USING SPATIOTEMPORAL STATISTICAL MODELS TO ESTIMATE ANIMAL ABUNDANCE AND INFER ECOLOGICAL DYNAMICS FROM SURVEY COUNTS

Paul B. Conn[1*], Devin S. Johnson[1], Jay M. Ver Hoef[1], Mevin B. Hooten[2,3,4], Josh M. London[1], and Peter L. Boveng[1]

[1] *National Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA National Marine Fisheries Service, Seattle, Washington 98115 U.S.A.*

[2] *U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, CO 80523 U.S.A.*

[3] *Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO 80523 U.S.A.*

[4] *Department of Statistics, Colorado State University, Fort Collins, CO 80523 U.S.A.*

Appendix B: Details of simulation studies investigating performance of spatiotemporal statistical models for abundance estimation

We conducted two simulation studies to evaluate the performance of different spatiotemporal statistical models for estimating animal abundance. In the first, we simulated "generic" populations, where surveys were assumed to occur over a simple 20×20 grid. The advantage of this approach was that estimator performance could be evaluated using a relatively small number of grid cells which resulted in relatively fast execution times. It also allowed us to investigate performance relative to different data generating models, levels of sampling effort, and degree of change in a hypothetical habitat covariate, thus allowing us to evaluate the robustness of different spatio-temporal statistical models over a variety of sampling situations.

The second simulation study used habitat covariates, effort, and estimated abundance from a recent survey of ice-associated seals (see e.g., Conn *et al.* Online Early) to investigate estimator performance under a more realistic sampling scenario. The purpose of this study was to determine whether different classes of spatio-temporal statistical models were likely to produce reliable estimates of apparent spotted seal abundance in 2012 (apparent in the sense that

estimates are presently unadjusted for incomplete detection). We now describe configuration of each of the simulation studies in greater detail.

1 Simulation study: generic populations

We configured generic simulations such that spatial and temporal variability in animal abundance was dependent (in part) on a hypothetical covariate. Each simulation involved several steps, as follows.

1. Generating a spatio-temporal habitat covariate

To start with, we generated a habitat covariate on a 30×30 grid by convolving a bivariate Gaussian kernel with independent AR1 time series processes occurring at each of $m = 36$ knots located on an even grid around the landscape (Fig. B1). This process is structurally similar to the type of spatiotemporal "dynamical process convolution model" described by Calder *et al.* (2002).

In particular, covariate values for time t are calculated as

$$\mathbf{x}_t = 0.5 + \log(1+\delta)(t-1) + \mathbf{K}\boldsymbol{\alpha}_t,$$

where δ is a fixed parameter influencing the degree of contraction or expansion of the habitat covariate, $\boldsymbol{\alpha}_t$ is a vector of length m giving time-specific weights associated with each knot, and the $S \times m$ dimensional matrix $\mathbf{K}$ maps the weights associated with each knot to S-dimensional space (in this case $S = 900$ gives the total number of grid cells). The entries in $\mathbf{K}$ are proportional to distance-specific kernel densities used to interpolate between the m knots and the S spatio-temporal locations being modeled. We set ($\sigma = 5$), the distance between knot locations, and set the ith row and jth column of $\mathbf{K}$ to $N(d_{ij}; 0, \sigma^2)$, where N denotes the Gaussian probability density function and d_{ij} gives the distance from the centroid of grid cell i to the location of knot j. The elements of $\mathbf{K}$ are then renormalized so that rows sum to 1.0.

We initialized the weights, $\boldsymbol{\alpha}_1$, by drawing them from an intrinsic conditionally autoregressive (ICAR) process with precision parameter $\tau = 15$ (see e.g., Rue & Held 2005). We

then let weights evolve over time according to independent AR1 processes, i.e.,

$$\boldsymbol{\alpha}_t = \boldsymbol{\rho}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\epsilon}, \tag{B.1}$$

where $\boldsymbol{\rho} = 0.9I_S$ is an $(m \times m)$ matrix with $\rho = 0.9$ on the main diagonal and zero elsewhere, and $\boldsymbol{\epsilon}$ is a vector of iid random normal variates with a standard deviation of 0.4. These values were selected to achieve a habitat covariate that would evolve over time at a reasonable speed (see Fig. 1 in the main journal article for an example). After habitat covariates were generated, we (1) divided each covariate by the covariate of maximum value for a particular time step (i.e., $\mathbf{x}_t^* = \mathbf{x}_t / \max(\mathbf{x}_t)$), and then truncated all negative valued covariates at zero (i.e., $\{x_{st}^* | x_{st} < 0\} = 0$). This process helped assure that a reasonable range of covariates would be observed over the range $(0, 1)$, and was an attempt to decrease the likelihood of abundance predictions outside the range of observed data.

