemetry

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Published 2022. This article is a U.S. Government work and is in the public domain in the USA. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

1 | INTRODUCTION

Spatio-temporal maps of animal density or relative abundance are fundamental to many applications in conservation and ecology. Ecologists, for instance, frequently study spatially explicit drivers of animal occurrence (Elith & Leathwick, 2009; MacArthur, 1972). Conservation examples include biodiversity monitoring, ecosystem management, and natural resource impact assessments, where spatial distributions inform biodiversity indicators and conservation planning (e.g. through GEOBON; Scholes et al., 2008), spatial harvesting (McCullough, 1996) and reserve designs (Wilson et al., 2005), and determining the number of animals affected by anthropogenic activities (e.g. oil spills, naval exercises; Ver Hoef et al., 2021). Our own interest in relative abundance maps stems from a desire to produce seasonal at-sea density maps of marine mammals to inform decision-making and to calculate 'takes', a requirement of the U.S. Marine Mammal Protection Act.

A variety of data sources have been used to construct models to predict animal density and distribution, including scientific surveys, citizen science observations, and geolocation data from telemetry devices. These data are often analysed in isolation, employing different types of models. For instance, researchers use species distribution models (SDMs) to produce species occurrence maps from presence-only records (Elith & Leathwick, 2009), density surface models (DSMs) to generate animal density maps from count data obtained during transect surveys, or effort-adjusted models to produce relative abundance maps from count data generated from citizen science monitoring programs (Ver Hoef et al., 2021). Similarly, telemetry data can be analysed with resource selection functions (RSFs; Manly et al., 2010) to produce utilization distributions (UDs) that represent the predicted probabilities of space use (often interpreted as a surrogate for density). Alternatively, animal movement models using step selection functions (SSFs; Thurfjell et al., 2014) can be applied to telemetry data; in this case, steady state solutions are often available that approximate predicted space use.

A recent focus in SDM modelling has been to develop integrated models that combine inferences from multiple data types (e.g. citizen science and occupancy models; Dorazio, 2014; Fithian et al., 2015; Pacifici et al., 2017). Such models leverage high quality (but low quantity) scientific survey data with lower quality (but high quantity) citizen science observations to produce relative abundance maps that have better predictive performance than those generated by analysing each data source in isolation. These models have relied on an underlying inhomogeneous Poisson process (a spatial model where animal density changes over space) as the common currency to simultaneously model observations of both data types. However, this approach is not readily extensible to other data types (e.g. telemetry records from satellite tags) because more complicated movement models are often needed to account for the temporal autocorrelation invariably present in animal tracking data (Hooten et al., 2017). This is unfortunate, as utilization distributions produced from animal movement models are often interpreted as relative abundance maps and clearly include information about distribution and density that

would be useful to meld with other available data into unified relative abundance maps.

Although there has been promising recent work that seeks to reconcile macro- and micro-scale parameters in analyses of animal movement (McClintock et al., 2021; Michelot et al. 2019a, 2019b; Wilson et al., 2018), the development of fully integrated models for multiple data types that share the same set of parameters may still be prohibitive to implement. Further, the comparison of abundance surfaces derived from Eulerian (macroscale) and Lagrangian (microscale) perspectives (Turchin, 1998) have been known to disagree. For instance, resource selection functions developed from telemetry data can differ from surfaces developed from survey data (see e.g. Conn et al., 2021; Phillips et al., 2019), presumably because tagged individuals represent a biased subset of the population. Similarly, when citizen scientists base decisions about where to sample on knowledge of the target population it can potentially bias SDM surfaces (this is a version of 'preferential sampling'; Diggle et al., 2010; Conn et al., 2017; Johnston et al., 2020). For these reasons, methods to combine surfaces should ideally include some capacity to account for potential bias in individual surfaces.

We consider methods for producing a single relative abundance prediction map when several potential surfaces are available. Rather than developing fully integrated models with shared parameters, our approach is to use a two-step method where individual data sources are analysed separately to produce spatial surfaces. Individual surfaces (and attendant variance-covariance matrices) are then used as data in a second step where a single relative abundance surface is produced. After defining this modelling framework, we conduct a simulation study to examine best practices for combining surfaces, including cases where all surfaces are assumed to be biased and cases where one surface can safely be assumed to be unbiased (such as one developed from scientific survey data). We then apply our methods to Steller sea lion (SSL; *Eumetopias jubatus*) telemetry and citizen science platform-of-opportunity (POP) data from Alaska, USA. In each case, our interest is in developing a single relative abundance map that provides a more accurate representation of population-level distribution than would be possible by analysing each data set alone.

2 | MATERIALS AND METHODS

2.1 | Notation and basic model structures

Suppose there are n individual relative abundance surfaces that one wishes to combine into a single surface. We assume that each surface has relative abundance values measured on a discrete spatial domain (i.e. pixel), and that they are proportional to absolute abundance (they could variously consist of density, abundance, or relative abundance). Relative abundance values in each pixel s ($s \in 1, 2, \dots, S$) are given by $d_i(s)$ for surface i ($i \in 1, 2, \dots, n$), and we assume that $d_i(s) > 0$ for all i and s (see Table 1 for a complete list of notation). The restriction $d_i(s) > 0$ is necessary to conduct modelling on the

log scale; note that $d_i(s) = 0$ can be accommodated by replacing zero values with an arbitrarily small number. We could also think of restricting the analysis to locations where $d_i(s) > 0$ or employing a zero-inflated mixture model, but these are not considered further in the present treatment.

We prefer to work on the log scale rather than the absolute scale; this transformation allows for an unconstrained parameter space, much in the way that log-link functions are often used in generalized linear models (GLMs) for count data (McCullagh & Nelder, 1989). Accordingly, we define $y_i(s) = \log(d_i(s))$, and assume that a variance–covariance matrix Σ_i is also available that represents uncertainty in log-scale relative abundance values. Ideally, Σ_i would be available directly from the initial analysis of dataset i ; if not, the delta method (Dorfman, 1938; Ver Hoef, 2012) could be used to translate variance–covariance matrices on the real scale to the log scale, but our experience is that the accuracy of the delta method approximation can be poor for log transformations of imprecise surfaces. If surface i is generated with a generalized linear mixed model (GLMM) or generalized additive model (GAM), it will also often be necessary to add a small value (e.g. 10^{-5}) to the diagonal of the prediction variance–covariance matrix to make it positive-definite. In our experience, this procedure, sometimes called ‘Tikhonov regularization’ (Golub et al., 1999) is not very sensitive to the choice of (small) constant.

Using bold notation to represent vectors (so for instance \mathbf{y}_i represents a vector of log-relative abundance values for surface i), we construct a GLMM-type model for the surfaces as

$$\mathbf{y}_i = \alpha_i + \boldsymbol{\mu} + \boldsymbol{\xi}_i + \boldsymbol{\varepsilon}_i, \quad (1)$$

where α_i is a scaling random effect for each dataset, $\boldsymbol{\mu}$ is a vector of mean zero, spatially autocorrelated mean effects, $\boldsymbol{\xi}_i$ is a vector of mean zero, spatially correlated random effects that represents the bias of

each dataset relative to the mean and $\boldsymbol{\varepsilon}_i$ is random error, assumed to have a multivariate normal distribution with mean zero and known variance–covariance matrix Σ_i . This formulation is somewhat similar to a hierarchical model previously proposed for making joint inference from relative abundance time series (Conn, 2010), but extended to the spatial realm. In particular, it acknowledges that surface values are subject to observation error (through $\boldsymbol{\varepsilon}_i$), but can also be subject to unknown levels of process error (through $\boldsymbol{\xi}_i$). Inference focuses on $\boldsymbol{\mu}$ which represents mean log-scale relative abundance. A single surface of normalized abundance computed from $\boldsymbol{\mu}$ is

$$\pi(s) = \frac{\exp(\boldsymbol{\mu}(s))}{\sum_{k=1}^n \exp(\boldsymbol{\mu}(k))}. \quad (2)$$

