

ARTICLE TEMPLATE

An Application of Spatiotemporal Modeling to Finite Population Abundance Prediction

Matt Higham^a, Carly Hammond^b, Jeff Wells^b

^aSt. Lawrence University Canton, NY 13617; ^bAlaska Department of Fish and Game Fairbanks, AK 99701

ARTICLE HISTORY

Compiled November 29, 2022

ABSTRACT

Insert abstract here.

KEYWORDS

spatial; temporal; kriging;

1. Introduction

1.1. Motivation

Moose surveys in Alaska and western Canada are often performed annually in many regions. The primary goal of these surveys is to predict moose abundance, the total number of moose, in the region. Because of time and money constraints, only some areas (sites) in the region of interest are selected to be in the survey. Biologists fly to these selected sites, count the number of moose, and can then use a spatial statistical model to find a prediction for the finite abundance for that year (Ver Hoef 2008).

Though these surveys are annual, each survey is analysed completely independently of surveys from previous years (e.g. Gasaway et al. 1986; Kellie and DeLong 2006; Boertje et al. 2009; Peters et al. 2014). For example, a model for a survey conducted in the year 2019 only uses counts on sites that were sampled in that year. However, using counts from previous years in a model that incorporates both spatial and temporal correlation (spatiotemporal) could result in a prediction for the realized total or mean that is more precise than predictions from a spatial model using only counts from the most recent survey year.

Though the framework of the motivation is given with an example on moose surveys, this type of analysis could be useful for other monitoring systems with a finite number of sites that are regularly surveyed.

1.2. Background

- Add paragraph about background of spatiotemporal models

CONTACT Matt Higham. Email: mhigham@stlawu.edu, Carly Hammond. Email: carly.hammond@alaska.gov, Jeff Wells. Email: jeff.wells@alaska.gov

Prediction for a total, a subset of the total, or a mean in a finite number of spatial locations should incorporate a finite population correction to the variance of the predictor (Ver Hoef 2008; Higham et al. 2021b). In the context of ecological monitoring in spatiotemporal prediction, we are often most interested in predicting the total abundance for the most recent year of the survey. In this case, the finite population correction should adjust based on the number of sites surveyed in the most recent year of the survey, so that, for example, the prediction variance is zero if all sites in the sampling frame in the most recent year are surveyed.

The rest of this paper is organized as follows. In Section 2, we couple spatiotemporal modeling with finite population prediction to develop the Best-Linear-Unbiased-Predictor for any linear function of a general response variable, including the total abundance across all sites. In Section 3, we apply the predictor to a moose data set in the TOC region of Alaska. In Section 4, we conduct a brief simulation study to examine the properties of the predictor. Finally, in Section 5, we conclude and give directions for future research.

2. Methods

We now give details on the development of the predictor for abundance. We first detail the spatiotemporal model. Because of the heavy use of notation in the spatiotemporal model development, we first introduce a purely spatial model (without temporal variability) and a purely temporal model (without spatial variability). We then build the spatiotemporal model from these two base components and develop a finite population correction factor to give a Best-Linear-Unbiased-Predictor (BLUP) and its prediction variance for any linear function of the response.

2.1. Spatial Model

First, we consider a spatial linear model for a response variable $Y_s(\mathbf{s}_i)$, $i = 1, 2, \dots, n_s$, where the vector \mathbf{s}_i contains the coordinates for the i^{th} spatial site location and n_s is the number of spatial locations. Then, a spatial model for $\mathbf{y}_s(\mathbf{s}_i)$, a vector of the $Y_s(\mathbf{s}_i)$, is

$$\mathbf{y}_s(\mathbf{s}_i) = \mathbf{X}_s \boldsymbol{\beta}_s + \boldsymbol{\epsilon}_s(\mathbf{s}_i), \quad (1)$$

where \mathbf{X}_s is a design matrix for the fixed effects and $\boldsymbol{\beta}_s$ is a parameter vector of fixed effects. The error $\boldsymbol{\epsilon}_s(\mathbf{s}_i)$ can be decomposed into spatial error and independent error components:

$$\boldsymbol{\epsilon}_s(\mathbf{s}_i) = \mathbf{Z}_s \boldsymbol{\delta} + \mathbf{Z}_s \boldsymbol{\gamma}. \quad (2)$$

In equation 2, \mathbf{Z}_s is an $n_s \times n_s$ matrix of 0's and 1's, where the values in a row corresponding to a data point at location \mathbf{s}_i are 1 in the i^{th} column and a 0 in all other columns. Note that, without temporal replication, \mathbf{Z}_s is the identity matrix so is not necessary to include in equation 2. $\boldsymbol{\delta}$ is a random vector independent of $\boldsymbol{\gamma}$ with mean $\mathbf{0}$ and covariance $\text{cov}(\boldsymbol{\delta}) = \sigma_\delta^2 \mathbf{R}_s$, where \mathbf{R}_s is a spatial correlation matrix and σ_δ^2 is sometimes called the spatial partial sill. $\boldsymbol{\gamma}$ is also a random vector with mean $\mathbf{0}$ but has covariance $\text{cov}(\boldsymbol{\gamma}) = \sigma_\gamma^2 \mathbf{I}_s$, where \mathbf{I}_s is the $n_s \times n_s$ identity matrix and σ_γ^2 is sometimes called the spatial nugget.

