ARTICLE TEMPLATE

An Application of Spatio-temporal Modeling to Finite Population Abundance Prediction

Matt Higham[a], Michael Dumelle[b], Carly Hammond[e], Jay Ver Hoef[d], Jeff Wells[c]

[a]St. Lawrence University Department of Mathematics, Computer Science, and Statistics Canton, NY 13617; [b]United States Environmental Protection Agency Corvallis, OR 97333; [c]Alaska Department of Fish and Game Fairbanks, AK 99701; [d]Marine Mammal Laboratory, Alaska Fisheries Science Center, National Oceanic and Atmospheric Administration Seattle, Washington 98115; [e]Bayer Crop Science Creve Coeur, MO 63141

ARTICLE HISTORY

Compiled February 5, 2023

ABSTRACT

Spatio-temporal predictors can incorporate both spatial and temporal information to predict means, totals, and proportions of a finite resource from a samples of spatial locations collected across time. We develop a spatio-temporal finite-population block kriging (ST-FPBK) predictor that incorporates an appropriate variance reduction for sampling from a finite population. Through an application to moose surveys in the east-central region of Alaska, we show that the predictor has a substantially smaller standard error compared to a predictor from the purely spatial model that is currently used to analyze moose surveys in the region. A separate simulation study shows that the spatio-temporal predictor is unbiased and that prediction intervals from the ST-FPBK predictor attain appropriate coverage. For ecological monitoring surveys completed with some regularity through time, use of ST-FPBK could improve precision. We also give an `R` package that ecologists and resource managers could use to incorporate data from past surveys in predicting a quantity from a current survey.

KEYWORDS

spatial; temporal; kriging; total; resource monitoring

1. Introduction

1.1. *Background*

Spatio-temporal data are indexed by both a spatial index, which we will refer to as a "site," and by a temporal index, which we will refer to as a "time point." Common examples of spatio-temporal data include infections from a disease in a country or region collected over a time period (e.g. Martínez-Beneito, López-Quilez, and Botella-Rocamora 2008; Sahu and Böhning 2022) or climate variables that are recorded through time at multiple locations (Lemos and Sansó 2009).

Models for spatio-temporal data have applications in a wide variety of scientific fields (see Wikle, Zammit-Mangion, and Cressie 2019, for many examples). One such

CONTACT Matt Higham. Email: `mhigham@stlawu.edu`, Michael Dumelle. Email: `Dumelle.Michael@epa.gov`, Carly Hammond. Email: `carly.hammond@bayer.com`, Jay Ver Hoef. Email: `jay.verhoef@noaa.gov`, Jeff Wells. Email: `jeff.wells@alaska.gov`

application is ecological monitoring of a particular resource, such as animal or plant abundance, rainfall, concentration of a compound in soil samples, etc.

In ecological monitoring, we are often interested in prediction of a total or a mean of a particular variable in a finite region at the most recent time point. Ver Hoef (2008) developed Finite Population Block Kriging (FPBK) to predict a linear function of the realized values of a response variable measured at one particular time point in a finite number of sampling units, incorporating a finite population correction to the variance of the predictor. Typically, the linear function is either a mean or a total of the realized values of the response.

1.2. *Motivating Example*

To motivate the development of the predictor in Section 2, we consider moose surveys, which are performed annually or every other year in many regions of Alaska and western Canada. The most common goal of these surveys is to predict moose abundance, the total number of moose, in some region to inform harvest regulations (Kellie, Colson, and Reynolds 2019). Because of time and money constraints, only some spatial indices, or sites, in the region of interest are selected to be in the survey at a particular time point. Biologists fly to these selected sites, count the number of moose, and then use FPBK to find a prediction for the finite abundance for that year. These surveys are historically analyzed with software developed by DeLong (2006), which calculates the "GeoSpatial Population Estimator" (GSPE) for a given survey. The GSPE is an application of the FPBK predictor developed by Ver Hoef (2008).

Though many of these surveys are completed regularly, most are analyzed 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 constructs a prediction for total abundance only from counts on sites that were sampled in that year. However, using counts from previous years in a model that incorporates both spatial and temporal (spatio-temporal) correlation while also using a finite population correction factor based on the proportion of sites surveyed in the most recent year could result in a prediction for the realized total that is more precise than predictions from a purely spatial model. Shortly, we describe such a predictor.

The rest of this paper is organized as follows. In Section 2, we couple spatio-temporal modeling with finite population prediction to develop the Best-Linear-Unbiased-Predictor (BLUP) and its prediction variance for any linear function of a general response variable, including the total abundance across all sites at a particular time point. We call this predictor the ST-FPBK (spatio-temporal Finite Population Block Kriging) predictor. In Section 3, we apply the ST-FPBK to a moose data set in the east-central region of Alaska. In Section 4, we conduct a simulation study to examine the properties of the ST-FPBK predictor and compare its performance to a predictor from a purely spatial model and a simple random sample design-based estimator. Finally, in Section 5, we offer additional thoughts on the application and simulation, and we give directions for future research.

2. Methods

We now give details on the development of the spatio-temporal model and subsequently use this model to 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 vector.

2.1. *Spatio-temporal Model*

Let $Y(\mathbf{s}_i, t_j)$, $i = 1, 2, \ldots, n_s$ and $j = 1, 2, \ldots, n_t$, be a random variable indexed by a spatial site and a time point, where the vector $\mathbf{s}_i$ contains the coordinates for the i^{th} spatial site, n_s is the number of unique sites, t_j is the time index for the j^{th} time point, and n_t is the number of unique time points. If each site 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 $n_s \cdot n_t \equiv N$. Note that, the above formulation assumes that each site is observed at each time point. We choose to make this assumption here because doing so ensures cleaner notation throughout the model development; however, in subsection 2.2, we no longer assume that the response is recorded at every site-time point combination. Then, a spatio-temporal model for $\mathbf{y}(\mathbf{s}_i, t_j)$ is

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

where $\mathbf{X}$ is a design matrix for the fixed effects and $\boldsymbol{\beta}$ is a parameter vector of fixed effects. As in Dumelle et al. (2021), we can decompose the error vector $\boldsymbol{\epsilon}(\mathbf{s}_i, t_j)$ into spatial, temporal, and spatio-temporal components, each of which will be explained in detail in the subsequent paragraphs:

$$\boldsymbol{\epsilon}(\mathbf{s}_i, t_j) = \mathbf{Z}_s\boldsymbol{\delta} + \mathbf{Z}_s\boldsymbol{\gamma} + \mathbf{Z}_t\boldsymbol{\tau} + \mathbf{Z}_t\boldsymbol{\eta} + \boldsymbol{\omega} + \boldsymbol{\nu}. \tag{2}$$

In the spatial component of equation 2 ($\mathbf{Z}_s\boldsymbol{\delta} + \mathbf{Z}_s\boldsymbol{\gamma}$), the matrix $\mathbf{Z}_s$ is an $N \times n_s$ matrix of 0's and 1's, where the values in a row corresponding to a data point at site $\mathbf{s}_i$ are 1 in the i^{th} column and 0 in all other columns. $\boldsymbol{\delta}$ is a random vector with mean $\mathbf{0}$ and covariance $\text{cov}(\boldsymbol{\delta}) = \sigma_\delta^2\mathbf{R}_s$, where $\mathbf{R}_s$ is an $n_s \times n_s$ spatial correlation matrix and σ_δ^2 is called the spatial dependent error variance (or spatial partial sill). The random vector $\boldsymbol{\gamma}$ also has 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 called the spatial independent error variance (or spatial nugget).