2. Simulating animal abundance conditional on hypothetical habitat covariate

For each simulated landscape, we then simulated spatially and temporally specific abundance across the landscape. We used three different data-generating models, including (1) an open population model with restricted dispersal (OPRD), (2) a closed population model with restricted dispersal (CPRD), and (3) a closed population model with unlimited dispersal (CPUD).

The OPRD and CPRD models were constructed in an identical fashion, the difference being that OPRD dynamics were simulated over a larger area than was surveyed, while the CPRD model assumed that surveys covered the entire area. In particular, OPRD dynamics were simulated over the entire (30×30) grid of covariate values while surveys were only conducted over a central (20×20) grid. By contrast, CPRD simulations started with the inner (20×20) grid for generating abundance data. In effect, the OPRD generating model allows animal abundance to cross the boundaries of the the survey area from time step to time step, while CPRD does not. Underlying dynamics proceeded as follows:

A. Initialize latent abundance at the first time step as $\log(\boldsymbol{\lambda}_1) = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{\epsilon}$, where $\mathbf{X}$ denotes an $(S \times 3)$ design matrix, $\boldsymbol{\beta}$ give regression parameters, $\boldsymbol{\eta}$ give spatial random effects, and $\boldsymbol{\epsilon}$ represent iid Gaussian error. In particular, we include linear and quadratic effects of the

habitat covariate, such that the rows of $\mathbf{X}$ are given by $[1 \;\; x_{s1} \;\; x_{s1}^2]$ and set $\boldsymbol{\beta} = [3 \;\; 10 \;\; -10]'$ so that expected abundance would have a unimodal, quadratic relationship with the hypothetical covariate (Fig. B2). Spatial random effects were simulated by drawing from an ICAR($\tau = 15$) distribution, and the variance for exchangable, iid errors (i.e., $\boldsymbol{\epsilon}$) was set to 0.01. This procedure resulted in true abundances that were in the 30,000 - 80,000 range.

B. Simulate latent abundance after time 1 via a resource selection model operating on the real scale. In particular, $\boldsymbol{\lambda}_t = \mathbf{M}_t \boldsymbol{\lambda}_{t-1}$, where $\mathbf{M}_t$ is defined in exactly the same manner as for the OPRS model presented in Appendix S1. Entries of the transition matrix M_t depend on the relative favorability of a given cell as a function of the habitat covariate, and the Euclidean distance between the centroid of the originating cell to the target cell. We again used a linear regression coefficient of 10.0 and a quadratic regression coefficient of -10.0 (Fig. B2) as a measure of habitat favorability, and used a standard deviation of 2.0 cells to set the bivariate Gaussian redistribution kernel (see Appendix A1).

We note that this specification is quite similar to that used in parameter estimation using the OPRS model (Appendix A1). However, an important distinction is that the resource selection model occurs here on the real scale (i.e., $\mathbf{M}_t$ operates on $\boldsymbol{\lambda}$), while in the OPRS estimation model it operates of the log scale. This distinction is important, in that abundance can be held roughly constant during simulations (either on the central (20×20) grid for CPRD or the full (30×30) grid for OPRD). Also, conditional on abundance at the first time step, evolution of the latent abundance process is deterministic in the simulation setting, but not in the estimation setting.

The CPUD model was constructed using the same closed population ideal free (CPIF) formulation as described in Appendix A1. Namely, abundance at time t was simulated as

$$\boldsymbol{\lambda}_t \;\; \sim \;\; \text{Multinomial}(\lambda; \boldsymbol{\pi}_t),$$

where we set $N = 70,000$ and calculated $\boldsymbol{\pi}_t$ as

$$\boldsymbol{\pi}_t \propto \exp(\beta_1 \mathbf{x}_t + \beta_2 \mathbf{x}_t^2 + \mathbf{K}\boldsymbol{\alpha}_t + \boldsymbol{\epsilon}_t). \tag{B.2}$$