There are many common parameterizations of \mathbf{R}_s . One common assumption is to assume the covariance function generating \mathbf{R}_s is stationary and isotropic, depending only on the spatial distance between the data points. For example, the exponential covariance function is defined as follows. For observations at locations i and i' at $h_{ii'}$ distance apart, row i and column i' of \mathbf{R}_s is equal to

$$\exp(-h_{ii'}/\phi), \quad (3)$$

where ϕ is a spatial range parameter controlling the decay rate of the covariance as distance between two data points increases.

2.2. Temporal Model

Next, we consider a temporal linear model for a response variable $Y_t(t_j)$, $j = 1, 2, \dots, n_t$, where t_j contains the time index for the j^{th} time point and n_t is the number of time points in the data. Then, a temporal model for $\mathbf{y}_t(t_j)$, a vector of the $Y_t(t_j)$, is

$$\mathbf{y}_t(t_j) = \mathbf{X}_t \boldsymbol{\beta}_t + \boldsymbol{\epsilon}_t(t_j), \quad (4)$$

where \mathbf{X}_t is a design matrix for the fixed effects and $\boldsymbol{\beta}_t$ is a parameter vector of fixed effects. The error $\boldsymbol{\epsilon}_t(t_j)$ can be decomposed into temporal error and independent error components:

$$\boldsymbol{\epsilon}_t(t_j) = \mathbf{Z}_t \boldsymbol{\tau} + \mathbf{Z}_t \boldsymbol{\eta}. \quad (5)$$

In equation 5, \mathbf{Z}_t is an $n_t \times n_t$ matrix of 0's and 1's, where the values in a row corresponding to a data point at time point t_j are a 1 in the j^{th} column and a 0 in all other columns. Note that, without spatial replication, \mathbf{Z}_t is the identity matrix so is not necessary to include in equation 5. $\boldsymbol{\tau}$ is a random vector independent of $\boldsymbol{\eta}$ with mean $\mathbf{0}$ and covariance $\text{cov}(\boldsymbol{\tau}) = \sigma_\tau^2 \mathbf{R}_t$, where \mathbf{R}_t is a temporal correlation matrix and σ_τ^2 is sometimes called the temporal partial sill. $\boldsymbol{\eta}$ is also a random vector with mean $\mathbf{0}$ but has covariance $\text{cov}(\boldsymbol{\eta}) = \sigma_\eta^2 \mathbf{I}_t$, where \mathbf{I}_t is the $n_t \times n_t$ identity matrix and σ_η^2 is sometimes called the temporal nugget.

There are many common parameterizations of \mathbf{R}_t . One common assumption is to assume the covariance function generating \mathbf{R}_t is stationary, depending only on the temporal distance between the data points. For example, the exponential covariance function is defined as follows. For observations at time points j and j' at $m_{jj'}$ units apart, row j and column j' of \mathbf{R}_t is equal to

$$\exp(-m_{jj'}/\rho), \quad (6)$$

where ρ is a temporal range parameter controlling the decay rate of the covariance as time units between two data points increases. Note that the exponential form of \mathbf{R}_t is equivalent to an AR(1) time series model if the time points are equally spaced and the correlation parameter is greater than zero (Schabenberger and Gotway 2017).

2.3. Spatiotemporal Model

We now combine the spatial error components and temporal error components to formulate a model for data collected across both space and time. Let $Y(\mathbf{s}_i, t_j)$, $i = 1, 2, \dots, n_s$ and $j = 1, 2, \dots, n_t$, be a random variable, where \mathbf{s}_i and n_s are defined in subsection 2.1 and t_j and n_t are defined in subsection 2.2. If each spatial location is represented at every time point, a vector of the $Y(\mathbf{s}_i, t_j)$, denoted $\mathbf{y}(\mathbf{s}_i, t_j)$, has length $n_s \cdot n_t \equiv N$. Then, a spatiotemporal model for $\mathbf{y}(\mathbf{s}_i, t_j)$ is

$$\mathbf{y}(\mathbf{s}_i, t_j) = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}(\mathbf{s}_i, t_j), \quad (7)$$

where \mathbf{X} is a design matrix for the fixed effects and $\boldsymbol{\beta}$ is a parameter vector of fixed effects. The error $\boldsymbol{\epsilon}(\mathbf{s}_i, t_j)$ can be decomposed into spatial and temporal components, as in Dumelle et al. (2021). A simple model incorporates spatial error and temporal error in $\boldsymbol{\epsilon}(\mathbf{s}_i, t_j)$ by summing the spatial and temporal errors defined in equation 2 and equation 5:

$$\boldsymbol{\epsilon}(\mathbf{s}_i, t_j) = \mathbf{Z}_s\boldsymbol{\delta} + \mathbf{Z}_s\boldsymbol{\gamma} + \mathbf{Z}_t\boldsymbol{\tau} + \mathbf{Z}_t\boldsymbol{\eta}. \quad (8)$$

With spatial and temporal replication, \mathbf{Z}_s is an $N \times n_s$ matrix while \mathbf{Z}_t is an $N \times n_t$ matrix. However, even when the spatial covariance function generating $\mathbf{Z}_s\boldsymbol{\delta} + \mathbf{Z}_s\boldsymbol{\gamma}$ and the temporal covariance function generating $\mathbf{Z}_t\boldsymbol{\tau} + \mathbf{Z}_t\boldsymbol{\eta}$ are strictly positive definite, the sum of the spatial and temporal components is not necessarily strictly positive definite (Myers and Journel (1990)).