In the temporal component of equation 2 ($\mathbf{Z}_t\boldsymbol{\tau} + \mathbf{Z}_t\boldsymbol{\eta}$), $\mathbf{Z}_t$ is an $N \times n_t$ matrix of 0's and 1's, where the values in a row corresponding to a data point at time point t_j are 1 in the j^{th} column and 0 in all other columns. $\boldsymbol{\tau}$ is a random vector with mean $\mathbf{0}$ and covariance $\text{cov}(\boldsymbol{\tau}) = \sigma_\tau^2\mathbf{R}_t$, where $\mathbf{R}_t$ is an $n_t \times n_t$ temporal correlation matrix and σ_τ^2 is called the temporal dependent error variance (or 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 called the temporal independent error variance (or temporal nugget).

In the spatio-temporal component of equation 2 ($\boldsymbol{\omega} + \boldsymbol{\nu}$), $\boldsymbol{\omega}$ is a random vector with mean $\mathbf{0}$ and covariance $\text{cov}(\boldsymbol{\omega}) = \sigma_\omega^2\mathbf{R}_{st}$, where $\mathbf{R}_{st}$ is an $N \times N$ spatio-temporal correlation matrix and σ_ω^2 is sometimes called the spatio-temporal dependent error variance (or spatio-temporal partial sill). $\boldsymbol{\nu}$ is also a random vector 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 spatio-temporal independent error variance (or spatio-temporal nugget).

Though there are a few types of models for the errors that can be built from 2 by setting certain error variances to 0 (e.g. a sum-with-error model sets $\sigma^2_\omega = 0$) and/or by allowing $\mathbf{R}_{st}$ to take certain forms, we focus only on the product-sum model (De Cesare, Myers, and Posa 2001; De Iaco, Myers, and Posa 2001). In a common formulation of the product-sum model, $\mathbf{R}_{st}$ 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. Note that, in order to save on the number of parameters, we will assume that the $\mathbf{R}_s$ and $\mathbf{R}_t$ that form $\mathbf{R}_{st}$ are the same as the $\mathbf{R}_s$ and $\mathbf{R}_t$ associated with $\boldsymbol{\delta}$ and $\boldsymbol{\tau}$, respectively, although this is not necessary in general. $\mathbf{R}_s$ can be parameterized in different ways, but one common assumption is to assume the covariance function generating $\mathbf{R}_s$ is second-order stationary (ie. the covariance between two data points is a function only of the separation vector between two sites) and isotropic (ie. the covariance is a function of the distance only and does not depend on the direction of the separation vector). For example, the exponential covariance function is defined as follows. For observations at sites i and i' at $h_{ii'}$ distance apart, row i and column i' of $\mathbf{R}_s$ is equal to

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

where $\exp(x)$ is equivalent to e^x and ϕ is a spatial range parameter controlling the decay rate of the covariance as distance between two sites increases (Cressie 2015).

Similarly, one common assumption when parameterizing $\mathbf{R}_t$ is to assume the covariance function generating $\mathbf{R}_t$ is second-order stationary (ie. the covariance is a function only of the temporal distance). 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

$$\exp(-m_{jj'}/\rho), \tag{4}$$

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 in the AR(1) series is greater than zero (Schabenberger and Gotway 2017).

The product-sum model for $\mathbf{y}(\mathbf{s}_i, t_j)$ is then

$$\mathbf{y}(\mathbf{s}_i, t_j) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_s\boldsymbol{\delta} + \mathbf{Z}_s\boldsymbol{\gamma} + \mathbf{Z}_t\boldsymbol{\tau} + \mathbf{Z}_t\boldsymbol{\eta} + \boldsymbol{\omega} + \boldsymbol{\nu}, \tag{5}$$

where $\boldsymbol{\delta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\tau}$, $\boldsymbol{\eta}$, $\boldsymbol{\omega}$, and $\boldsymbol{\nu}$ are mutually independent, $\mathbf{y}(\mathbf{s}_i, t_j)$ has mean $\mathbf{X}\boldsymbol{\beta}$, and $\mathbf{y}(\mathbf{s}_i, t_j)$ has covariance

$$\text{var}(\mathbf{y}) \equiv \boldsymbol{\Sigma} = \sigma^2_\delta\mathbf{Z}_s\mathbf{R}_s\mathbf{Z}_s' + \sigma^2_\gamma\mathbf{Z}_s\mathbf{I}_s\mathbf{Z}_s' + \sigma^2_\tau\mathbf{Z}_t\mathbf{R}_t\mathbf{Z}_t' + \sigma^2_\eta\mathbf{Z}_t\mathbf{I}_t\mathbf{Z}_t' + \sigma^2_\omega\mathbf{R}_{st} + \sigma^2_\nu\mathbf{I}_{st}. \tag{6}$$

There are a few reasons for why we choose to solely focus on the product-sum model. First, as long as $\mathbf{R}_s$ and $\mathbf{R}_t$ are positive definite and either $\sigma^2_\omega > 0$ or $\sigma^2_\nu > 0$, then the covariance matrix in equation 6 is also positive definite (De Cesare, Myers, and Posa 2001; De Iaco, Myers, and Posa 2001). Also, the product-sum model is flexible in its ability to model many kinds of spatial and temporal correlation (De Iaco, Palma, and Posa 2015; Dumelle et al. 2021). Xu and Shu (2015) claim that the product-sum model is the most widely used spatio-temporal model used in practical applications.