Here, we include spatio-temporal random effects through process convolution using the same set of equations as we did for simulating habitat covariates. In particular, $\mathbf{K}$ is set as before, and $\boldsymbol{\alpha}_t$ evolve as in Eq. B.1. For simulating abundance data, we once again set $\beta_1 = 10$ and $\beta_2 = -10$. We set initial $\boldsymbol{\alpha}$ by drawing them from an ICAR($\tau = 15$) distribution, and set the AR1 autocorrelation parameters for $\boldsymbol{\alpha}$ to $\rho = 0.5$ and $\sigma = 0.1$. The precision for exchangeable errors (i.e., the $\boldsymbol{\epsilon}_t$ in Eq. B.2) was set to 20. These values were subjective but not arbitrary; they were obtained via trial and error with the objective that latent abundance evolve neither too fast nor too slow.

3. Simulating counts from transect surveys

At each time step of simulation, transects were simulated as straight, vertical lines across the entire study area, beginning at randomly selected grid cells along the southern terminus (see e.g., Fig. 1 of main article). Each transect thus covered 20 grid cells; we configured simulations such that transects covered 5% of each grid cell they passed through. In the event of > 1 transect per time step, starting grid cells on the southern terminus were sampled via a uniform distribution without replacement. Transect placement was re-randomized at each time step. Although strategic options for spatially balanced sampling would have been possible here, our experience with conducting aerial transect surveys in dynamic environments is that logistical factors such as weather often prevent one from adhering to 'optimal' sampling designs. Random transect placement was an attempt to mimic such a situation.

Following selection of the grid cells and times where sampling occurred, we drew animal counts from a Poisson distribution at times and locations where sampling was conducted:

$$C_{s,t} \sim \text{Poisson}(0.05\lambda_{s,t}).$$

4. Estimation of animal abundance

For each simulated dataset we used 1-4 estimation models (see `Simulation design and performance metrics`) to estimate animal abundance. These included AST, CPIF, OPRS, and STPC (for further detail on these models and prior distributions used, see Appendix A1). Each model was provided with the spatio-temporal covariates used in in abundance generation for use

in the linear predictor part of each estimation model.

1.1 Simulation design and performance metrics

We conducted a small simulation study where simulation replicates varied by (1) the number of transects conducted per time step (1 or 5), and (2) the data generating model (CPRD, CPUD, or OPRD). We generated 100 simulated animal count data sets per design point, each of which had a different value of δ (recall that δ controls the propensity for the habitat covariate to decrease or increase as a function of time). In particular, we let δ range from -0.02 (a two percent expected decrease in the habitat covariate between times steps) to 0.02 in equally spaced increments.

Unfortunately, estimating abundance for each such simulation replicate with each type of estimation model was not feasible owing to computational limitations. To set the number of simulation replicates analyzed by each estimation method, we first conducted a single analysis with each combination of estimation model and number of transects. Desiring to have Monte Carlo error coefficient of variation (CV) below 0.01, we calculated the number of MCMC iterations (n) necessary to obtain this goal. This was accomplished using the relation

$$n = 10,000(1 + 2\sum_{k} \hat{\rho}_k)\hat{\sigma}^2_N \hat{N}^{-2},$$

where $\hat{\rho}_k$ indicates lag k autocorrelation, $\hat{\sigma}^2_N$ gives posterior variance for estimated abundance, and $\hat{N}$ gives the posterior mean for abundance. We then adjusted this estimate upwards depending on appearance of MCMC trace plots to arrive at MCMC run times (Table B1). Including a "burn-in" phase and a 5,000 iteration tuning phase for each analysis, we then estimated requisite run times for each analysis type for the two levels of sample coverage (1 or 5 transects; Table B1).

Examination of Table B1 indicated extremely slow run times for the OPRS model, and that the AST model was appreciably faster than the CPIF or STPC models. Longer run times were needed when there were fewer data (i.e., when there was only one transect for time step). We thus limited the number of simulation replicates analyzed for certain combinations of estimation model and sample coverage (Table B1). For instances where only 10 simulations were analyzed, we chose data sets to cover a range of δ values (by selecting simulations

$5, 15, 25, \ldots, 95$); for those employing 19 simulations, simulations $5, 10, 15, \ldots, 95$ were selected.