A more flexible option that is always positive definite, given in Dumelle et al. (2021), is the product-sum linear mixed model:

$$\mathbf{y}(\mathbf{s}_i, t_j) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_s\boldsymbol{\delta} + \mathbf{Z}_s\boldsymbol{\gamma} + \mathbf{Z}_t\boldsymbol{\tau} + \mathbf{Z}_t\boldsymbol{\eta} + \boldsymbol{\omega} + \boldsymbol{\nu}. \quad (9)$$

In equation 9, $\boldsymbol{\omega}$ is a random vector of length N with mean $\mathbf{0}$ and covariance $\text{cov}(\boldsymbol{\omega}) = \sigma_\omega^2 \mathbf{R}_{st}$ where \mathbf{R}_{st} is a spatiotemporal correlation matrix and σ_ω^2 is sometimes called the spatiotemporal partial sill. $\boldsymbol{\nu}$ is also a random vector of length n with mean $\mathbf{0}$ but has covariance $\text{cov}(\boldsymbol{\nu}) = \sigma_\nu^2 \mathbf{I}_{st}$, where \mathbf{I}_{st} is the $N \times N$ identity matrix and σ_ν^2 is sometimes called the spatiotemporal nugget.

In the product-sum linear mixed model, the formulation for \mathbf{R}_{st} is

$$\mathbf{R}_{st} \equiv \mathbf{Z}_s \mathbf{R}_s \mathbf{Z}'_s \odot \mathbf{Z}_t \mathbf{R}_t \mathbf{Z}'_t,$$

where \odot is the Hadamard product operator.

If we assume that $\boldsymbol{\delta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\tau}$, $\boldsymbol{\eta}$, $\boldsymbol{\omega}$, and $\boldsymbol{\nu}$ are mutually independent of each other, then

$$\text{var}(\mathbf{y}) \equiv \boldsymbol{\Sigma} = \sigma_\delta^2 \mathbf{Z}_s \mathbf{R}_s \mathbf{Z}'_s + \sigma_\gamma^2 \mathbf{Z}_s \mathbf{I}_s \mathbf{Z}'_s + \sigma_\tau^2 \mathbf{Z}_t \mathbf{R}_t \mathbf{Z}'_t + \sigma_\eta^2 \mathbf{Z}_t \mathbf{I}_t \mathbf{Z}'_t + \sigma_\omega^2 \mathbf{R}_{st} + \sigma_\nu^2 \mathbf{I}_{st}. \quad (10)$$

Note that the model in equation 9 does not have any distributional assumptions: we only need to specify the mean and variance of \mathbf{y} . However, if we also assume that \mathbf{y} is multivariate normal (with mean $\mathbf{X}\boldsymbol{\beta} \equiv \boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$ (Equation 10)), then all model parameters can be easily estimated with Maximum Likelihood (ML) or Restricted Maximum Likelihood (REML).

2.4. Finite Population Kriging

The model in equation 9 is for the N -length vector \mathbf{y} . However, often we do not have the resources to sample or observe every spatial site in every year. Therefore, we may have an interest in prediction of the response values on sites that were not observed. Throughout this section, let the subscript o denote data points that were surveyed (both past and present), the subscript u denote data points that were not surveyed, and the subscript a denote all observations. Then, we can re-order the response vector so that

$$\mathbf{y}_a = [\mathbf{y}'_u, \mathbf{y}'_o]'. \quad (11)$$

Our primary goal is to use the model developed in equation 9 to find optimal weights \mathbf{q}' to apply to the observed realizations of \mathbf{y}_o such that $\mathbf{q}'\mathbf{y}_o$ is the Best Linear Unbiased Predictor (BLUP) for $\mathbf{b}'_a\mathbf{y}_a$. For example, if we are interested in the total of the response across all years, then \mathbf{b}_a would be a column vector of 1's, so that we are adding up all values of the response for a predictor of total abundance across all spatial sites and time points.

Unbiasedness implies that $E(\mathbf{q}'\mathbf{y}_o) = E(\mathbf{b}'_a\mathbf{y}_a)$ for all β . So, denoting \mathbf{X}_o as the design matrix for surveyed data points, $\mathbf{q}'\mathbf{X}_o\beta = \mathbf{b}'\mathbf{X}\beta$ for every β , implying that $\mathbf{q}'\mathbf{X}_o = \mathbf{b}'_a\mathbf{X}_a$.