Several notes are worth making at this point. First, it seems likely that $\boldsymbol{\mu}$ and $\boldsymbol{\xi}_i$ will be poorly identified in cases where there are only a few surfaces available. Possible constraints would be to set $\boldsymbol{\xi}_i = \mathbf{0}$ for one of the datasets (e.g. for a scientific survey where no systematic bias is expected). Another approach would be to set $\boldsymbol{\xi}_i = \mathbf{0}$ for all i (i.e. to assume all surfaces are unbiased), in which case the model will reduce to something very much like inverse variance weighting often used in meta-analysis (e.g. Fleiss, 1993). Second, we note that there is no need to actually estimate $\boldsymbol{\varepsilon}_i$. Rather, it implies a multivariate probability distribution for each data source, specifically

$$\mathbf{y}_i \sim \text{Multivariate Normal}(\alpha_i + \boldsymbol{\mu} + \boldsymbol{\xi}_i, \Sigma_i). \quad (3)$$

In addition to the data likelihood (i.e. Equation 3), we imposed spatial autocorrelation in $\boldsymbol{\mu}$ and $\boldsymbol{\xi}_i$ by specifying Matérn covariance structures for each. There are thus $3n + 2$ fixed effects in any unconstrained model, including n scaling parameters (α), n Matérn range parameters for $\boldsymbol{\xi}$ random effects, n Matérn smoothness parameters for $\boldsymbol{\xi}$ random

Parameter	Definition
$\mu(s)$	‘True’ log-scale relative abundance in location s
$\pi(s)$	Integrated relative abundance estimator (derived parameter)
$\xi_i(s)$	Spatially autocorrelated random effect representing bias of surface i in location s (log-scale)
α_i	Scaling parameter for surface i
Data	Definition
$y_i(s)$	Log-scale relative abundance surface value at location s from analysis of data source i
Σ_i	Variance–covariance matrix of $y_i(s)$ values; we assume this is available as a byproduct of analysis of data source i
Simulation quantity	Definition
N	Population size (abundance) used to simulate counts
$N(s)$	True population size in location s
$d_i(s)$	Relative abundance value for surface i in location s (note $y_i(s) = \log(d_i(s))$ are used as data for integrated modelling)
$c_i(s)$	Observed count at location s for dataset i

TABLE 1 Parameter, data and simulation definitions. Notation used in models for spatial surface integration.

effects, as well as a smoothness and range parameter for the μ random effects. To implement our model, we used maximum integrated likelihood. In particular, we developed custom code in Template Model Builder (TMB; Kristensen et al., 2016) to evaluate the integrated likelihood, which could then be optimized in the R Programming Environment (R Development Core Team, 2022). TMB uses Laplace approximation to integrate out random effects. To increase numerical tractability and improve computing times, we took advantage of sparse matrix routines by modelling random effects using Gaussian Markov random fields (GMRF; Rue & Held, 2005), with basis functions constructed using the 'SPDE approach' (Lindgren et al., 2011) as implemented in the R package R-INLA (Rue et al., 2009, www.r-inla.org).

2.2 | Extensions for differing spatial support

The methodology described thus far has implicitly assumed that relative abundance surfaces are developed with identical spatial support; that is, the grid resolution and spatial extent was assumed to be the same for each surface. However, this is unlikely to be the case if relative abundance surfaces are developed from disparate data sets or by different organizations. For example, high intensity scientific surveys might be conducted on smaller study areas than the overall region of interest. For these cases, we developed an extension of our methods and show how to implement it in a vignette-like Appendix S1.

3 | SIMULATION STUDY

We conducted a small simulation study to determine the accuracy of our proposed method, and to compare our results to simpler integration approaches (e.g. just using a single surface or a mean of relative abundance values). Our simulations varied by the number of surfaces (2, 4, or 8), and three data generating models: one in which all surfaces had process error ('All biased'), one in which the first surface was simulated without bias (as in a scientific survey) while the remainder were subject to process error ('One unbiased'), and a third where no surfaces included process error ('All unbiased'). For each of these design points, we conducted 500 simulations, fitting several estimation models to each. These included (1) our model, with all ξ_i estimated ('All biased'); (2) our model, with no bias assumed for the first surface ('One unbiased'; $\xi_1 = \mathbf{0}$); (3) our model, with no bias assumed ('All unbiased'; $\xi_i = \mathbf{0} \forall i$); (4) just using the first surface as an estimator of relative abundance ('First surface'); (5) simply taking the arithmetic mean relative abundance over all surfaces ('Mean'); and (6) our full model in (1), but with a diagonal covariance matrix, Σ_i^* (the diagonal elements of which were set equal to those from Σ_i and off-diagonal elements were set to zero; 'AB - diag'). We considered the final model because it may be difficult to get a full covariance matrix for surfaces obtained from the literature, and also because it may be numerically prohibitive to model dense covariance matrices for large spatial surfaces. For each estimator, we quantified performance using

root mean squared error (RMSE) of relative abundance predictions, computed as

$$\text{RMSE} = \sqrt{S^{-1} \sum_{s=1}^S (\pi(s) - \hat{\pi}(s))^2},$$

where $\pi(s)$ is calculated as in Equation 2.

To generate individual surface values, we used the following approach (also see Figure 1). First, for each simulation, we generated a true abundance surface on a 15×15 grid (distance between neighbouring centroids was set to 1.0). Population size was set to $N = 10,000$, with true abundance values in cell s set to $N(s) = N\pi(s)$, where $\pi(s)$ represents relative abundance, such that $\sum_s \pi(s) = 1$. For each simulation, we set $\pi(s) = a^{-1} \exp(v(s))$ where $a = \sum_s \exp(v(s))$ and $v \sim \text{Multivariate Normal}(\mathbf{0}, \Sigma_v)$. Here, Σ_v represents a variance-covariance matrix with an exponential spatial autocorrelation function with a range parameter of 10, as generated using the `exp.cov` function in the `fields` R package (Nychka et al., 2017). Second, for each surface i that was assumed to be biased (representing preferential sampling, for instance), we generated a biased representation of true surface values as $N_i(s) = N\pi_i(s)$, where $\pi_i(s) = b^{-1} \exp(v(s) + \eta(s))$ where $b = \sum_s \exp(v(s) + \eta(s))$. Here $\eta \sim \text{Multivariate Normal}(\mathbf{0}, \Sigma_\eta)$; we parameterized Σ_η in the same manner as Σ_v except that we used a smaller range parameter of 6. This choice means that process errors varied on a finer spatial scale than true abundance (Figure 1). Next, we simulated sampling on our grid by assuming that a total of 25 randomly selected grid cells (out of 225) were sampled, and that counts in each cell were Poisson distributed, with $c_i(s) \sim \text{Poisson}(N_i(s))$. Finally, each of the count datasets were analysed with a spatial model to produce y_i and Σ_i . This was accomplished in TMB using an estimation model, where $c_i(s) \sim \text{Poisson}(\exp(y_i(s)))$, with $y_i(s) = \beta_0 + \kappa_i$ and $\kappa_i \sim \text{Multivariate normal}(\mathbf{0}, \Sigma_k)$. The variance-covariance matrix Σ_k was given a Matérn covariance structure and implemented using a GMRF with an SPDE basis. After all count data sets had been analysed, we used predictions of $y_i(s)$ and its attendant covariance matrix output from TMB for integrated modelling of multiple surfaces (Figure 2).