For each design point, we calculated proportional relative bias (PRB), mean squared error (MSE), coefficient of variation (CV), and 90% posterior credible interval coverage (Cov90). Proportional bias and mean squared error were calculated differently depending upon the data generating model (open or closed) and estimation model (open or closed). When the generating model was closed but the estimation model was open (i.e., different estimates of abundance were available for each time step), we calculated a single best "average abundance estimator," $\bar{N}$, to calculate PRB and MSE. In particular, we used numerical optimization (via the 'BFGS' minimizer in function `optim` in the R programming environment R Development Core Team 2013) to solve for the $\bar{N}$ that minimizes

$$E\left[\sum_t (\hat{N}_t^{(m)} - \bar{N})^2 | \text{Data}\right],$$

where $\hat{N}_t^{(m)}$ is a posterior realization of abundance at time t (e.g., at the mth iteration of the Markov chain). This approach seeks to minimize Bayes risk for mean squared error loss. To calculate PRB and MSE when an open population model was used to simulate abundance, we simply took the grand mean of PRB and MSE over simulation run and time step. We summarized CV by (a) taking the median CV (posterior SE / posterior mean) with respect to time step for each individual simulation, and (b) averaging median CV across simulation runs. For Cov90, we again took the grand mean of 90% credible interval coverage across both time and simulation run.

In addition to typical summary statistics, we also analyzed whether proportional bias was a function of the strength of habitat changes. In particular, we used simple linear regression to analyze whether bias was a function of δ, the expected proportional change in simulated habitat covariate values (recall this value ranged from -0.02 to 0.02 depending on simulation run).

1.2 Results

Estimator performance summaries are provided in Table B2. Discussion of these results is provided in the main article text. The only design point where there was evidence for a non-zero regression coefficient relating proportional bias to δ was for the STPC estimation model, OPRD

data generating model, and one transect per time period. In this case, there was a significantly negative slope coefficient ($p < 0.001$), indicating that when the habitat covariate is decreasing, one tends to overestimate abundance, and when the habitat covariate is decreasing one tends to overestimate abundance. All twenty other design points had $p > 0.05$.

2 Simulation study: spotted seals in the eastern Bering Sea

Initial analyses of seal abundance in the Bering Sea (e.g., Conn *et al.* Online Early, Ver Hoef *et al.* 2014) indicated several important predictors of abundance, including linear and quadratic effects of sea ice concentration (*ice*), distance from 1000m depth countour (*contour*), and distance from the southern ice edge (*edge*). Spotted seals exhibited the highest concentrations of abundance near the southern ice edge, and were rare in the northern reaches of the surveyed area. We thus developed formulations for spotted seal abundance that mirrored these findings (Fig. B3). We initially considered using all three data generating models to simulate spotted seal abundance; however those employing resource selection (CPRD,OPRD) led to unrealistic distributions. In particular, the rapid dissipation of ice in Bristol Bay often led to large, unrealistic abundance levels in individual sample units. As such, we only used the CPUD data generating model (as defined in preceding sections) to simulate spotted seal abundance. We modeled fixed effects for abundance at time t using the following general structure:

$$\boldsymbol{\theta}_t = 20ice - 13ice^2 - 0.7edge - 0.7edge^{0.5} - 2.5contour. \tag{B.3}$$

Note that $\boldsymbol{\theta}_t$ is in log space, that the *edge* and *contour* covariates were standardized to have a mean of 1.0 over the study area. We set expected abundance for the CPUD model to 280,000. This value is approximately half of estimated abundance from a previous analysis of seal data (Conn *et al.* Online Early) to reflect incomplete detection of seals on sea ice.

In addition to fixed effects for abundance, we again included spatiotemporal random effects in the CPUD model. These were induced in the same manner as for the "generic"

simulations, albeit with a different set of knots (Fig. B4). Sample units (grid cells in Fig. B4) were defined at the resolution of sea ice imagery, which was obtained from the National Snow and Ice Data Center, Boulder, CO on an EASE Grid 2.0 projection. Each sample unit is approximately 625km^2. We simulated a total of 100 data sets.

For each simulation, we used the same procedure as for the "generic" simulations to simulate transect counts, conditioning on the day, time, and effective area sampled by each transect. Transect counts were then analyzed by the three estimation models (AST, CPIF, STPC); each was provided with the same set of covariates and "correct" fixed effects model structure. We then used pilot runs to determine adequate MCMC run time lengths, requisite computing time, and to determine a reasonable number of simulations to analyze for each estimation model (Table B3).

For each simulation, we used the same set of prior distributions specified in Appendix A1 and computed the same set of performance measures as for the "generic" simulation study. However, in this case we also computed these measures on a day-by-day basis. In particular, we were interested in whether estimates made for April 27 (the only day a flight was made into the epicenter of simulated seal abundance along the southwestern ice edge) were more accurate than "average" estimates for the entire time series.