Kriging weights are then found by finding $\boldsymbol{\lambda}_o$, an $n_o \times 1$ vector, such that

$$E\{(\mathbf{q}'\mathbf{y}_o - \mathbf{b}'_a\mathbf{y}_a)(\mathbf{q}'\mathbf{y}_o - \mathbf{b}'_a\mathbf{y}_a)\} - E\{(\boldsymbol{\lambda}'_o\mathbf{y}_o - \mathbf{b}'_a\mathbf{y}_a)(\boldsymbol{\lambda}'_o\mathbf{y}_o - \mathbf{b}'_a\mathbf{y}_a)\} \quad (12)$$

is greater than 0 for all \mathbf{q}' . The prediction equations are

$$\begin{pmatrix} \Sigma_{o,o} & \mathbf{X}_o \\ \mathbf{X}'_o & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ m \end{pmatrix} = \begin{pmatrix} \Sigma_{o,o} & \Sigma_{o,u} \\ \mathbf{X}'_o & \mathbf{X}'_u \end{pmatrix} \begin{pmatrix} \mathbf{b}_o \\ \mathbf{b}_u \end{pmatrix}, \quad (13)$$

where again the subscripts o and u denote observed and unobserved data points. For example, letting n_o denote the number of observed data points, $\Sigma_{o,o}$ denotes the $n_o \times n_o$ submatrix of Σ corresponding only to rows and columns of observed data points and $\Sigma_{u,o}$ denotes the $(N - n_o) \times n_o$ submatrix of Σ corresponding to rows of data points that were not observed and columns of data points that were observed. Then, the optimal prediction weights are

$$\boldsymbol{\lambda}'_o = \mathbf{b}'_o + \mathbf{b}'_u (\Sigma_{u,o} \Sigma_{o,o}^{-1}) - \mathbf{b}'_u (\Sigma_{u,o} \Sigma_{o,o}^{-1}) \mathbf{X}_o (\mathbf{X}'_o \Sigma_{o,o}^{-1} \mathbf{X}_o)^{-1} \mathbf{X}'_o \Sigma_{o,o}^{-1} + \mathbf{b}'_u \mathbf{X}'_u (\mathbf{X}'_o \Sigma_{o,o}^{-1} \mathbf{X}_o)^{-1} \mathbf{X}_o \Sigma_{o,o}^{-1}. \quad (14)$$

The BLUP for $\mathbf{b}'_a\mathbf{y}_a$ is

$$\widehat{\mathbf{b}'_a\mathbf{y}_a} = \boldsymbol{\lambda}'_o \mathbf{y}_o, \quad (15)$$

with a prediction variance of

$$E((\boldsymbol{\lambda}'_o \mathbf{y}_o - \mathbf{b}'_a \mathbf{y}_a)(\boldsymbol{\lambda}'_o \mathbf{y}_o - \mathbf{b}'_a \mathbf{y}_a)) = \boldsymbol{\lambda}'_o \Sigma_{o,o} \boldsymbol{\lambda}_o - 2\mathbf{b}'_a \Sigma_{a,o} \boldsymbol{\lambda}_o + \mathbf{b}'_a \Sigma_{a,a} \mathbf{b}_a. \quad (16)$$

A common predictor of interest is the total abundance in the most recent time point of the survey. Then, \mathbf{b}_a is a vector of 1's and 0's, where the k^{th} element of \mathbf{b}_a is a 1

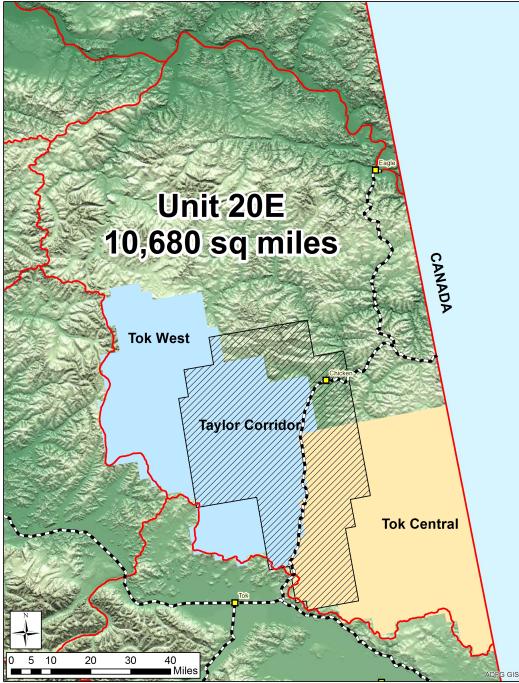


Figure 1. A map of the Taylor Corridor in the TOK region of Alaska.

if the k^{th} element of \mathbf{y}_a is from the most recent time point of the survey and the k^{th} element of \mathbf{b}_a is a 0 otherwise. If we order \mathbf{y}_a by (1) the unobserved, past data points, (2) the unobserved, current data points, (3) the observed, past data points, and (4) the observed, current data points, then

$$\mathbf{b}_a = [\mathbf{b}'_{up}, \mathbf{b}'_{uc}, \mathbf{b}'_{op}, \mathbf{b}'_{oc}]' = [\mathbf{0}', \mathbf{1}', \mathbf{0}', \mathbf{1}']', \quad (17)$$

where the subscripts up , uc , op , and oc denote unobserved sites in past years, unobserved sites in current years, observed sites in past years, and observed sites in current years, respectively.

$\boldsymbol{\lambda}_o$ can then be rewritten as

$$\boldsymbol{\lambda}'_o = \mathbf{b}'_o + \mathbf{b}'_{uc}(\Sigma_{uc,o}\Sigma_{o,o}^{-1}) - \mathbf{b}'_{uc}(\Sigma_{uc,o}\Sigma_{o,o}^{-1})\mathbf{X}_o(\mathbf{X}'_o\Sigma_{o,o}^{-1}\mathbf{X}_o)^{-1}\mathbf{X}'_o\Sigma_{o,o}^{-1} + \mathbf{b}'_{uc}\mathbf{X}'_{uc}(\mathbf{X}'_o\Sigma_{o,o}^{-1}\mathbf{X}_o)^{-1}\mathbf{X}_o\Sigma_{o,o}^{-1}. \quad (18)$$

with a prediction variance of

$$\boldsymbol{\lambda}'_o\Sigma_{o,o}\boldsymbol{\lambda}_o - 2\mathbf{b}'_c\Sigma_{c,o}\boldsymbol{\lambda}_o + \mathbf{b}'_c\Sigma_{c,c}\mathbf{b}_c, \quad (19)$$

where c denotes observations in the most current time point.