4 | EXAMPLE: STELLER SEA LIONS

To demonstrate our surface integration approach at a realistic scale, we consider the case of trying to combine two relative abundance maps of Steller sea lions (SSLs) in Alaska. The two surfaces were based on (i) a utilization distribution generated from geolocations of satellite-tagged individuals, and (ii) an analysis of incidental observations recorded by citizen scientists during non-targeted survey operations (POP data). We limited survey boundaries to those initially analysed by Ver Hoef et al. (2021), further restricting records to those west of 144°W longitude [the delineation between western and eastern distinct population segment (DPS) of SSLs]. Within this boundary, we developed an analysis raster composed of $13,658$ 15×15 km grid cells for common analysis. Next, we describe how

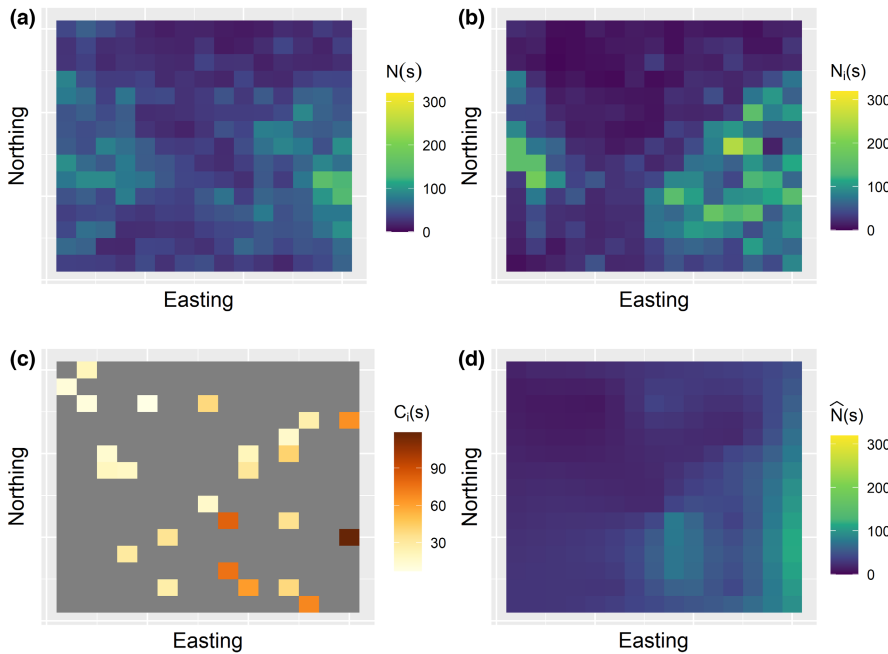


FIGURE 1 A visual depiction of how relative abundance surfaces were generated in the simulation study. Each simulation started with a true animal abundance surface (a). Based on (a), expected abundance surfaces that incorporated spatially autocorrelated bias (b) were then simulated (for surfaces that were assumed to be biased). Next, we simulated count data (c) based on (b), assuming that counts in a simple random sample of grid cells were Poisson distributed. Finally, data in (c) were used to estimate a species abundance surface using a spatial model (d). Note that it is actually the log-scale version of (d) that is used as input to integrated surface modelling.

each of these datasets were analysed before turning our attention to surface integration.

4.1 | Utilization distribution surface

We used telemetry locations from 36 SSLs captured near 11 rookeries and haul-out sites in the central Aleutian Islands, Alaska, USA (Figure 3). Ten of the individuals were adult females captured from 2011 to 2015 and equipped with both Fastloc-GPS satellite transmitters and archival tags that transmitted to Argos satellites (Wildlife Computers). Argos and GPS location data were available for these individuals. The remaining sea lions consisted of 10 juvenile females and 16 juvenile males. Juvenile sea lions were captured during 2000, 2004 and 2005 and equipped with either satellite-linked depth records (SDRs; Wildlife Computers) or satellite relay data loggers (SRDLs; SMRU Instrumentation) that only recorded Argos locations. Details concerning capture protocols are available in Lander et al. (2019, 2021). On average, individuals were monitored for 78 days (range: 7–253 days) during which time a mean of 1283 locations (range 68–8278 locations) were recorded. Relevant field work was conducted under the authority of MMPA/ESA Permits 782-1532, 782-1768, 14,326 and 18,528; IACUC Protocol Nos. A/NW-2010-4, A/NW-2013-2, and 2887-09; and ACUC Protocol No. 2010-14R 09-28.

In addition to positional data, the satellite telemetry devices also collected conductivity (wet/dry) information that indicated when a sea lion was at sea or hauled out on land. The devices from Wildlife Computers recorded wet/dry status over 20-min intervals, whereas the SRDLs (SMRU Instrumentation) recorded the start and end times of haul-out (dry) periods only. We aligned all available dive activity data by converting the SRDL haul-out periods to 20-min intervals following a majority rule. In other words, 20-min intervals were assumed dry if > 10 min fell within a recorded haul-out period, and wet otherwise.

We used the habitat-driven Langevin diffusion model of Michelot, Gloaguen, et al. (2019) to estimate the utilization distribution from the telemetry data. This continuous time model links animal movement to habitat covariates and has an explicit stationary distribution that is synonymous with the utilization distribution. Three habitat covariates were included in the SSL movement model: depth, sea floor slope and the Euclidean distance to the nearest rookery or haul-out site (Wilson et al., 2018). Similar to Michelot, Gloaguen, et al. (2019), we used a two-stage approach to account for Fastloc/ARGOS measurement error and missing locations (e.g. when animals were diving below the surface or hauled out on land). We first predicted locations to match the 20-min resolution of the wet/dry activity data using a continuous-time correlated random walk model that accounts for haul-out behaviour (Johnson et al., 2008). We then fit the Langevin diffusion model to the predicted tracks using an Euler discretization approximation of the likelihood (Michelot, Gloaguen, et al., 2019), where the predicted location at time $t + 1$ (\mathbf{s}_{t+1}) is modelled as

$$\mathbf{s}_{t+1} = \mathbf{s}_t + \frac{\sigma^2 \Delta_t}{2} \sum_{k=1}^K \delta_k \nabla x_k(\mathbf{s}_t) + \epsilon_t,$$

where $K = 3$ is the number of covariates, $x_k(\mathbf{s}_t)$ is the k th covariate evaluated at \mathbf{s}_t , δ_k is the resource selection coefficient for covariate k , $\epsilon_t \sim \mathcal{N}(\mathbf{0}, \sigma^2 \Delta_t \mathbf{I})$, Δ_t is the duration of the interval from time t to $t + 1$, and ∇ is the gradient operator. The limiting distribution of this model has the same form as a standard resource selection function,

$$\pi(\mathbf{s}) = \frac{\exp\left(\sum_{k=1}^K \delta_k x_k(\mathbf{s})\right)}{\int \exp\left(\sum_{k=1}^K \delta_k x_k(\mathbf{z})\right) d\mathbf{z}},$$

and it can be interpreted as the utilization distribution. All steps of this analysis were performed in R (R Development Core Team, 2022). The `crawl` package (Johnson & London, 2018) was used for path