2.1 Results

Simulation performance results for "Bering Sea" simulations are presented in Table B4.

Literature Cited

Calder, C. A., Holloman, C. & Higdon, D. (2002). Exploring space-time structure in ozone concentration using a dynamic process convolution model. *Case Studies in Bayesian Statistics 6* (eds. C. Gatssonis, R. E. Kass, A. Carriquiry, A. Gelman, D. Higdon, D. K. Pauler & I. Verdinelli), pp. 165–176. Springer-Verlag.

Conn, P. B., Ver Hoef, J. M., McClintock, B. T., Moreland, E. E., London, J. M., Cameron,

M. F., Dahle, S. P. & Boveng, P. L. (Online Early). Estimating multispecies abundance using automated detection systems: ice-associated seals in the eastern Bering Sea. *Methods in Ecology and Evolution*, pp. doi: 10.1111/2041–210X.12127.

R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rue, H. & Held, L. (2005). *Gaussian Markov Random Fields*. Chapman & Hall/CR, Boca Raton, Florida, USA.

Ver Hoef, J. M., Cameron, M. F., Boveng, P. L., London, J. M. & Moreland, E. E. (2014). A hierarchical model for abundance of three ice-associated seal species in the eastern Bering Sea. *Statistical Methodology*, **17**, 46–66.

Table B1. Execution details for MCMC analysis of spatiotemporal abundance data as a function of estimation model and number of transects for "generic" simulations. Estimation model and transects/time step are indicated in columns (e.g., AST-5 indicates the AST model applied to data from 5 transects/time step). Presented are the minimum number of MCMC iterations needed to achieve a Monte Carlo CV of 0.01 ('Min iterations'), the number of MCMC iterations used to summarize the posterior distribution ('MCMC length'), the number of additional iterations at the beginning of the MCMC chain that are discarded as a burn-in ('Burn-in'), the number of additional iterations conducted per simulation in order to tune Metropolis-Hastings proposal distributions prior to start of the main MCMC run ('Adapt length'), the estimated number of hours per simulation ('Hrs/sim'), and the actual number of simulations that were analyzed for each estimation model ('# Sims').

	Estimation model							
	AST-1	AST-5	CPIF-1	CPIF-5	OPRS-1	OPRS-5	STPC-1	STPC-5
Min iterations:	13,000	1,014	810	60	45,081	6,218	2560	520
MCMC length:	20,000	10,000	10,000	10,000	50,000	10,000	10,000	10,000
Burn-in:	10,000	10,000	5,000	5,000	10,000	10,000	5,000	5,000
Adapt length:	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000
Hrs/sim:	0.23	0.17	1.5	1.5	71.5	27.5	1.5	1.5
# Sims:	100	100	19	19	0	10	19	19

Table B2. Performance of posterior summaries of abundance for MCMC analysis of spatiotemporal abundance data for "generic" simulations. In particular, we report proportion relative bias (PRB), 90% credible interval coverage (Cov90), coefficient of variation (CV), and relative mean square error (rMSE) as a function of estimation model (Est.mod), number of transects per time step, and generating model (Gen.mod). Details on performance metric calculations are provided in the text.

Est.mod	Transects	Gen.mod	PRB	Cov90	CV	rMSE
AST	1 Transect	CPRD	0.02	0.73	0.07	16.4
AST	1 Transect	CPUD	0.00	0.80	0.06	2.8
AST	1 Transect	OPRD	0.01	0.78	0.07	43.6
AST	5 Transects	CPRD	-0.00	0.76	0.04	0.9
AST	5 Transects	CPUD	0.00	0.86	0.04	0.4
AST	5 Transects	OPRD	0.00	0.80	0.03	6.8
CPIF	1 Transect	CPRD	0.01	0.61	0.03	12.7
CPIF	1 Transect	CPUD	0.01	0.84	0.03	6.5
CPIF	1 Transect	OPRD	-0.00	0.48	0.03	22.2
CPIF	5 Transects	CPRD	0.00	0.72	0.01	0.6
CPIF	5 Transects	CPUD	0.00	0.89	0.01	0.4
CPIF	5 Transects	OPRD	0.01	0.18	0.01	14.3
OPRS	5 Transects	CPRD	0.02	0.88	0.05	1.4
OPRS	5 Transects	CPUD	0.03	0.85	0.05	5.7
OPRS	5 Transects	OPRD	0.02	0.92	0.05	7.5
STPC	1 Transect	CPRD	0.02	0.64	0.05	8.1
STPC	1 Transect	CPUD	0.02	0.57	0.06	9.9
STPC	1 Transect	OPRD	0.01	0.70	0.05	21.0
STPC	5 Transects	CPRD	0.00	0.57	0.02	0.4
STPC	5 Transects	CPUD	0.00	0.34	0.02	0.4
STPC	5 Transects	OPRD	-0.00	0.54	0.02	6.6