3. Application

3.1. Data Description

Abundance surveys are performed in the Taylor Corridor of the TOK region of Alaska annually (Figure 1). In particular, surveys were conducted every year from 2014

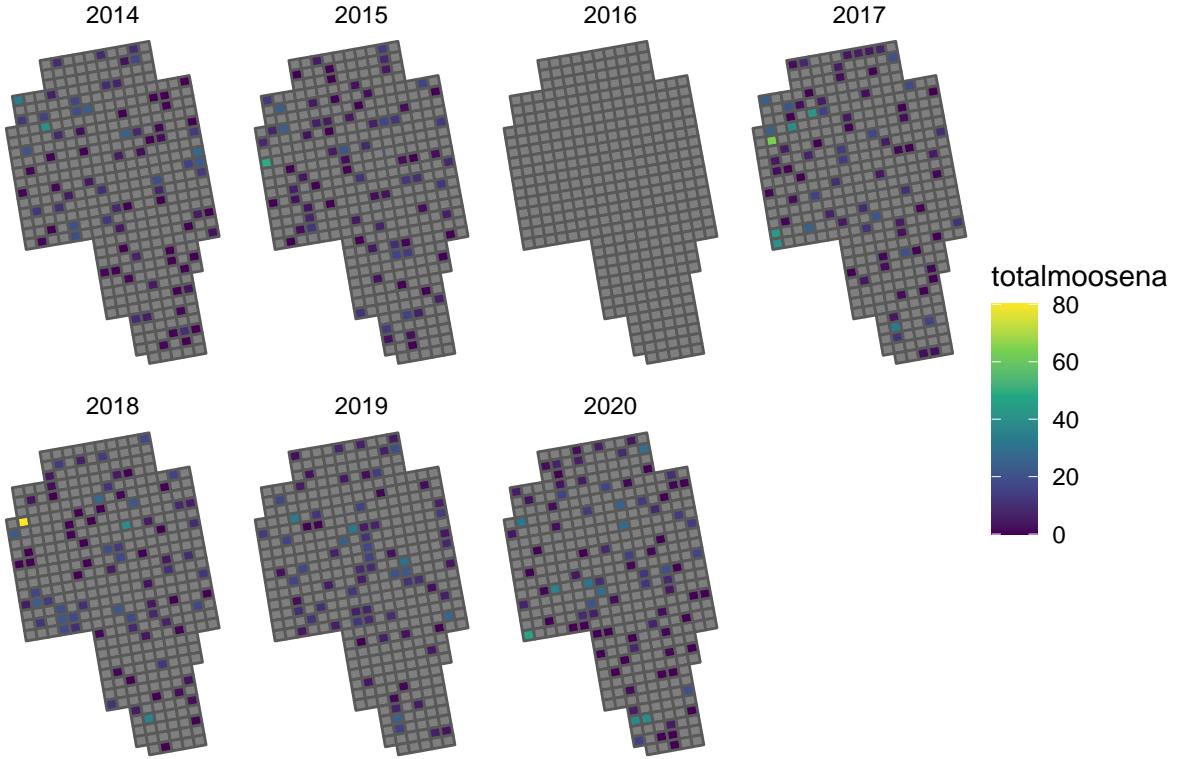


Figure 2. Layout of the spatial sites used to survey moose in the Taylor corridor of the TOK region of Alaska.

through 2020, except for the year 2016, during which there was not sufficient snow cover to perform a survey. There are a total of 381 unique spatial locations, which we refer to as “sites,” and a total of 7 unique time points in the data set, including the missing year of 2016.

In each year of the survey, an aerial team of biologists selects some of the 381 sites to survey. The number of sites in the sampling frame that are selected varies from a low of 76 in the year 2019 to a high of 90 in the year 2020. Throughout the 7 unique time points, some sites are surveyed as many as four or five different times while others are never surveyed (Figure 2).

Before the survey begins in each year, biologists stratify the sites into a “HIGH” stratum and a “LOW” stratum. There are 230 sites in the “HIGH” stratum while there are 151 sites in the “LOW” stratum. The goal of the following analysis is to predict the total abundance of moose across all spatial sites in the year 2020, the most recent year of the survey.

3.2. Model Fitting

We fit the product-sum covariance model defined in equation 9 using REML, with stratum as a covariate in the design matrix, an exponential spatial correlation structure defined in 3, and an exponential temporal correlation structure defined in 6. Table 1 gives the estimated parameters from the model fit.