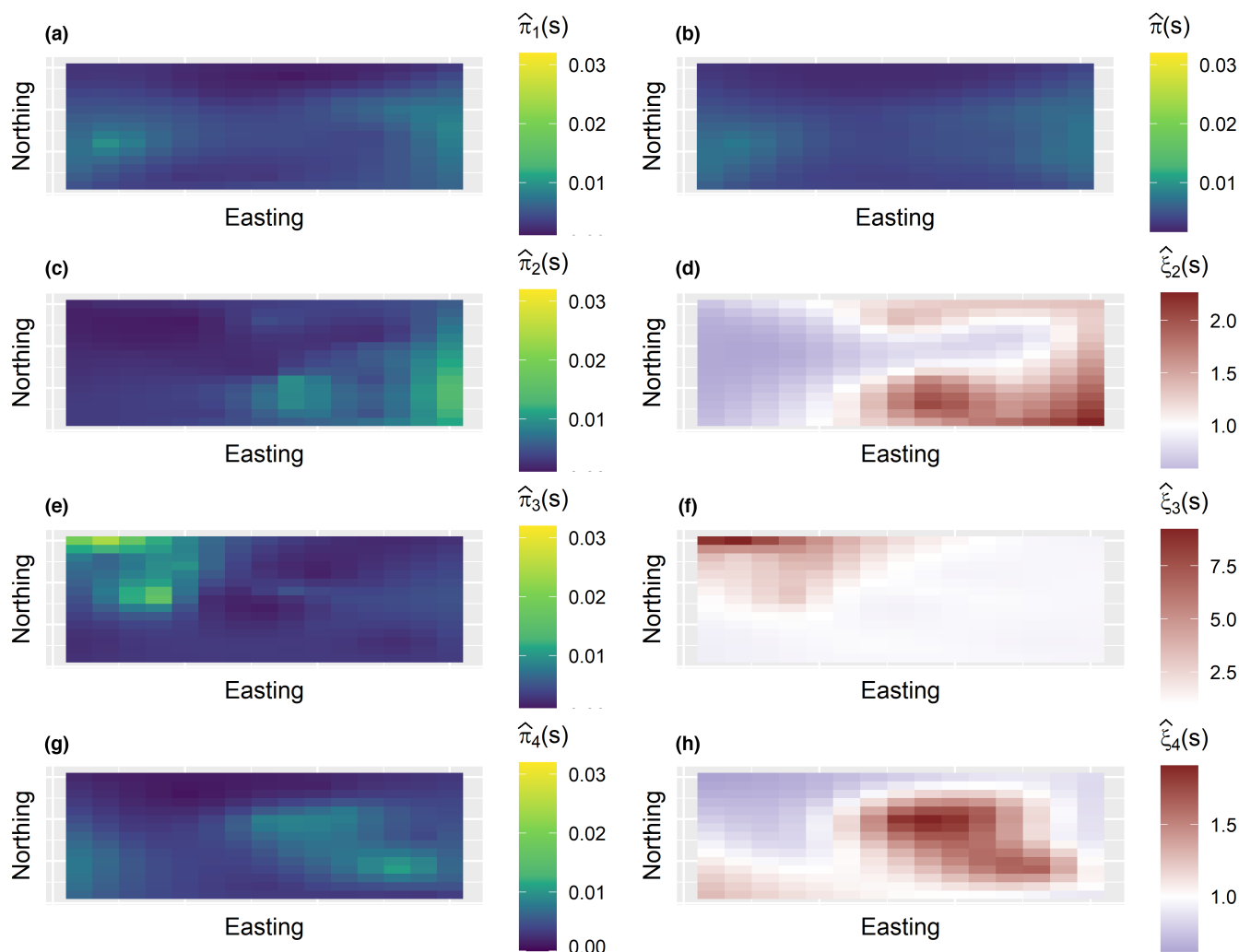


FIGURE 2 An example of the process of fitting integrated surface models to multiple datasets. In this example, panels in the first column (a, c, e, g) give relative abundance surfaces that are subject to integrated modelling (note that it is actually log-scale versions of these that are modelled). In this particular simulation, we assumed the first surface (a) was unbiased, but that surfaces c, e, and g were subject to potential bias. Estimated log-scale process errors (ξ_i) for these surfaces are shown in panels d, f, and h, respectively. The final, integrated surface (on the relative abundance scale) is shown in b. Note that it looks very similar to the surface that was assumed to be unbiased (i.e. a). The integrated surface also resembles the true abundance map (Figure 1a), albeit a highly smoothed version.

prediction, and the `momentuHMM` package (version 2.0.0; McClintock & Michelot, 2018) was used for fitting the Langevin diffusion using maximum likelihood methods. The `Rhabbit` package was used for approximating the gradients of the covariates using bilinear interpolation (Michelot, Gloaguen, et al., 2019). The estimated habitat selection coefficients were $\hat{\delta}_1 = 0.77$ (SE = 0.39) for depth, $\hat{\delta}_2 = 0.05$ (SE = 0.11) for slope, and $\hat{\delta}_3 = -0.08$ (SE = 0.02) for distance to nearest rookery or haul-out site, where a positive (respectively negative) coefficient indicates selection for larger (smaller) values of the covariate.

4.2 | POP surface

A platform of opportunity (POP) dataset of marine mammal sightings was opportunistically collected on ships from 1958 to 2016, containing 109,465 records from the Bering Sea and Gulf of Alaska,

situated between Alaska (United States) and Siberia (Russia), which we subsetting to the months from May to September for SSLs. These data were analysed by Himes Boor and Small (2012) and Ver Hoef et al. (2021). Here, we reproduce the results of Ver Hoef et al. (2021) but on the 15km × 15km raster grid developed for the combined analysis. After subsetting, there were 13,658 SSL records.

In brief, the method of Ver Hoef et al. (2021) attempts to create a smooth spatial surface of effort by using the records for *all* species, and then using this surface as an offset in a spatial Poisson regression model for the total number of SSLs counted in a given grid cell (counts were summed across repeat visits). The resulting surface is the expected count of SSLs for an average transit of a ship through the spatial polygon. For the purposes of defining the neighbours for the conditional autoregressive model (CAR) model used in Ver Hoef et al. (2021), here we used the eight nearest neighbours (including diagonal neighbours).

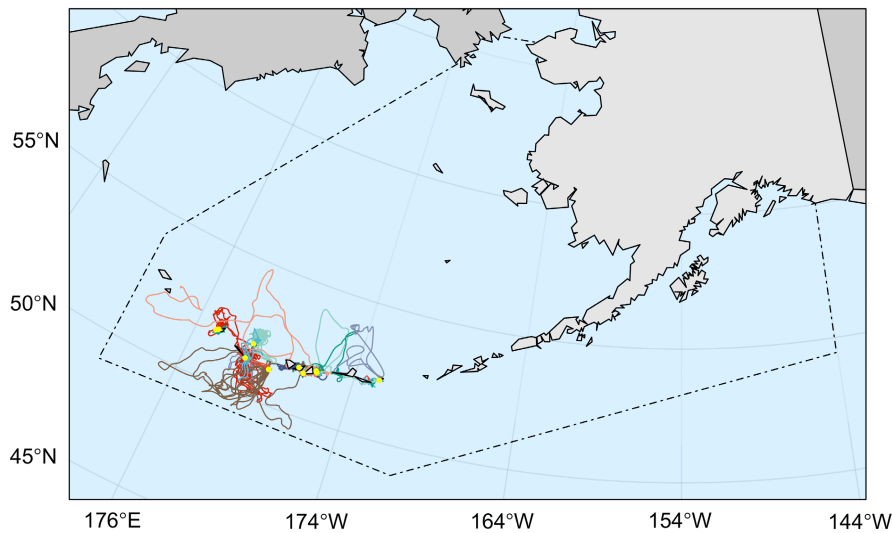


FIGURE 3 A map of predicted Steller sea lion tracks from a random walk movement model, colour-coded by individual. Initial tagging locations are represented by yellow circles, and the study area used for surface melding is shown as a black polygon (dashed line).

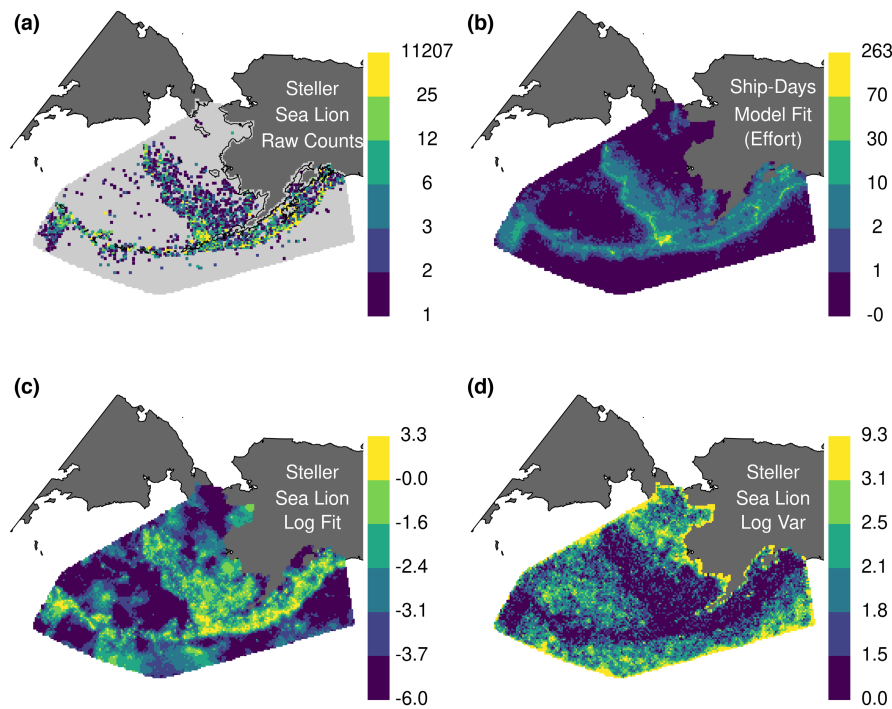


FIGURE 4 Steller sea lion (SSL) platform-of-opportunity (POP) data and model fits. (a) Raw counts per polygon. (b) Modelled effort surface (as ship-days), used as an offset when modelling SSL counts. (c) Modelled SSL counts, expressed as log of expected counts per ship-day. (d) Variance of the modelled SSL counts. This is the diagonal of the variance–covariance matrix, and note that the full covariance matrix is available as described in the text.