2.2. *Finite Population Block Kriging*

The model that we developed in the previous section in equation 5 is for the N-length vector $\mathbf{y}$. However, often we do not have the resources to sample or observe every spatial site during every time point. Therefore, we may have an interest in prediction of the response values on sites that were not observed, particularly sites in the most recent time point. Throughout this section, let the subscript o denote data points that were "observed" or sampled, the subscript u denote data points that were "unobserved" or not sampled, and the subscript a denote "all" data points. Then, we can re-order the response vector $\mathbf{y}$ so that

$$\mathbf{y} \equiv \mathbf{y}_a = [\mathbf{y}_u', \mathbf{y}_o']'. \tag{7}$$

Our primary goal is to use the model developed for $\mathbf{y}_a$ in equation 5 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 $\mathbf{b}_a'\mathbf{y}_a$, a linear function of $\mathbf{y}_a$. The N-length vector $\mathbf{b}_a'$ is, for example, a vector of 1's, in which case we would be predicting the total response across all sites and all time points.

Unbiasedness implies that $E(\mathbf{q}'\mathbf{y}_o) = E(\mathbf{b}_a'\mathbf{y}_a)$ for all $\boldsymbol{\beta}$. So, denoting $\mathbf{X}_o$ as the design matrix for the observed data points and $\mathbf{X}_a$ as the design matrix for all data points, $\mathbf{q}'\mathbf{X}_o\boldsymbol{\beta} = \mathbf{b}'{}_a\mathbf{X}_a\boldsymbol{\beta}$ for every $\boldsymbol{\beta}$, implying that $\mathbf{q}'\mathbf{X}_o = \mathbf{b}'{}_a\mathbf{X}_a$. Kriging weights are then found by finding $\boldsymbol{\lambda}_o$, an $n_o \times 1$ column vector, where n_o is the number of observed data points, such that

$$E\{(\mathbf{q}'\mathbf{y}_o - \mathbf{b}'{}_a\mathbf{y}_a)^2\} - E\{(\boldsymbol{\lambda}_o'\mathbf{y}_o - \mathbf{b}'{}_a\mathbf{y}_a)^2\} \tag{8}$$

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

$$\begin{pmatrix} \boldsymbol{\Sigma}_{o,o} & \mathbf{X}_o \\ \mathbf{X}_o' & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ m \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{o,o} & \boldsymbol{\Sigma}_{o,u} \\ \mathbf{X}_o' & \mathbf{X}_u' \end{pmatrix} \begin{pmatrix} \mathbf{b}_o \\ \mathbf{b}_u \end{pmatrix}, \tag{9}$$

where again the subscripts o and u denote observed and unobserved data points. For example, $\boldsymbol{\Sigma}_{o,o}$ denotes the $n_o \times n_o$ submatrix of $\boldsymbol{\Sigma}$ (from equation 6) corresponding only to rows and columns of observed data points and $\boldsymbol{\Sigma}_{u,o}$ denotes the $(N-n_o)\times n_o$ submatrix of $\boldsymbol{\Sigma}$ corresponding to rows of data points that were not observed and columns of data points that were observed. Solving the prediction equations, the optimal prediction weights that are both unbiased and have the smallest possible prediction variance compared to any other linear predictor are

$$\boldsymbol{\lambda}_o' = \mathbf{b}_o' + \mathbf{b}_u' \left[(\boldsymbol{\Sigma}_{u,o}\boldsymbol{\Sigma}_{o,o}^{-1}) - (\boldsymbol{\Sigma}_{u,o}\boldsymbol{\Sigma}_{o,o}^{-1})\mathbf{X}_o\mathbf{W}_o^{-1}\mathbf{X}_o'\boldsymbol{\Sigma}_{o,o}^{-1} + \mathbf{X}_u'\mathbf{W}_o^{-1}\mathbf{X}_o\boldsymbol{\Sigma}_{o,o}^{-1}\right], \tag{10}$$

where $\mathbf{W}_o = \mathbf{X}_o'\boldsymbol{\Sigma}_{o,o}^{-1}\mathbf{X}_o$. The BLUP for $\mathbf{b}_a'\mathbf{y}_a$ is then

$$\widehat{\mathbf{b}_a'\mathbf{y}_a} = \boldsymbol{\lambda}_o'\mathbf{y}_o, \tag{11}$$

which is equivalent to

$$\mathbf{b}_o'\mathbf{y}_o + \mathbf{b}_u'\hat{\mathbf{y}}_u,$$

where $\hat{\mathbf{y}}_u = \mathbf{\Sigma}_{o,s}\Sigma_{o,o}^{-1}(\mathbf{y}_o - \hat{\boldsymbol{\mu}}_o) + \hat{\boldsymbol{\mu}}_u$ with $\hat{\boldsymbol{\mu}}_o = \mathbf{X}_o\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\mu}}_u = \mathbf{X}_u\hat{\boldsymbol{\beta}}$. $\hat{\boldsymbol{\beta}}$ is the generalized least squares estimator $(\mathbf{X}_o'\mathbf{\Sigma}_{o,o}^{-1}\mathbf{X}_o)^{-1}\mathbf{X}_o'\mathbf{\Sigma}_{o,o}^{-1}\mathbf{y}_o$. We can see then that the predictor multiplies the observed data $\mathbf{y}_o$ with relevant weights from the $\mathbf{b}_o$ vector, and then adds in the kriged predictions $\hat{\mathbf{y}}_u$ multiplied with relevant weights from the $\mathbf{b}_u$ vector.

The prediction variance of the predictor in equation 11 is

$$E((\boldsymbol{\lambda}_o'\mathbf{y}_o - \mathbf{b}_a'\mathbf{y}_a)(\boldsymbol{\lambda}_o'\mathbf{y}_o - \mathbf{b}_a'\mathbf{y}_a)) = \boldsymbol{\lambda}_o'\mathbf{\Sigma}_{o,o}\boldsymbol{\lambda}_o - 2\mathbf{b}_a'\mathbf{\Sigma}_{a,o}\boldsymbol{\lambda}_o + \mathbf{b}_a'\mathbf{\Sigma}_{a,a}\mathbf{b}_a. \tag{12}$$

We call the predictor in equation 11 with $\mathbf{\Sigma}$ in equation 6 the ST-FPBK predictor.