Table B3. Execution details for MCMC analysis of spatiotemporal abundance data for spotted seal simulations. For each estimation model, we present the minimum number of MCMC iterations needed to achieve a Monte Carlo CV of 0.01 ('Min iterations'), the number of MCMC iterations ultimately used to summarize the posterior distribution ('MCMC length'), the number of additional iterations at the beginning of the MCMC chain that are discarded as a burn-in ('Burn-in'), the number of additional iterations conducted per simulation in order to tune Metropolis-Hastings proposal distributions prior to start of the main MCMC run ('Adapt length'), the estimated number of hours per simulation ('Hrs/sim'), and the actual number of simulations that were analyzed for each estimation model ('# Sims').

	Estimation model		
	AST	CPIF	STPC
Min iterations:	9,281	7,400	130,548
MCMC length:	30,000	60,000	160,000
Burn-in:	10,000	10,000	10,000
Adapt length:	5,000	5,000	5,000
Hrs/sim:	0.3	9.3	22.0
# Sims:	100	10	10

Table B4. Performance of posterior summaries of abundance for MCMC analysis of spatiotemporal abundance data for simulations constructed to resemble seal sampling in the Bering Sea. In particular, we report proportion relative bias (PRB), 90% credible interval coverage (Cov90), coefficient of variation (CV), and relative mean square error (rMSE) as a function of estimation model ('Est.mod'). Separate performance summaries are presented for the average abundance estimator (i.e., Est.type = $\bar{N}$) and for posterior estimates of abundance on April 27, 2012 (the day a flight transited over the epicenter of spotted seal abundance). Details on performance metric calculations are provided in the text.

Est.type	Est.mod	rMSE	PRB	Cov90	CV
27 Apr	AST	0.56	0.03	1.00	0.18
27 Apr	CPIF	0.43	0.06	0.80	0.06
27 Apr	STPC	0.79	0.03	1.00	0.11
$\bar{N}$	AST	7.45	0.26	0.77	0.26
$\bar{N}$	CPIF	0.44	0.06	0.80	0.06
$\bar{N}$	STPC	0.60	0.04	0.86	0.23

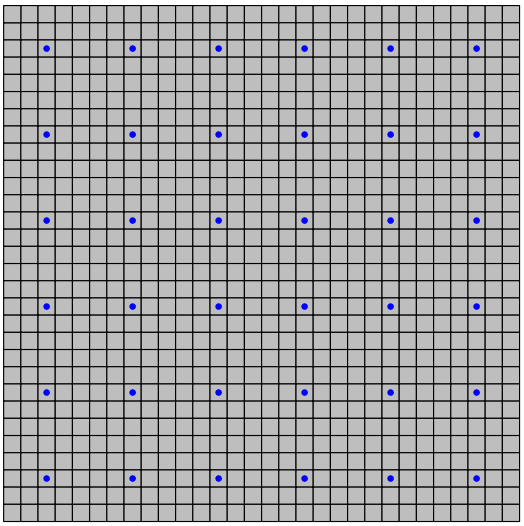

Figure B1. Spatial grid and knot positions (blue points) for simulating spatiotemporal variation in a hypothetical habitat covariate.