The raw SSL counts are shown in Figure 4a. Because these are citizen-science data, there is effort bias in the SSL counts. Figure 4b shows the modelled effort surface, where details on estimating this surface are provided in Ver Hoef et al. (2021). The counts were modelled as,

$$[\mathbf{c} \mid \beta_0, \mathbf{z}, \mathbf{e}, v_c] = \text{NB}(\exp(\beta_0 + \log(\mathbf{e}) + \mathbf{z}), v_c), \quad (4)$$

where \mathbf{c} is a vector of conditionally independent observed counts in the raster polygons, $\text{NB}(\mathbf{a}, v_c)$ is a negative binomial distribution with mean vector \mathbf{a} and common variance v_c , β_0 is an intercept, \mathbf{e} is a spatially explicit effort surface (Figure 4b), and \mathbf{z} is a spatially autocorrelated random effect, whose distribution was a multivariate conditional-autoregression (CAR) model (e.g. Ver Hoef et al., 2018).

The model in Equation 4 was treated as a Bayesian hierarchical model and fit with Markov chain Monte Carlo (MCMC) methods. Let $\beta_0^{[k]}$ be the k th MCMC sample from the posterior distribution of β_0 , where $k = 1, 2, \dots, 1000$ and likewise let $\mathbf{z}^{[k]}$ be the k th MCMC sample from the posterior distribution of \mathbf{z} . Then, the k th MCMC sample modelled relative abundance of SSLs, on the log scale, standardized to a single ship-day, is $\mathbf{a}^{[k]} \equiv \beta_0^{[k]} + \mathbf{z}^{[k]}$. The mean of $\mathbf{a}^{[k]}$, call it $\bar{\mathbf{a}}$, over all k is given in Figure 4c, and the polygon-wise variance of $\mathbf{a}^{[k]}$ over all k is given in Figure 4d. Then, for the purposes of combining spatial surfaces in Equation 1, we let $\mathbf{y}_i = \bar{\mathbf{a}}$, where i is the index for the POP surface, and

$$\Sigma_i = \frac{1}{1000} \sum_{k=1}^{1000} (\mathbf{a}^{[k]} - \bar{\mathbf{a}})(\mathbf{a}^{[k]} - \bar{\mathbf{a}})^T.$$

4.3 | Integrated Steller sea lion surfaces

We used three approaches to create an integrated surface from UD and POP surfaces. First, we simply took the mean relative abundance value for each grid cell; that is, $\pi(s) = 0.5(\pi_1(s) + \pi_2(s))$, where $\pi(s)$ gives the mean relative abundance value and $\pi_1(s)$ and $\pi_2(s)$ give relative abundance values for the UD and POP surfaces, respectively. This corresponds to the 'Mean' estimator in the simulation study (cf. Figure 5). Second, we fitted a TMB model, where process errors (i.e. ξ_i) were estimated for each surface and where diagonal versions of covariance matrices Σ_i were used (i.e. the 'AB - diag' estimator from the simulation study). We used the diagonalized version because it allowed for sparse matrix routines in TMB, which was important since the full, dense $13,658 \times 13,658$ covariance matrices quickly exhausted memory on the Dell Precision laptop (30GB of RAM) used to conduct analysis. Third, we used a model with ξ_i fixed to zero, corresponding to a scenario where each surface is viewed as unbiased (i.e. subject to observation error only). This option also used diagonalized covariance matrices. For both TMB estimation approaches, we employed a reduced-rank predictive process model (Banerjee et al., 2008) for μ and ξ parameters to reduce computational burden. Specifically, we modelled μ as $\mu = A\mu^*$, where μ^* is log-scale relative abundance at 1000 knots and A is a $13,658 \times 1000$ prediction weight matrix. We used a `kmeans` procedure to set knot locations and the `inla.spde.make.A` function in INLA to determine A . This is the same workflow for implementing predictive process models included in the `sdmTMB` package (Anderson et al., 2022).

5 | RESULTS

5.1 | Simulation study

The performance of different estimators of relative abundance (as measured by RMSE) was dependent on the number of surfaces available as well as whether individual surfaces were biased or not (Figure 5). In particular, the accuracy of relative abundance predictions increased (i.e. had lower RMSE) as the number of surfaces increased, and when individual surfaces were unbiased. Interestingly, just taking the mean relative abundance had similar or better performance than the full model with process errors (i.e. $\xi_i(s)$) in the case where all underlying surfaces were subject to bias. Also, there was evidence of increased performance when using multiple surfaces, even when only one was unbiased. Evidently, biased surfaces can still contribute information to relative abundance predictions, at least when potential bias is accounted for via random effects. Approximating the multivariate normal error distribution with a diagonal version often exhibited decreased performance relative to using full, dense covariance matrices.

5.2 | Steller sea lion analysis

Integrated models fitted to the two SSL surfaces exhibited similarities and differences with the original surfaces, depending on the

assumptions used (Figure 6). In particular, the model assuming each surface is subject to bias produced a log-scale estimate that more closely resembled the POP surface, whereas the model assuming no bias more strongly resembled the UD surface. Both of the combined surfaces represented a compromise between the original surfaces, however (both in terms of absolute value on the log scale, and in predictions of absolute relative abundance on the real scale). All of the integrated maps showed high relative abundance along Aleutian islands and along the Alaska coast into the Gulf of Alaska, with fewer sea lions present in the northern parts of the Bering Sea than the original POP and UD surfaces. One point of difference between the integrated maps was relative abundance in the central Aleutians, with the 'Biased' model agreeing with the POP surface that there were relatively few animals there, and the 'Unbiased' model agreeing with the UD surface that it was a location of reasonably high relative abundance (note that this is the location where SSLs were captured and fitted with satellite tags).

6 | DISCUSSION

In this study, we developed a GLMM-type method for combining relative abundance surfaces and applied it to simulated and Steller sea lion datasets. We found that melding multiple surfaces improved relative abundance predictions even if some of the surfaces are biased. This means that it will often make sense to combine spatial surfaces, even if one or more data sources are potentially biased—a finding that has also been made when studying the performance of integrated SDMs (Pacifi et al., 2019). However, the optimal type of estimation procedure depended on characteristics of the underlying data. For instance, if one surface was generated without bias (as with a scientific survey), the best course of action was to employ our model with process errors (i.e. ξ_i) set equal to zero for that surface. If all surfaces could be assumed unbiased, it was best to set all ξ values equal to zero. Surprisingly, if all surfaces were biased, averaging relative abundance surface values often resulted in the best performance.

For western DPS Steller sea lions, our integrated models produced estimates of relative abundance that were somewhere in between POP and UD surfaces (Figure 6). However, the degree to which integrated surfaces resembled the original UD and POP surfaces depended upon modelling assumptions. Based on the results of our simulations, it would appear that with just two surfaces that the 'Biased' surface and the arithmetic mean surface would be the most reliable to use for management purposes (e.g. to calculate 'takes' under the U.S. Marine Mammal Protection Act). Given an estimate N of overall SSL abundance (e.g. from adjusted aerial survey counts), one approach would simply be to compute predictions of at-sea absolute abundance as $N_s = N\pi_s$, where π_s is estimated using our GLMM approach.