A common predictor of interest is the total abundance in the most current time point of the survey. In this scenario, $\mathbf{b}_a$ is a vector of 1's and 0's, where the k^{th} element of $\mathbf{b}_a$ is equal to 1 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 equal to 0 otherwise. If we order $\mathbf{y}_a$ by (1) the unobserved data points from past surveys, (2) the unobserved data points from the current survey, (3) the observed data points from past surveys, and (4) the observed data points from the current survey, then

$$\mathbf{b}_a = [\mathbf{b}_{up}', \mathbf{b}_{uc}', \mathbf{b}_{op}', \mathbf{b}_{oc}']' = [\mathbf{0}', \mathbf{1}', \mathbf{0}', \mathbf{1}']', \tag{13}$$

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

2.3. *Estimation*

In practical applications, the covariance matrix $\mathbf{\Sigma}$ in equation 6 that is partitioned into the various sub-matrices in equations 11 and 12 needs to be estimated from the observed data $\mathbf{y}_o$. The spatio-temporal model in equation 5 does not have any distributional assumptions: we only need to specify the mean and variance of $\mathbf{y}_o$. Restricted Maximum Likelihood (REML) can be used to estimate the covariance parameters in $\mathbf{\Sigma}$, which we will refer to as $\boldsymbol{\theta} \equiv [\sigma_\delta^2,\ \sigma_\gamma^2,\ \phi,\ \sigma_\tau^2,\ \sigma_\eta^2,\ \rho,\ \sigma_\omega^2,\ \sigma_\nu^2]'$ (Patterson and Thompson 1971; Harville 1977). Even if $\mathbf{y}_a$ is not multivariate normal, the REML estimator for the parameter vector $\boldsymbol{\theta}$ is still unbiased (Heyde 1994; Cressie and Lahiri 1993).

However, REML estimation can be computationally burdensome, particularly for large spatio-temporal data sets with many observed sites and time points. Therefore, we use developments from Dumelle et al. (2021) in the application, the simulations described in the next section, and the accompanying `R` package to speed up estimation of $\boldsymbol{\theta}$.

3. Application

We now apply the ST-FPBK predictor to a moose data set described below. Moose surveys throughout Alaska and Canada are often conducted regularly, making them good candidates for incorporating temporal correlation.

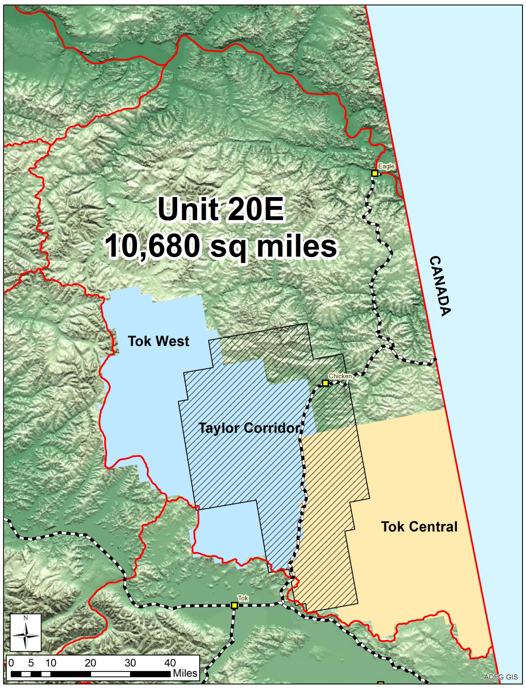

Figure 1. A map of the Taylor Corridor in the east-central region of Alaska.

3.1. *Data Description*

The Taylor Corridor in the east-central region of Alaska is a popular habitat for moose and other wildlife. Abundance surveys for moose are performed in the Taylor Corridor of the Tok region of Alaska annually (Figure 1) so that biologists have an idea about the abundance of moose each year. In particular, surveys were conducted from 2014 through 2020 in every year except 2016, during which there was not sufficient snow cover to perform a survey. The spatial sampling frame for our study area consists of 381 sites. There are a total of 7 unique time points represented in the data, including the missing year of 2016. Therefore, N is 2667.

In each year of the survey, a team of biologists selected some of the 381 sites to survey. The number of sites that were selected varies from a low of 76 in the year 2019 to a high of 90 in the year 2020. Throughout the 7 unique years, some sites were sampled as many as five different times while others were never sampled at all (Figure 2). Figure 2 and all remaining figure graphics are constructed with the `ggplot2` `R` package (Wickham 2016). The number of units sampled throughout all survey years, n, was 487 units.

Before the survey begins in each year, biologists stratified the sites into a "High" stratum and a "Low" stratum composed. The goal of the following analysis is to predict the total abundance of moose across all sites in the year `2020`, the most recent year of the survey, using stratum as a covariate in the spatio-temporal model.

3.2. *Model Fitting*

We fit the product-sum covariance model defined in equation 5 using REML with stratum as a covariate in the design matrix, an exponential spatial correlation structure

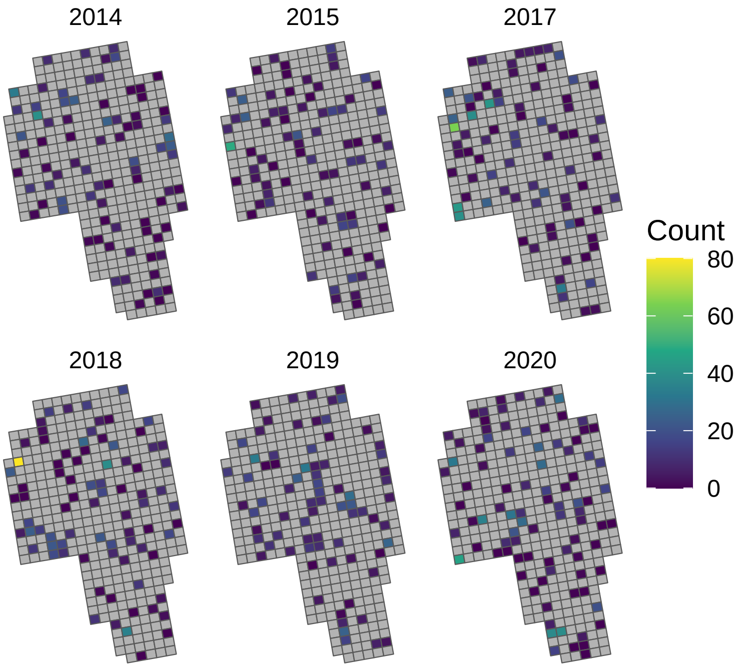

Figure 2. Layout of the spatial sites used to survey moose in the Taylor corridor in eastern-central Alaska, coloured by moose count. Sites coloured grey were not sampled in that year. The year 2016 is excluded because no survey was performed in that year.

defined in equation 3, and an exponential temporal correlation structure defined in equation 4. Table 1 gives the estimated parameters from the model fit.