To help interpret what some of these fitted covariance parameter estimates mean, we can construct a fitted covariance plot (Figure 3). Note that the centroids of two sites directly adjacent to one another are about 4 units apart. As the spatial distance

$\hat{\sigma}_\delta^2$	$\hat{\phi}$	$\hat{\sigma}_\gamma^2$	$\hat{\sigma}_\tau^2$	$\hat{\rho}$	$\hat{\sigma}_\eta^2$	$\hat{\sigma}_\omega^2$	$\hat{\sigma}_\nu^2$
16.9	4.44	3.8	0.9	2.29	0.2	30.8	24.0

Table 1. Estimated covariance parameters in the model. $\hat{\sigma}_\delta^2$, $\hat{\sigma}_\tau^2$, and $\hat{\sigma}_\eta^2$ are the estimated spatial, temporal, and spatiotemporal partial sill, respectively, $\hat{\phi}$ and $\hat{\rho}$ are the estimated spatial and temporal range parameters, and $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_\omega^2$, and $\hat{\sigma}_\nu^2$ are the estimated spatial, temporal, and spatiotemporal nuggets.

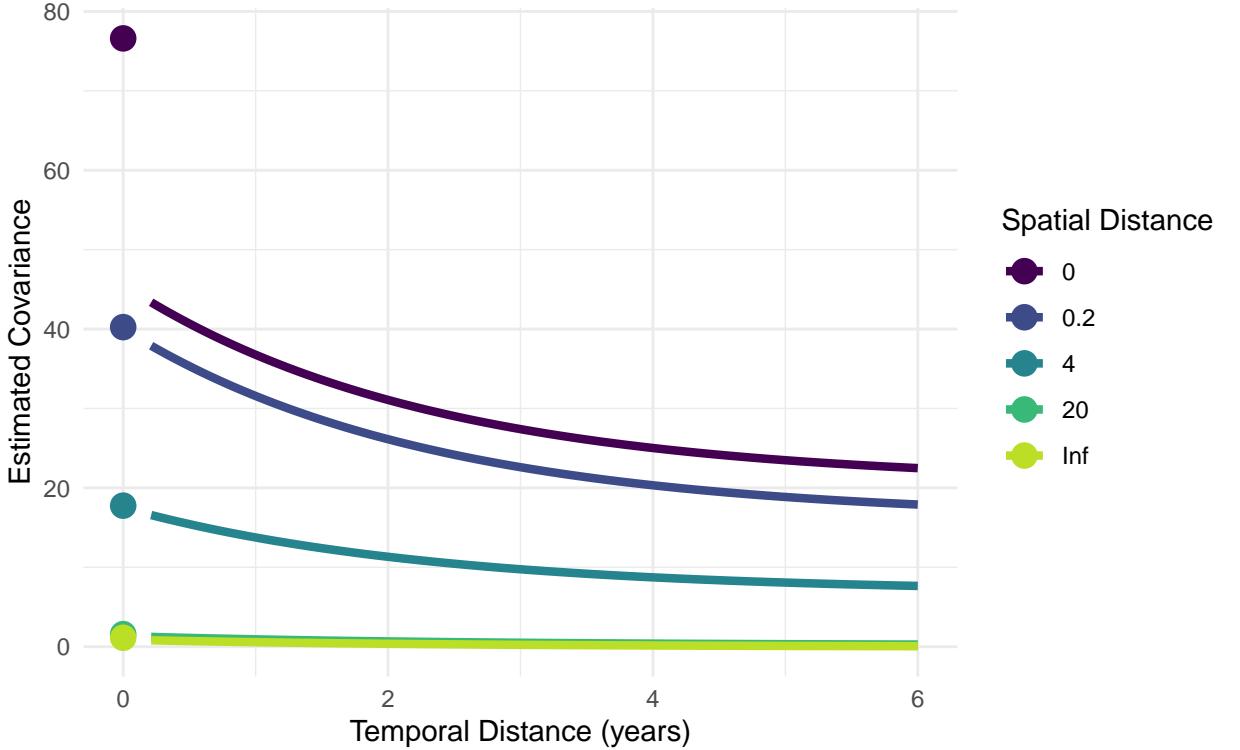


Figure 3. Estimated covariance of the errors from the fitted parameters in a product-sum model.

between two sites increases (dark colour to light colour), the covariance between the response values decreases to 0. In fact, the model estimates the covariance to be nearly 0 when two sites are 20 or more units apart, no matter what the temporal distance is. Note that the fact that the covariance is larger than 0 when temporal distance is 6 and the spatial distance is either 0 implies that including surveys before 2014 could improve precision of the predictor for the total abundance in 2020 even more.

The estimated vector of fixed effects, using "HIGH" as the reference group, is $\beta = (11.26, -9.76)$. Therefore the overall mean for sites in the "HIGH" stratum is estimated to be 11.26 moose while the overall mean for sites in the "LOW" stratum is estimated to be 1.5 moose.