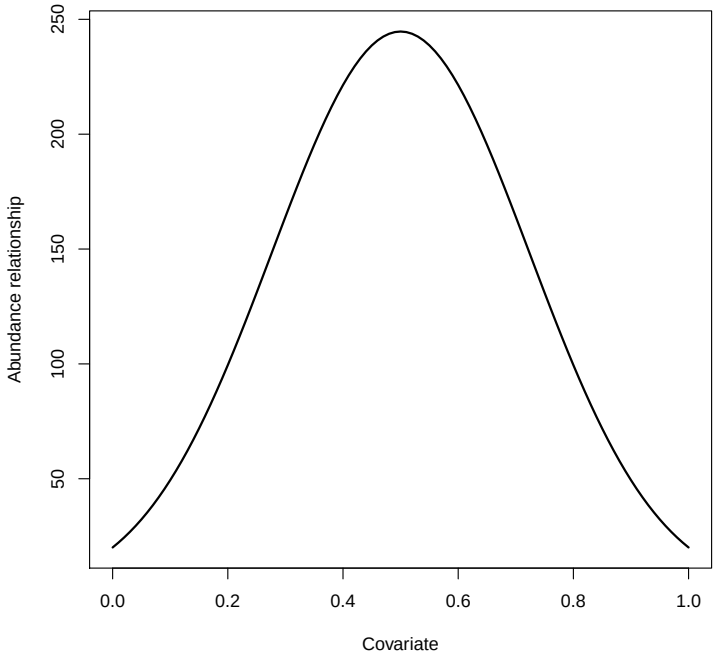

Figure B2. Relationship between expected initial abundance ($\boldsymbol{\lambda}$) and a simulated covariate for "generic" spatiotemporal abundance simulations. The absolute relationship applies to initial abundance (i.e., at the first time step) for both restricted dispersal data generating models (OPRD and CPRD); the same relative relationship (ignoring the intercept) is applied to the evolution of abundance via habitat selection for OPRD and CPRD, and to multinomial cell probabilities for the unlimited dispersal data generating model (CPUD).