Table 1. Estimated covariance parameters in the model. $\hat{\sigma}^2_\delta$, $\hat{\sigma}^2_\gamma$, and $\hat{\phi}$ are the spatial dependent error variance, independent error variance, and range parameters, respectively. $\hat{\sigma}^2_\tau$, $\hat{\sigma}^2_\eta$, and $\hat{\rho}$ are the temporal dependent error variance, independent error variance, and range parameters, respectively. $\hat{\sigma}^2_\omega$ and $\hat{\sigma}^2_\nu$ are the spatio-temporal dependent error variance and spatio-temporal independent error variance.

Spatial			Temporal			Spatio-temporal	
$\hat{\sigma}^2_\delta$	$\hat{\sigma}^2_\gamma$	$\hat{\phi}$	$\hat{\sigma}^2_\tau$	$\hat{\sigma}^2_\eta$	$\hat{\rho}$	$\hat{\sigma}^2_\omega$	$\hat{\sigma}^2_\nu$
16.37	7.78	4.51	0.29	0	3.68	25.53	36.47

To help interpret what some of these fitted covariance parameter estimates mean, we can construct a fitted covariance plot (Figure 3). As the spatial distance between two sites increases (dark colour to light colour), the covariance of two errors decreases to 0, with the $\hat{\phi}$ parameter estimate controlling the rate of decay. In fact, the model estimates the covariance to be nearly 0 when two sites are 20 or more kilometers apart, no matter what the temporal distance is. The covariance between two errors that are six years apart is still estimated to be positive if the two errors come from the same site or from adjacent sites.

The estimated vector of fixed effects, using "High" as the reference group, is $\hat{\boldsymbol{\beta}} =$ (9.62, -4.55). Therefore, the overall mean for sites in the "High" stratum is estimated to be 9.62 moose while the overall mean for sites in the "Low" stratum is estimated to be 5.07 moose.

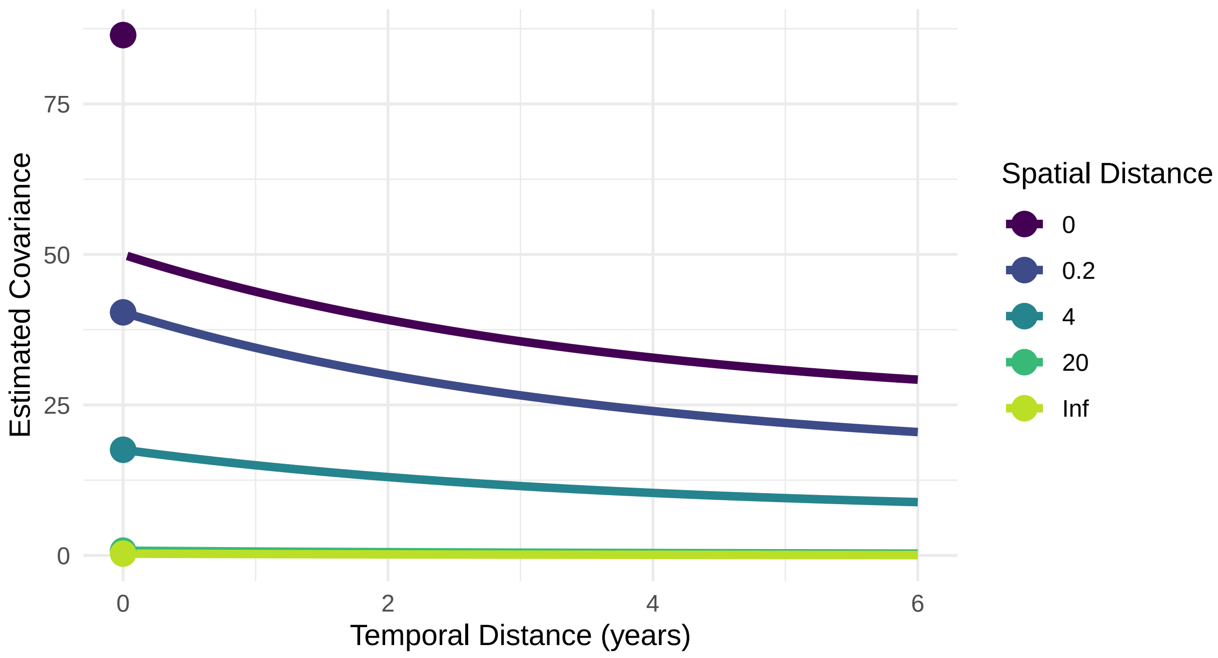

Figure 3. Estimated covariance of the errors from the estimated parameters in a spatio-temporal product-sum model. Distance between two sites is calculated from the site centroids; the centroids of two sites directly adjacent to one another are about 4 spatial kilometers apart.

3.3. *Prediction*

We now use the fitted spatio-temporal model with the BLUP from equation 11 and weights given in equation 13 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 11 and 12 and letting elements of $\mathbf{b}_a$ be equal to 1 for data points in 2020 and equal to 0 otherwise, we obtain a prediction of 3001 moose and a standard error (the square root of the prediction variance) of 217 moose. A 90% normal-based prediction interval for the total abundance in 2020 is (2644, 3357) moose. Note that, though the response in this example is a count, a normal-based prediction interval for the total is still appropriate through an application of the central limit theorem for dependent data (Smith 1980). 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. 2021) to compute the spatial FPBK prediction (Ver Hoef 2008) for the total abundance of moose in the year 2020 with stratum as a covariate. The spatial FPBK predictor is what is currently implemented in the widely used GSPE software for moose surveys (DeLong 2006).

We also use the stratified random sampling design-based estimator

$$\sum_{i=1}^{2} N_i \cdot \bar{y}_i$$

where $\bar{y}_i$ is the sample mean for the observed data in 2020 in the i^{th} stratum and N_i is the total number of sites in 2020 in the i^{th} stratum. The stratified random sampling