3.3. Prediction

We now use the model in subsection 3.2 to predict the total abundance across all sites in the year 2020, the most recent year of the survey. Plugging in estimates of the covariance parameters into equations 18 and 19 and letting elements of \mathbf{b}_a be 1's for data points in 2020 and 0's otherwise, we obtain a prediction of 2874 moose and a standard error of 234 moose. A 90% normal-based prediction interval for the total

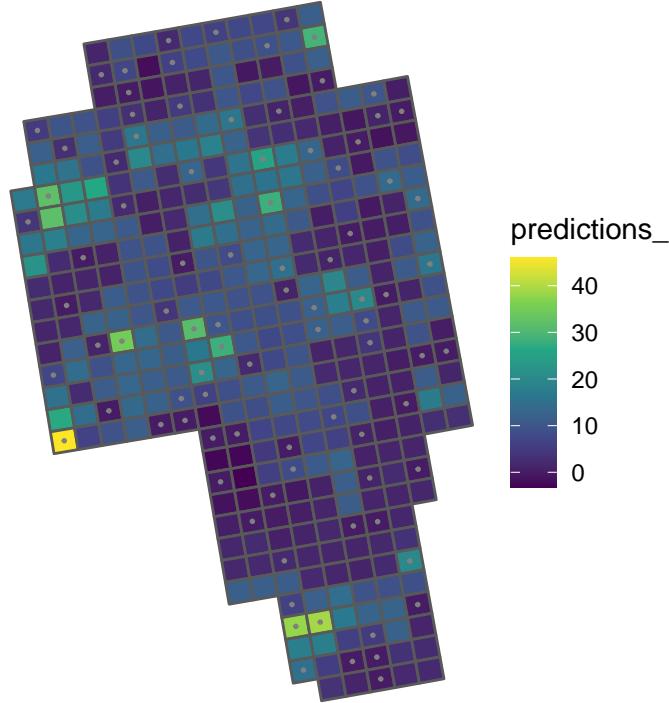


Figure 4. A map of the predictions for the sites in the year 2020. A site with a grey dot in the center means that the site was sampled in 2020.

abundance in 2020 is (2489, 3259) moose. Sitewise predictions for sites in 2020 are given in the map in Figure 4.

For comparison, we use the spatial **sptotal** package (Higham et al. 2021a) to compute the prediction for the total abundance of moose in the year 2020 (Ver Hoef 2008). We also use the standard simple random sampling estimator

$$\bar{y} \cdot \frac{n_s}{n_o},$$

where \bar{y} is the sample mean for the data points in 2020, n_s is the total number of sites in 2020, and n_o is the number of observed data points in 2020. The simple random sampling estimator has a variance for the total abundance of $n_s^2 \cdot \frac{\hat{\sigma}^2}{n_o} \cdot (1 - \frac{n_o}{n_s})$. Note that the purely spatial model fit with **sptotal** and the simple random sampling estimator **only** use data from 2020.

For the purely spatial model, the prediction for the total number of moose in 2020 in the region is 2870 moose with a standard error of 319 moose. For the simple random sampling estimator, the estimated total number of moose in 2020 in the region is 3052 moose with a standard error of 374 moose. While the predictions for the total are somewhat similar across the three methods, we see that the spatiotemporal model is most efficient ($SE = 234$ moose compared to 319 moose for the purely spatial model and 374 moose for the simple random sampling estimator that ignores both spatial and temporal information).

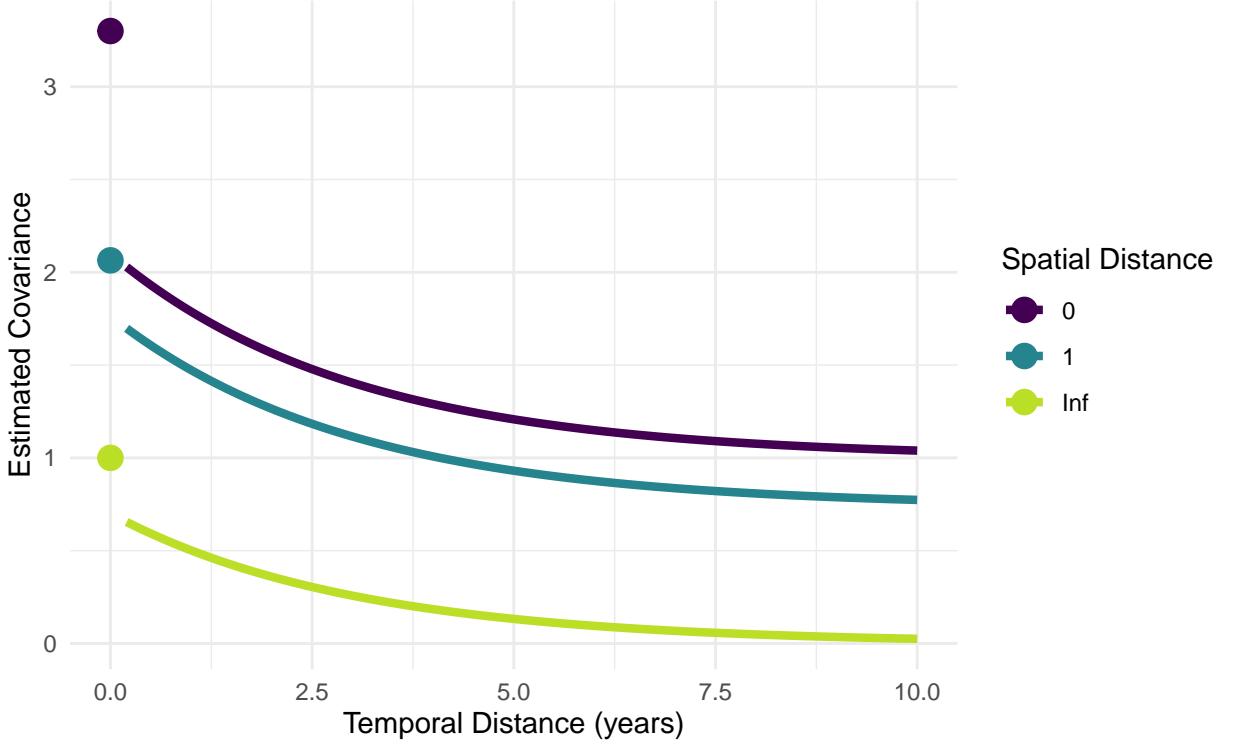


Figure 5. The model covariance used in the simulations for a few spatial distances.