Our objective with the SSL data in this paper was primarily to demonstrate integrated modelling approaches for combining spatial surfaces. As such, the maps we generated may need further work before they are appropriate for direct use in SSL management. In particular, the telemetry dataset was limited to the central Aleutians and the UD surface developed from these data relied heavily on distance from

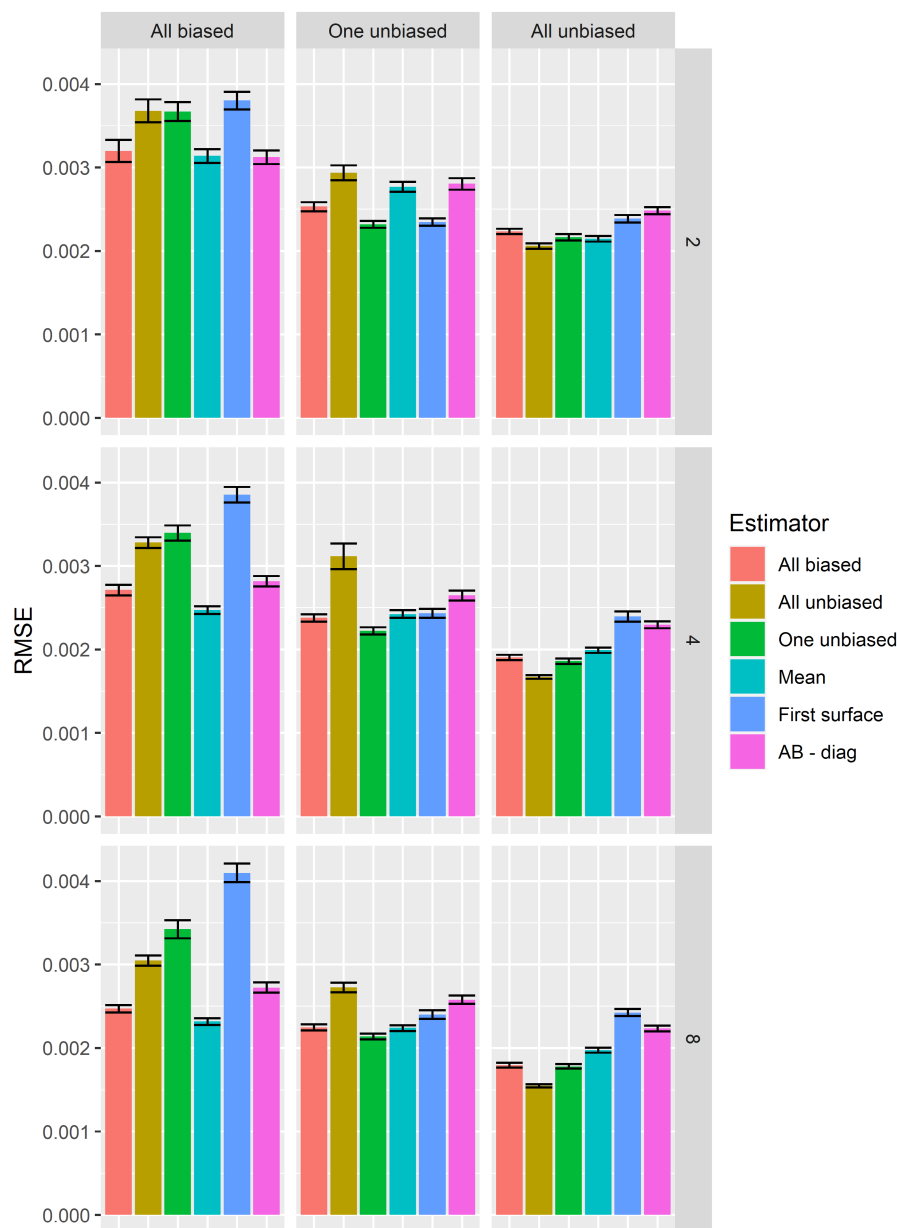


FIGURE 5 Root mean squared error (RMSE) from the simulation study, as a function of the number of surfaces modelled (rows), the underlying level of bias simulated (columns), and the estimator employed (bars). Lower RMSE values indicate more accurate relative abundance predictions. Error bars represent $\pm 2SE$. Note that 'first surface' corresponds to the unbiased surface in the 'One unbiased' scenario, and that 'AB-diag' gives performance for the 'all biased' data integration model in which the error covariance was approximated with a diagonal matrix.

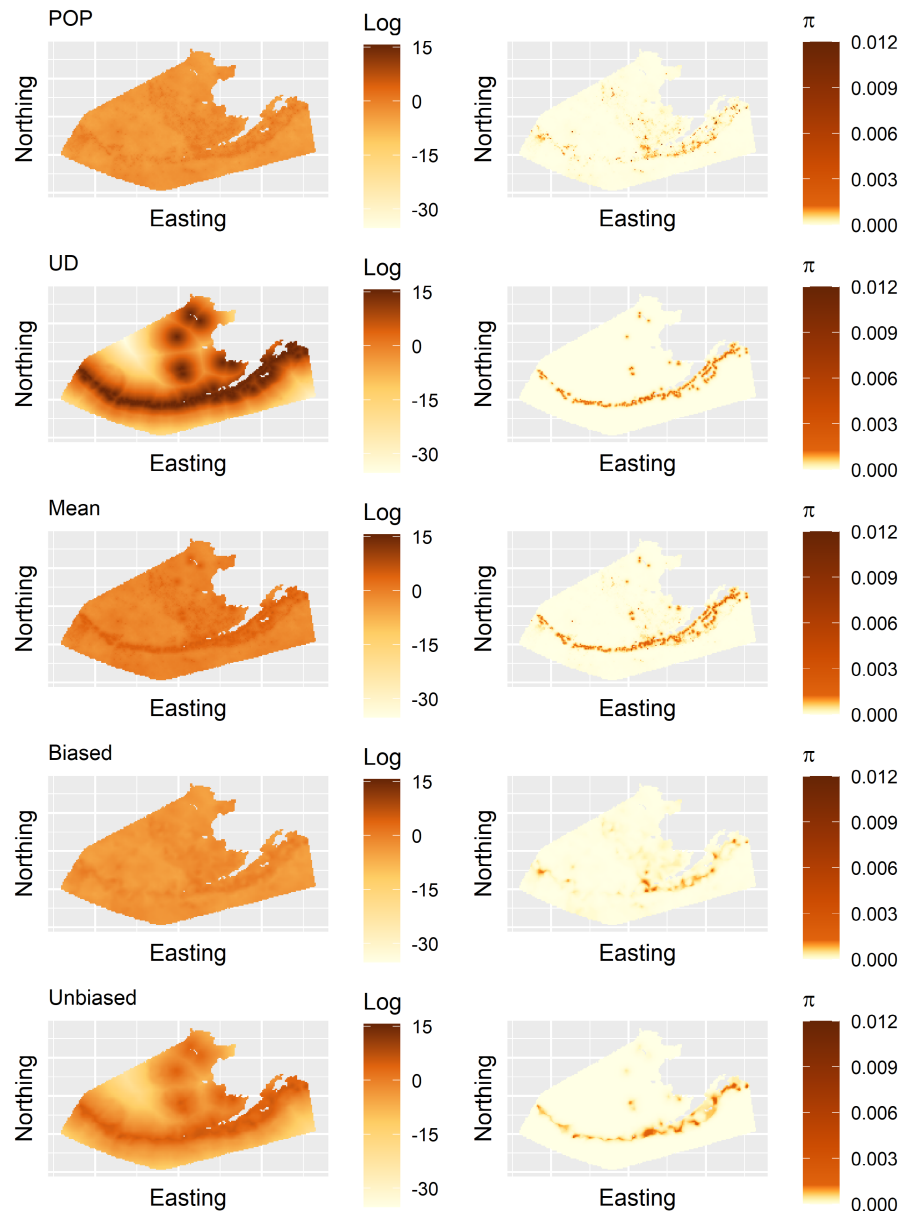
known rookery and haul-out locations to make predictions of space use. However, this covariate does not adequately account for the fact that there is considerable variation in the number of SSLs using individual rookeries or haul-outs. Future research should be devoted to trying to integrate aerial survey estimates into SSL relative abundance maps. These surveys provide 'snapshot' estimates of the number of sea lions using individual rookery or haul-out sites (typically in early summer), but it is unclear how they should be used in the context of 'at sea' distributions that vary seasonally. Future procedures will likely require assumptions about how such point-level data are distributed into surrounding water, perhaps using dispersal kernels.