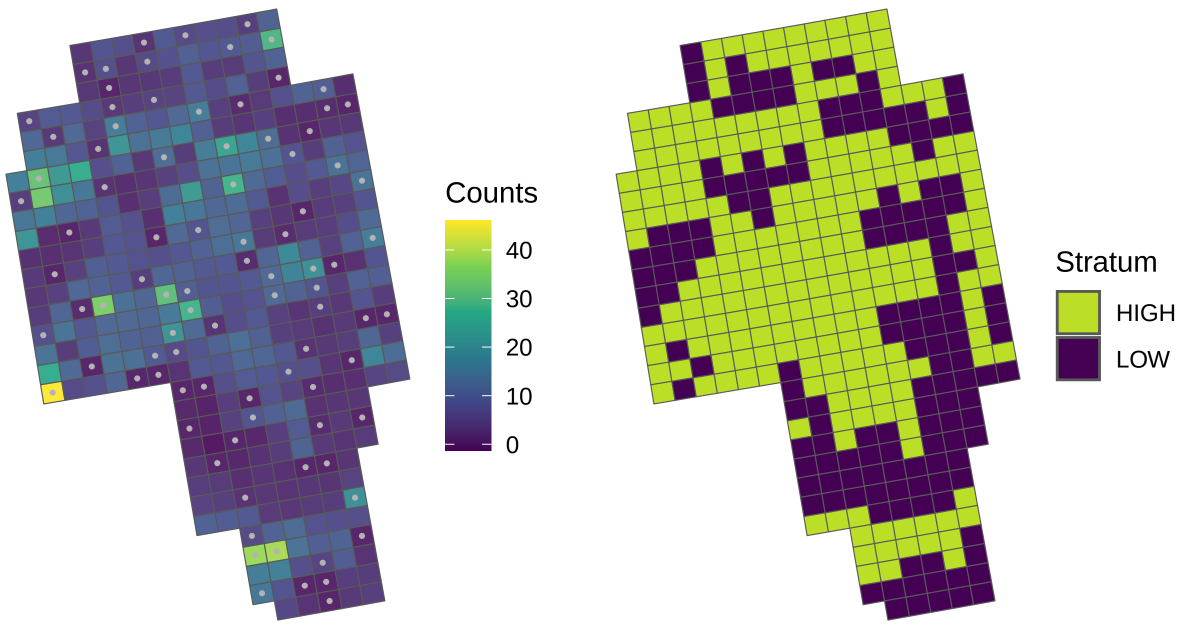

Figure 4. (Left). A map of the predictions for the sites in the year 2020 from the spatio-temporal model. A site with a grey dot in the center means that the site was sampled in 2020. (Right). A map of the stratification of sites in 2020. NOTE: need to update these to be proper sub-figures if we keep the stratification figure.

design-based estimator has a variance for the total abundance of

$$\sum_{i=1}^{2} N_i^2 \cdot (1 - \frac{n_i}{N_i}) \cdot \frac{s_i^2}{n_i},$$

where s_i^2 is the sample variance of the observed data points in 2020 in the i^{th} stratum and n_i is the number of observed data points in 2020 in the i^{th} stratum. Both the purely spatial model fit with `sptotal` and the stratified random sampling design-based estimator use data only 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 stratified random sampling design-based estimator, the estimated total number of moose in 2020 in the region is 2853 moose with a standard error of 371 moose. While the predictions for the total moose abundance are similar across the three methods, we see that the spatio-temporal model is most efficient ($SE = 217$ moose compared to 319 moose for the purely spatial model that ignores previous surveys and 371 moose for the stratified random sampling design-based estimator that ignores both previous surveys and spatial correlation in the current survey).

In addition to making a prediction for the abundance in the most recent survey, we can also use the spatio-temporal model to backcast predictions for the abundance in past survey years, interpolate predictions for years during which a survey was not completed, and forecast predictions for future years. For example, in the Taylor Corridor surveys, there was no survey conducted in the year 2016 because of insufficient snow cover. Leveraging the temporal structure of the st-FPBK predictor, we can still construct a prediction and corresponding standard error though, as expected, this standard error is larger than the standard errors of years where a survey was com-

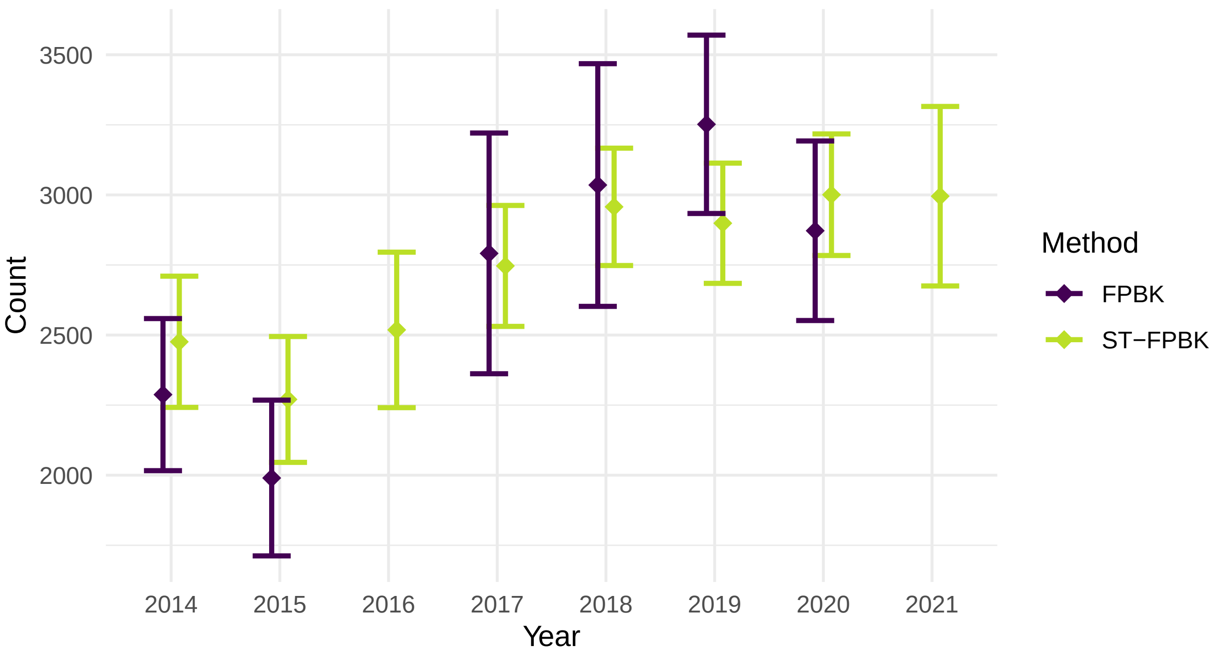

Figure 5. Moose abundance predictions for the Taylor Corridor from 2014 through 2021 with the spatial FPBK predictor and the ST-FPBK predictor. Predictions are given with a diamond symbol; the bars surrounding each prediction are standard error bars. Because surveys were not conducted in 2016 and 2021, there is no spatial FPBK predictor and the standard errors for the ST-FPBK predictor for those years is larger than the standard errors in the other years.

pleted (Figure 5). Also, in Figure 5, we see a forecasted prediction and corresponding standard error for the abundance in 2021. Again, the standard error associated with the forecasted prediction is larger than the standard errors for years with surveys.