4. Simulation

For a preliminary simulation, we simulate a response vector \mathbf{y} on a 20×20 grid of 400 spatial sites on the unit square and 10 equally-spaced integer time points from 1 to 10. \mathbf{y} is multivariate normal with mean **10** and product-sum covariance matrix Σ defined in equation 10 with the following covariance parameters:

σ_δ^2	σ_γ^2	ϕ	σ_τ^2	σ_η^2	ρ	σ_ω^2	σ_ν^2
0.9	0.1	5	0.7	0.3	3	0.4	0.9

Table 2. Covariance parameters used to simulate data. σ_δ^2 , σ_τ^2 , and σ_η^2 are the estimated spatial, temporal, and spatiotemporal partial sill, respectively, ϕ and ρ are the estimated spatial and temporal range parameters, and σ_γ^2 , σ_ω^2 , and σ_ν^2 are the estimated spatial, temporal, and spatiotemporal nuggets.

Both \mathbf{R}_s and \mathbf{R}_t are generated from the exponential correlation function for the simulations. Of the 4000 total data points, $n = 500$ are selected to be sampled in each iteration. Figure 5 shows the model covariance of the errors used to generate data sets.

Table 3 shows the results from the short simulation study.

coverage	rmspe	median PI length
0.89	39.59	126.77

Table 3. Summary metrics from the simulation study. Coverage is the prediction interval coverage (nominal of 0.90), rmspe is root mean squared prediction error, and median PI length is the median length of a prediction interval.

Note that, as it stands, this set of simulations is more of a “proof of concept.” Below is a proposal for the full simulation study.

Methods to Compare:

- Spatiotemporal prediction
- Spatial prediction (`sptotal`)
- Simple Random Sample estimation with standard finite population correction for the variance of the total

Simulation Design:

The simulations will be run as a $2 \times 4 \times 3$ factorial design with the following parameters and settings:

- model correctly specified with gaussian errors and model incorrectly specified with right-skewed errors (using a similar strategy as what was used in Design v. Model simulation study).
- no spatial or temporal covariance (independent error only), spatial covariance only (no time covariance), temporal covariance only (no spatial covariance), and spatiotemporal covariance.
- small sample size, medium sample size, and large sample size.

5. Discussion

- substantial reduction of se in the application (and, presumably, the simulations).
- normal-based-related limitations
- Bayesian approach, and its drawbacks
- forecasting potential
- take-home message: monitoring programs that use regularly-scheduled surveys might consider incorporating time into their analysis to improve precision of predictors for the mean or total.

References

- Boertje, Rodney D, Mark A Keech, Donald D Young, Kalin A Kellie, and C Tom Seaton. 2009. “Managing for elevated yield of moose in interior Alaska.” *The Journal of Wildlife Management* 73 (3): 314–327.
- Dumelle, Michael, Jay M Ver Hoef, Claudio Fuentes, and Alix Gitelman. 2021. “A linear mixed model formulation for spatio-temporal random processes with computational advances for the product, sum, and product-sum covariance functions.” *Spatial Statistics* 43: 100510.
- Gasaway, William C, Stephen D DuBois, Daniel J Reed, and Samuel J Harbo. 1986. *Estimating moose population parameters from aerial surveys*. Technical Report. University of Alaska. Institute of Arctic Biology.
- Higham, Matt, Jay Ver Hoef, Bryce Frank, and Michael Dumelle. 2021a. *sptotal: Predicting Totals and Weighted Sums from Spatial Data*. R package version 1.0.0, <https://highamm.github.io/sptotal/index.html>.
- Higham, Matt, Jay Ver Hoef, Lisa Madsen, and Andy Aderman. 2021b. “Adjusting a finite population block kriging estimator for imperfect detection.” *Environmetrics* 32 (1): e2654.
- Kellie, Kalin A, and Robert A DeLong. 2006. *Geospatial survey operations manual*. Technical Report. Alaska Department of Fish and Game.
- Myers, Donald E, and Andre Journel. 1990. “Variograms with zonal anisotropies and noninvertible kriging systems.” *Mathematical Geology* 22 (7): 779–785.
- Peters, Wibke, Mark Hebblewhite, Kirby G Smith, Shevenell M Webb, Nathan Webb, Mike Russell, Curtis Stambaugh, and Robert B Anderson. 2014. “Contrasting aerial moose pop-

- ulation estimation methods and evaluating sightability in west-central Alberta, Canada.”
Wildlife Society Bulletin 38 (3): 639–649.
- Schabenberger, Oliver, and Carol A Gotway. 2017. *Statistical methods for spatial data analysis: Texts in statistical science*. Chapman and Hall/CRC.
- Ver Hoef, Jay M. 2008. “Spatial methods for plot-based sampling of wildlife populations.”
Environmental and Ecological Statistics 15 (1): 3–13.