Analysis of SSL data raised several additional questions that we had not previously considered when analysing simulated data. First, it was apparent that the range of log-scale relative abundance predictions was much wider for the UD surface than for the POP surface. Practically, this means that the UD surface was more definitive about where sea lions were present. Some companion analytical work with

linear models (J. Ver Hoef, unpublished data) suggests that surfaces with high variance in relative abundance values may dominate inference. Although this was the case for the TMB model with process error fixed to zero, it did not hold true for the integrated surface allowing process errors. Second, computational requirements with a fine raster mesh and dense covariance matrices can easily become prohibitive. This led us to consider reduced rank formulations for spatially auto-correlated random effects and diagonalized versions of error covariance matrices, which allowed sparse matrix routines and extremely fast computing times. However, our simulation study suggested the diagonal covariance approximation often resulted in poorer inferences than models with dense covariance matrix error structures (Figure 5). Although we employed numerous procedures intended to increase numerical efficiency, development of alternative procedures for fitting models with dense covariance matrices would clearly be beneficial.

We suspect that most applications of our approach will be for 2–4 surfaces (e.g. telemetry, citizen science and a scientific survey).

FIGURE 6 Estimates of log-scale (left column) and absolute (right column) relative abundance for western DPS Steller sea lions. Rows give original POP and UD surfaces, the arithmetic mean relative abundance, and two sets of estimates from TMB data integration models: one assuming each surface was subject to spatially coloured process error ('Biased'), and the other assuming that differences between the true log-scale relative abundance and the estimated surfaces were solely attributable to observation error ('Unbiased').



However, UD surfaces are often estimated for individual animals, and could themselves be considered individual surfaces. Although recursive Bayesian inference (Hooten et al., 2016, 2021; McCaslin et al., 2021) has shown great promise for combining UD surfaces that share parameters, our approach may be useful for some telemetry applications as well, where the number of surfaces could actually be quite large. In order to meld telemetry data with other sources, researchers would first need to decide whether to treat population-level UD surfaces as a single surface, or to treat individual-level UD surfaces as separate surfaces when applying our approach. For the latter, one might consider a hierarchical formulation for spatial random effects, in an analogous fashion to hierarchical meta-analysis (e.g. Rossetti et al., 2017).

Our focus here has been on species relative abundance surfaces - that is, data for which $\exp(y_i(s)) = a_i N_i(s)$ for some unknown constant a_i (possibly subject to spatially coloured bias). Other surfaces, such as occupancy probabilities on the (0,1) scale derived from logit- or probit-link functions (e.g. Johnson et al., 2013; Royle &

Dorazio, 2008), do not fit into this framework (i.e. the simple scalar a_i is no longer appropriate). However, recent work on integrated SDMs (e.g. Fithian et al., 2015) has emphasized that occupancy probabilities can be expressed as a function of Poisson process intensities by using a complimentary log-log (*cloglog*) link function. Since point process intensities are proportional to abundance (precisely the condition we require), we therefore suggest that researchers wishing to use our approach with occupancy surfaces first conduct inference on the *cloglog* scale. Then, predicted log-Poisson intensities can be used as the $y_i(s)$ values for surface melding.

In Appendix S1, we addressed the problem of spatial misalignment (Bradley et al., 2016; Gotway & Young, 2002; Pacifici et al., 2019), which can happen if surfaces are defined on different spatial support (e.g. different sizes or shapes of grid cells, or different spatial extent). In particular, we used 'proportional allocation' (Gotway & Young, 2002) to relate expected relative abundance in different grid cell configurations (i.e. on the data level) to relative

abundance levels on a reference grid (i.e. the μ level). In our view, this greatly increases the generality of our methods, in that relative abundance surfaces no longer need to be provided on the same set of pixels and same spatial extent. However, additional research would be useful for providing guidelines on the reference grid pixel sizing; our intuition is that the reference grid should have pixel sizing similar to the smallest observed pixel size, but whether this is optimal would require additional research.

7 | CONCLUSIONS

In this paper, we developed a GLMM-type method for estimating a single relative abundance map from multiple species abundance surfaces. In most cases, this method results in increased performance relative to inference based on the analysis of a single data set. Although it will not be appropriate to use in every integrated analysis, we believe it fulfils an important gap for ecologists wishing to combine inferences from multiple surveys. It should be particularly useful in cases where one or more surfaces are subject to bias (e.g. from preferential sampling), when there is no clear underlying process (e.g. point process) that can be used as a common currency with which to construct a fully integrated model, or when fitting fully integrated models is computationally infeasible.

AUTHORS' CONTRIBUTIONS

P.C., J.V.H., D.J., B.M., and B.B. conceived the ideas and designed methodology; P.C. designed and conducted the simulation study and performed the integrated analysis of SSL data; B.B., J.V.H., and D.J. coordinated and formatted SSL data; J.V.H. analysed SSL POP data; B.M. analysed SSL telemetry data; P.B.C., J.V.H., and B.M. drafted the manuscript. All authors edited drafts, provided critical responses during revisions, and gave final approval for publication.

ACKNOWLEDGEMENTS

We thank members of the Alaska Ecosystems Program, Alaska Fisheries Science Center for conducting fieldwork to obtain the telemetry data used in our Steller sea lion example. Views expressed are those of the authors and do not necessarily represent findings or policy of any government agency. Use of trade or brand names does not indicate endorsement by the U.S. government.

CONFLICT OF INTEREST

None declared.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/2041-210X.13948>.

DATA AVAILABILITY STATEMENT

R scripts and data necessary to recreate analyses are currently available at https://github.com/pconn/melding_spatial_surfaces and

have been permanently archived at <https://doi.org/10.5281/zenodo.6784008> (Conn et al., 2022).

ORCID

Paul B. Conn  <https://orcid.org/0000-0002-2801-299X>

Jay M. Ver Hoef  <https://orcid.org/0000-0003-4302-6895>

Brett T. McClintock  <https://orcid.org/0000-0001-6154-4376>

Devin S. Johnson  <https://orcid.org/0000-0002-0068-7098>

Brian Brost  <https://orcid.org/0000-0002-9810-3950>

REFERENCES

- Anderson, S. C., Ward, E. J., English, P. A., & Barnett, L. A. K. (2022). sdmmb: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. *bioRxiv*, <https://doi.org/10.1101/2022.03.24.485545>.
- Banerjee, S., Gelfand, A. E., Finley, A. O., & Sang, H. (2008). Gaussian predictive process models for large spatial datasets. *Journal of the Royal Statistical Society B*, 70, 825–848.
- Bradley, J. R., Wikle, C. K., & Holan, S. H. (2016). Bayesian spatial change of support for count-valued survey data with application to the American Community Survey. *Journal of the American Statistical Association*, 111(514), 472–487.
- Conn, P. B. (2010). Hierarchical analysis of multiple noisy abundance indices. *Canadian Journal of Fisheries and Aquatic Sciences*, 67, 108–120.
- Conn, P. B., Chernook, V. I., Moreland, E. E., Trukhanova, I. S., Regehr, E. V., Vasiliev, A. N., Wilson, R. R., Belikov, S. E., & Boveng, P. L. (2021). Aerial survey estimates of polar bears and their tracks in the Chukchi Sea. *PLoS ONE*, 16(5), e0251130.
- Conn, P. B., Thorson, J. T., & Johnson, D. S. (2017). Confronting preferential sampling when analysing population distributions: diagnosis and model-based triage. *Methods in Ecology and Evolution*, 8(11), 1535–1546.
- Conn, P. B., Ver Hoef, J. M., McClintock, B. T., Brost, B., & Johnson, D. S. (2022). v0.1 pconn\melding_spatial_surfaces: Melding_spatial_surfaces. <https://doi.org/10.5281/zenodo.6784008>
- Diggle, P. J., Menezes, R., & Su, T. I. (2010). Geostatistical inference under preferential sampling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 59(2), 191–232. <https://doi.org/10.1111/j.1467-9876.2009.00701.x>
- Dorazio, R. M. (2014). Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. *Global Ecology and Biogeography*, 23(12), 1472–1484.
- Dorfman, R. (1938). A note on the delta-method for finding variance formulae. *The Biometric Bulletin*, 1, 129–137.
- 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.
- Fithian, W., Elith, J., Hastie, T., & Keith, D. A. (2015). Bias correction in species distribution models: Pooling survey and collection data for multiple species. *Methods in Ecology and Evolution*, 6(4), 424–438.
- Fleiss, J. (1993). Review papers: The statistical basis of meta-analysis. *Statistical Methods in Medical Research*, 2(2), 121–145.
- Golub, G. H., Hansen, P. C., & O'Leary, D. P. (1999). Tikhonov regularization and total least squares. *SIAM Journal on Matrix Analysis and Applications*, 21(1), 185–194.
- Gotway, C. A., & Young, L. J. (2002). Combining incompatible spatial data. *Journal of the American Statistical Association*, 97(458), 632–648.
- Himes Boor, G. K., & Small, R. J. (2012). Steller sea lion spatial-use patterns derived from a Bayesian model of opportunistic observations. *Marine Mammal Science*, 28(4), E375–E403.