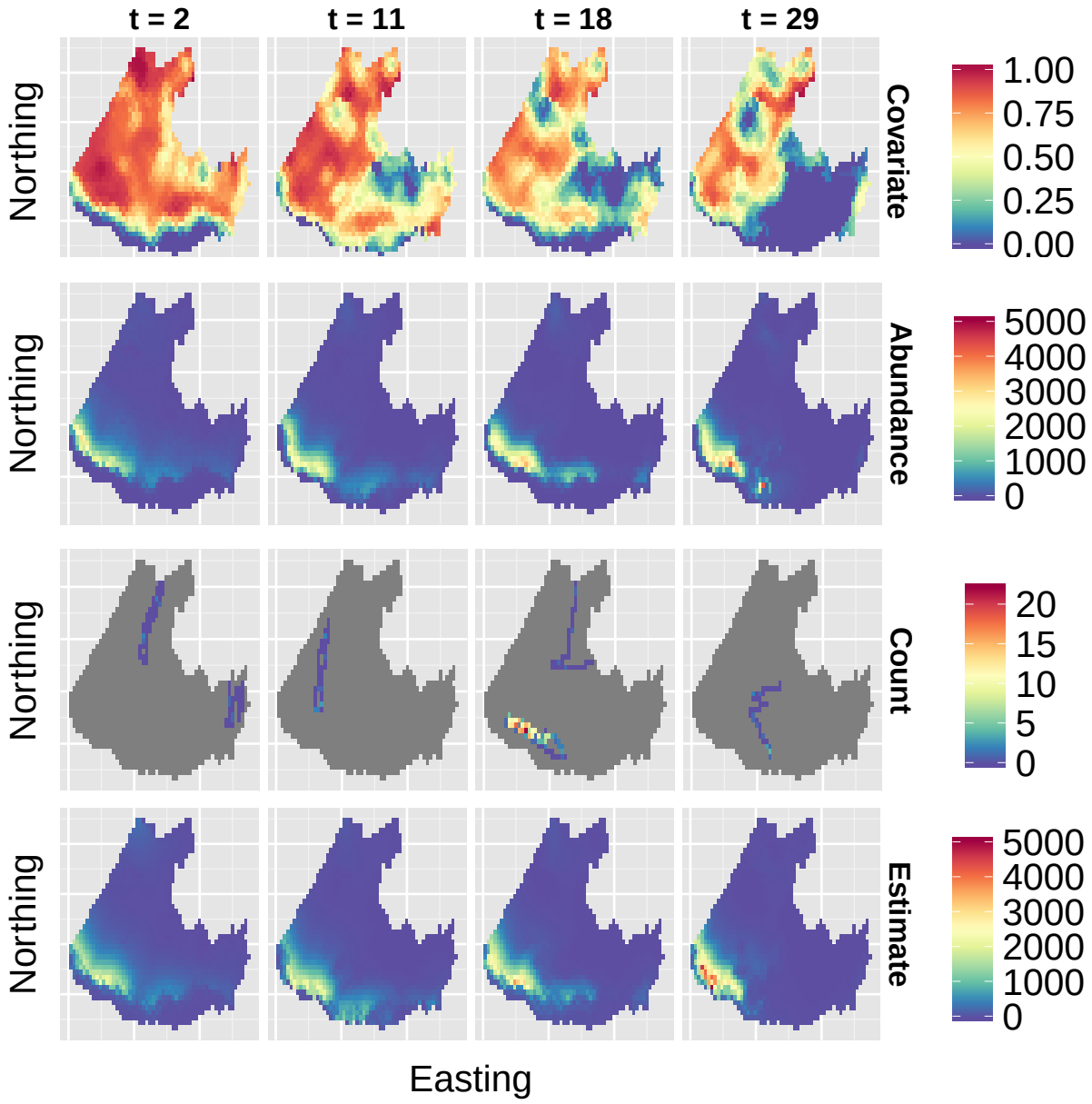

Figure B3. An example of a single spotted seal simulation, whereby 2012 habitat covariates (e.g., daily sea ice concentration, displayed in the first row of panels) are used to simulate animal abundance (second row of panels). Counts are then generated using effort information from 2012 transect surveys (third row of panels), and animal abundance is estimated via a spatiotemporal statistical model (fourth row of panels). Each column depicts results for a given time step. Day one of simulation corresponded to April 10, 2012, and the final day of simulation (day 29) corresponded to May 8, 2012. In this particular simulation, abundance was simulated according to the closed population unlimited dispersal (CPUD) specification, while estimation was conducted assuming the closed population, ideal free (CPIF) model (which included spatiotemporal random effects).

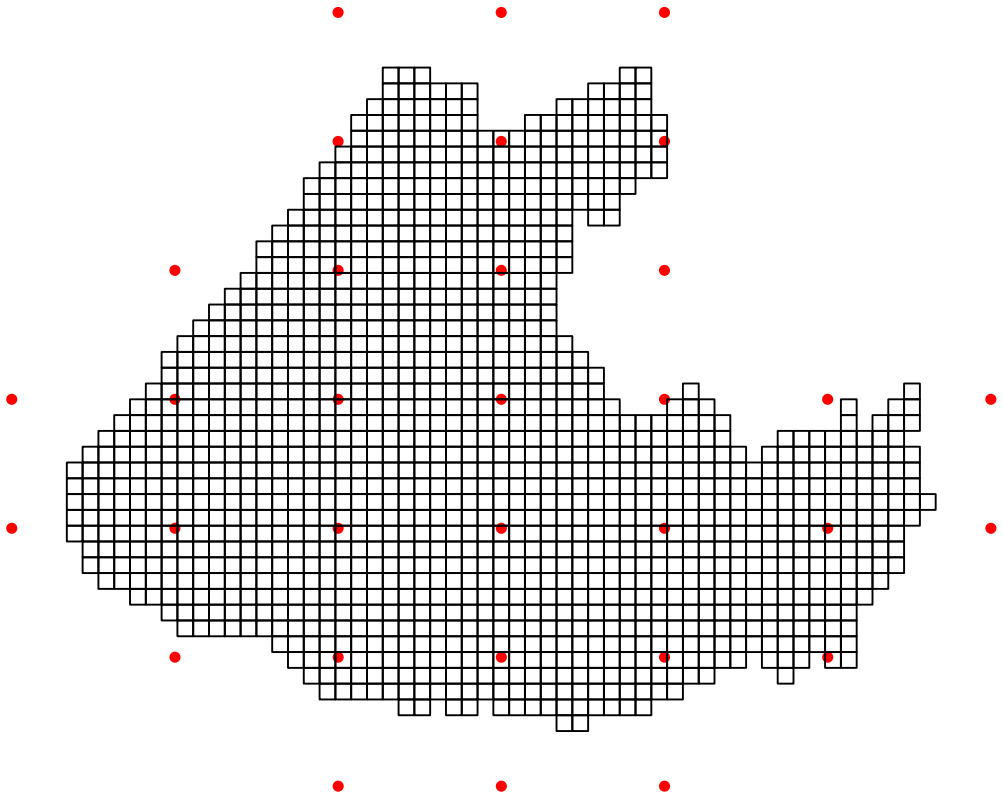

Figure B4. Sample units (grid cells) and knots (red dots) used in spatiotemporal analysis of spotted seal transect counts.