4. Simulation

4.1. *Description*

To evaluate performance of the ST-FPBK predictor, we conduct a simulation study. We simulate a response vector $\mathbf{y}$ of length $N = 1000$ on a 10×10 grid of 100 spatial sites on the unit square ($[0, 1] \times [0, 1]$) and 10 equally-spaced time points in the interval $[0, 1]$, so that each spatial site has a response value at each time point. $\mathbf{y}$ is multivariate normal with mean $\mathbf{0}$ and product-sum covariance matrix $\boldsymbol{\Sigma}$ defined in equation 6 with the covariance parameters given in Table 2.

Table 2. Covariance parameters used to simulate data. σ^2_δ, σ^2_γ, and ϕ are the spatial dependent error variance, independent error variance, and range parameters, respectively. σ^2_τ, σ^2_η, and ρ are the temporal dependent error variance, independent error variance, and range parameters, respectively. σ^2_ω and σ^2_ν are the spatio-temporal dependent error variance and spatio-temporal independent error variance. Note that both ϕ (and ρ) appear in $\mathbf{R}_{st}$; therefore, their values can change the underlying covariance even when σ^2_δ (and σ^2_τ) are equal to 0.

	Spatial			Temporal			Spatio-temporal	
scenario	σ^2_δ	σ^2_γ	ϕ	σ^2_τ	σ^2_η	ρ	σ^2_ω	σ^2_ν
all-dev	0.5	0.17	0.47	0.5	0.17	0.33	0.50	0.17
t-iev	0	0	0.47	0	1.50	0	0.25	0.25
spt-iev	0	0	0	0	0	0	0	2.00

The three scenarios in the table correspond to (1) **all-dev**: a scenario where a substantial proportion of the overall variance comes from the spatial, temporal, and spatio-temporal dependent error variance parameters $\sigma^2_\delta, \sigma^2_\tau$, and σ^2_ω; (2) **t-iev**: a scenario where there the overall variance is dominated by the temporal independent error variance parameter, σ^2_η; and (3) **spt-iev**: a scenario where all of the variability comes from σ^2_ν so that errors are independent regardless of spatial and time indices. In all scenarios, summing all six variance parameters gives a total variance equal to two.

Both $\mathbf{R}_s$ and $\mathbf{R}_t$ are generated from the exponential correlation function with ϕ and ρ as the range parameters in equations 3 and 4. The values 0.471 and 0.3333 are chosen for ϕ and ρ, respectively, so that the effective ranges, 3ϕ and 3ρ, are equal to the maximum distance between two data points in space ($\sqrt{2} = 1.414$) and the maximum distance between two data points in time (1). A value of 0 for ϕ (or ρ) sets the $\mathbf{R}_s$ (or the $\mathbf{R}_t$) matrix to the identity matrix. Figure 6 shows the model covariance of the errors used to generate data for the “all-dev” scenario.

Each of these three scenarios is replicated for two different sample sizes: $n = 250$ and $n = 500$. A simple random sample as chosen from the 1000 total data points.

Finally, the simulation experiment is repeated for a skewed response variable. To create the skewed response variable, a normally-distributed response is simulated according to the parameters given in Table 2, except that each of the variance parameters (not including ϕ and ρ) is divided by 2.89 so that the total variance is equal to 0.6931. This variable is then exponentiated so that the total variance after exponentiation is equal to 2. Note that, not only does exponentiation result in a right-skewed response variable, but exponentiating also allows for an assessment of how the ST-FPBK predictor performs when the covariance is mis-specified, as the resulting response variable is now simulated with an intractable covariance function that is not used in the model fitting.

Therefore, the simulation study has 12 total settings coming from a $3\times2\times2$ (scenario $\times$ sample size $\times$ distribution shape) factorial design. For each setting, we simulate 1000 realizations of the response vector $\mathbf{y}$. For each realization, we use three methods to predict the total response for the “most current” time point, which is when the time index is equal to 1 on the interval $[0, 1]$). We will henceforth call this “total response for the most current time point quantity” the “current total.”

The first method uses the ST-FPBK predictor in equation 11 with the spatio-temporal model covariance in equation 6. REML estimation with the observed data $\mathbf{y}_o$ is used to obtain estimates for the covariance parameter vector $\boldsymbol{\theta}$. The second method is the FPBK spatial model fit with the `sptotal` `R` package (Higham et al. 2021) that only uses data from the most current time point.

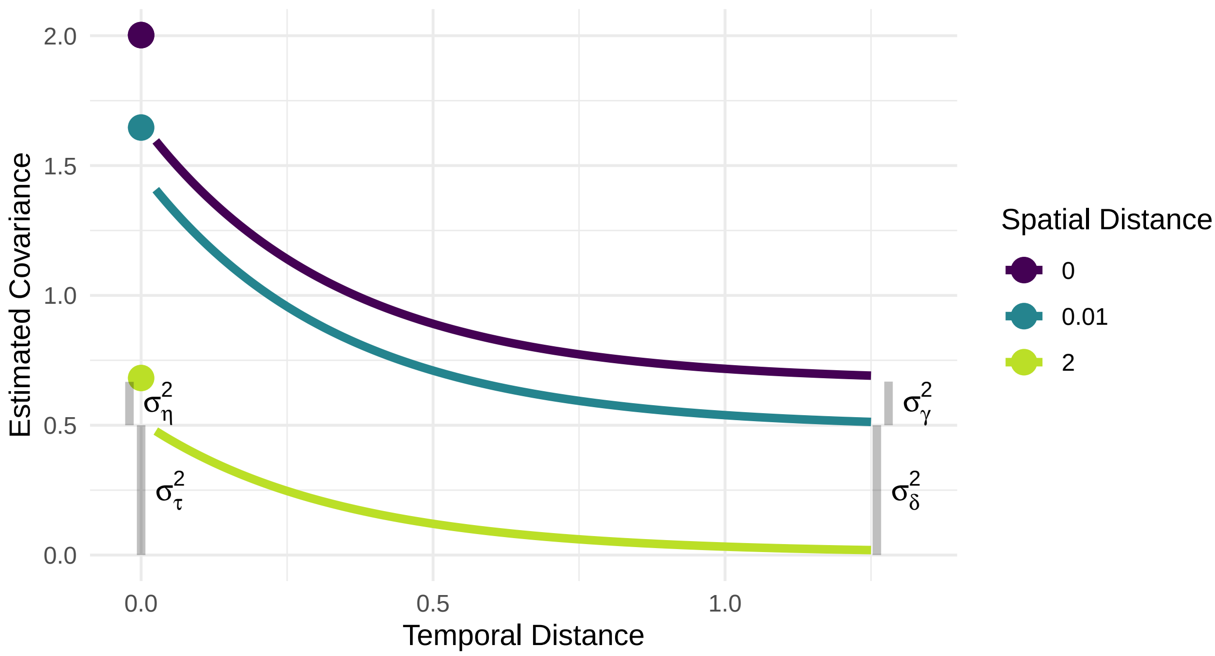

Figure 6. The model covariance used in the simulations for the spatio-temporal scenario. Covariance is approximately 0 for errors from data points that are $\sqrt{2}$ distance units apart in space and 1 distance unit apart in time. The spatial dependent error variance (σ_δ^2), spatial independent error variance (σ_γ^2), temporal dependent error variance (σ_τ^2), and temporal independent error variance (σ_η^2) are shown with grey lines.