- Hooten, M. B., Buderman, F. E., Brost, B. M., Hanks, E. M., & Ivan, J. S. (2016). Hierarchical animal movement models for population-level inference. *Environmetrics*, 27(6), 322–333.
- Hooten, M. B., Johnson, D. S., & Brost, B. M. (2021). Making recursive Bayesian inference accessible. *The American Statistician*, 75(2), 185–194.
- Hooten, M. B., Johnson, D. S., McClintock, B. T., & Morales, J. M. (2017). *Animal movement: Statistical models for telemetry data*. CRC Press.
- Johnson, D. S., Conn, P. B., Hooten, M., Ray, J., & Pond, B. (2013). A probit approach for spatio-temporal modeling of ecological occupancy data. *Ecology*, 94, 801–808.
- Johnson, D. S., & London, J. M. (2018). crawl: An R package for fitting continuous-time correlated random walk models to animal movement data. <https://doi.org/10.5281/zenodo.596464>
- Johnson, D. S., Thomas, D. L., Ver Hoef, J. M., & Christ, A. (2008). A general framework for the analysis of animal resource selection from telemetry data. *Biometrics*, 64(3), 968–976.
- Johnston, A., Moran, N., Musgrove, A., Fink, D., & Baillie, S. R. (2020). Estimating species distributions from spatially biased citizen science data. *Ecological Modelling*, 422, 108927.
- Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. M. (2016). TMB: Automatic differentiation and Laplace approximation. *Journal of Statistical Software*, 70, 1–21.
- Lander, M., Fadely, B., Gelatt, T., & Sterling, J. (2021). Satellite telemetry data used to determine habitat use of adult female Steller sea lions (*Eumetopias jubatus*) in Alaska, 2010–2016. Research Workspace. <https://doi.org/10.24431/rw1k583>
- Lander, M., Fadely, B., Johnson, D., Gelatt, T., Sterling, J., Rehberg, M., & Burns, J. (2019). Steller sea lion (*Eumetopias jubatus*) satellite telemetry data used to determine at-sea distribution in the western-central Aleutian islands, Alaska 2000–2013. Research Workspace, version. <https://doi.org/10.24431/rw1k31z>
- Lindgren, F., Rue, H., & Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4), 423–498.
- MacArthur, R. H. (1972). *Geographical ecology; Patterns in the distribution of species*. Harper & Row.
- Manly, B., McDonald, L., Thomas, D., McDonald, T., & Erickson, W. (2010). *Resource selection by animals* (2nd ed.). Kluwer Academic Publishers.
- McCaslin, H. M., Feuka, A. B., & Hooten, M. B. (2021). Hierarchical computing for hierarchical models in ecology. *Methods in Ecology and Evolution*, 12(2), 245–254.
- McClintock, B. T., Abrahms, B., Chandler, R. B., Conn, P. B., Converse, S. J., Emmet, R., Gardner, B., Hostetter, N. J., & Johnson, D. S. (2021). An integrated path for spatial capture-recapture and animal movement modeling. *Ecology*, e03473.
- McClintock, B. T., & Michelot, T. (2018). momentuHMM: R package for generalized hidden Markov models of animal movement. *Methods in Ecology and Evolution*, 9(6), 1518–1530. <https://doi.org/10.1111/2041-210X.12995>
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*. Chapman and Hall.
- McCullough, D. R. (1996). Spatially structured populations and harvest theory. *The Journal of Wildlife Management*, 60, 1–9.
- Michelot, T., Blackwell, P. G., & Matthiopoulos, J. (2019). Linking resource selection and step selection models for habitat preferences in animals. *Ecology*, 100(1), e02452.
- Michelot, T., Gloaguen, P., Blackwell, P. G., & Étienne, M. P. (2019). The Langevin diffusion as a continuous-time model of animal movement and habitat selection. *Methods in Ecology and Evolution*, 10(11), 1894–1907.
- Nychka, D., Furrer, R., Paige, J., & Sain, S. (2017). *fields: Tools for spatial data*. R package version 12.5. <https://doi.org/10.5065/D6W957CT>; <https://github.com/NCAR/Fields>
- Pacifici, K., Reich, B. J., Miller, D. A., Gardner, B., Stauffer, G., Singh, S., McKerron, A., & Collazo, J. A. (2017). Integrating multiple data sources in species distribution modeling: a framework for data fusion. *Ecology*, 98(3), 840–850.
- Pacifici, K., Reich, B. J., Miller, D. A., & Pease, B. S. (2019). Resolving misaligned spatial data with integrated species distribution models. *Ecology*, 100(6), e02709.
- Phillips, E. M., Horne, J. K., Zamon, J. E., Felis, J. J., & Adams, J. (2019). Does perspective matter? A case study comparing Eulerian and Lagrangian estimates of common murre (*Uria aalge*) distributions. *Ecology and Evolution*, 9(8), 4805–4819.
- R Development Core Team. (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing. <http://www.R-project.org>. ISBN 3-900051-07-0.
- Rossetti, M. R., Tschardt, T., Aguilar, R., & Batáry, P. (2017). Responses of insect herbivores and herbivory to habitat fragmentation: A hierarchical meta-analysis. *Ecology Letters*, 20(2), 264–272.
- Royle, J. A., & Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology*. Academic Press.
- Rue, H., & Held, L. (2005). *Gaussian Markov random fields*. Chapman & Hall/CRC.
- Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2), 319–392.
- Scholes, R. J., Mace, G. M., Turner, W., Geller, G. N., Jürgens, N., Larigauderie, A., Muchoney, D., Walther, B. A., & Mooney, H. A. (2008). Toward a global biodiversity observing system. *Science*, 321(5892), 1044–1045.
- Thurfjell, H., Ciuti, S., & Boyce, M. S. (2014). Applications of step-selection functions in ecology and conservation. *Movement Ecology*, 2(1), 1–12.
- Turchin, P. (1998). *Quantitative analysis of movement: Measuring and modeling population redistribution in animals and plants*. Sinauer Associates.
- Ver Hoef, J. M. (2012). Who invented the delta method? *The American Statistician*, 66(2), 124–127.
- Ver Hoef, J. M., Johnson, D. S., Angliss, R., & Higham, M. (2021). Species density models from opportunistic citizen science data. *Methods in Ecology and Evolution*, 12, 1911–1925.
- Ver Hoef, J. M., Peterson, E. E., Hooten, M. B., Hanks, E. M., & Fortin, M. J. (2018). Spatial autoregressive models for statistical inference from ecological data. *Ecological Monographs*, 88(1), 36–59.
- Wilson, K., Hanks, E., & Johnson, D. (2018). Estimating animal utilization densities using continuous-time Markov chain models. *Methods in Ecology and Evolution*, 9(5), 1232–1240.
- Wilson, K. A., Westphal, M. I., Possingham, H. P., & Elith, J. (2005). Sensitivity of conservation planning to different approaches to using predicted species distribution data. *Biological Conservation*, 122(1), 99–112.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Conn, P. B., Ver Hoef, J. M., McClintock, B. T., Johnson, D. S., & Brost, B. (2022). A GLMM approach for combining multiple relative abundance surfaces. *Methods in Ecology and Evolution*, 13, 2236–2247. <https://doi.org/10.1111/2041-210X.13948>