The third method uses a simple random sample (SRS) design-based estimator with data from the most current time point. The SRS design-based estimator for the total is $100 \cdot \bar{y}$, where $\bar{y}$ is the sample mean of the response in the most current time point. The variance of the estimator (Lohr 2021) is $100^2 \cdot \frac{s^2}{n_1} \cdot (1 - \frac{n_1}{100})$, where s^2 is the sample variance of the response variable in the most current time point and n_1 is the number of sampled locations in the most current time point.

The SRS method gives an estimator, not a predictor, and a corresponding confidence interval, not a prediction interval, because the SRS design-based estimator treats the observed data as fixed, not as a random realization from a process (Brus 2021; Dumelle et al. 2022). However, in the remaining text and tables, we refer to the "current total" response quantity obtained from the three methods as a "prediction" and to the corresponding interval as a "prediction interval" to limit unnecessarily verbose text and tables.

For each method, we calculate the root-mean-squared-prediction-error (rMSPE) as $\sqrt{\frac{1}{1000}(\sum_{i=1}^{1000}(T_i - \hat{T}_i)^2)}$, where T_i and $\hat{T}_i$ are the realized and predicted current totals, respectively, in the i^{th} iteration. Bias is recorded as $\frac{1}{1000}\sum_{i=1}^{1000}(T_i - \hat{T}_i)$. We also create a normal-based 90% prediction interval for the realized current total and record $\frac{1}{1000}\sum_{i=1}^{1000} I(LB_i < T_i < UB_i)$, where $I(LB_i < T_i < UB_i)$ is an indicator variable that is equal to 1 if the realized total in iteration i, T_i, is between the lower bound, LB_i, and the upper bound, UB_i, of the i^{th} prediction interval.

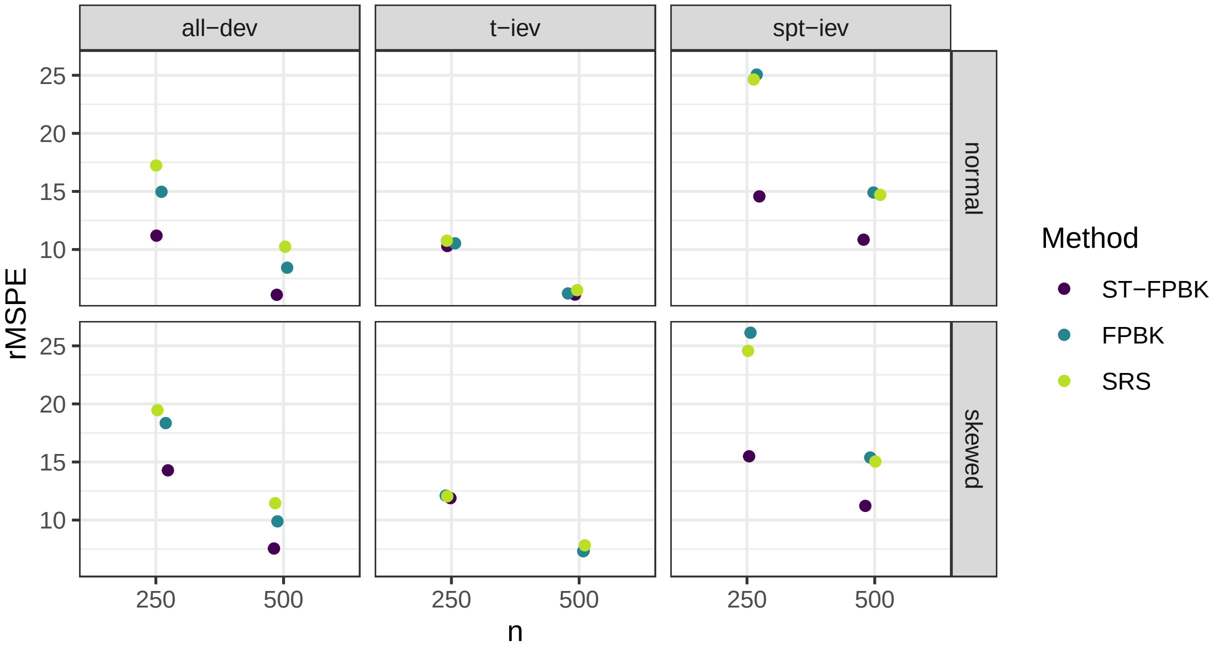

Figure 7. root-mean-squared-prediction-error (rMSPE) for all simulation settings. The ST-FPBK predictor has the smallest rMSPE in all settings tested, though it is similar to the rMSPE of the other two methods in the t-iev scenario.

4.2. *Results*

Tables A1, A2, and A3 in the Appendix give the rMSPE, bias, and interval coverage of the three methods in all 12 simulation settings. In Figure 7, we see that the ST-FPBK predictor outperforms both the purely spatial FPBK predictor and the simple random sample design-based estimator in all of the “all-dev” and “spt-iev” scenarios. In general, rMSPE improvement is larger for the smaller sample size.

We see little gains in rMSPE for the ST-FPBK predictor in the “t-iev” scenario. This setting was chosen to explore how the spatio-temporal model would perform when most of the variability in the response comes from σ_η^2, which allows for data collected in different time points to be uncorrelated, and, for different time points to have very different realized totals. As expected, the ST-FPBK predictor performs no better than a purely spatial model or the SRS design-based estimator for this scenario; however, we can also say that the added complexity of the spatio-temporal model is not detrimental.

All methods appear relatively unbiased in all simulation settings: Table A2 shows that the bias of each method is small compared to the squares of the rMSPE values given in Table A1.

Figure 8 shows the interval coverage for the normal-based prediction intervals (Smith 1980), where the nominal level is 0.90. We see that the ST-FPBK predictor for the current total has approximate 90% coverage in all settings tested. The spatial model and the SRS design-based estimator have lower than nominal coverage in some settings because of the small sample size used (recall that the $n = 250$ observed samples span 10 unique time points so that, on average, the spatial model and SRS design-based estimator only have 25 observed responses to use in the current time point).

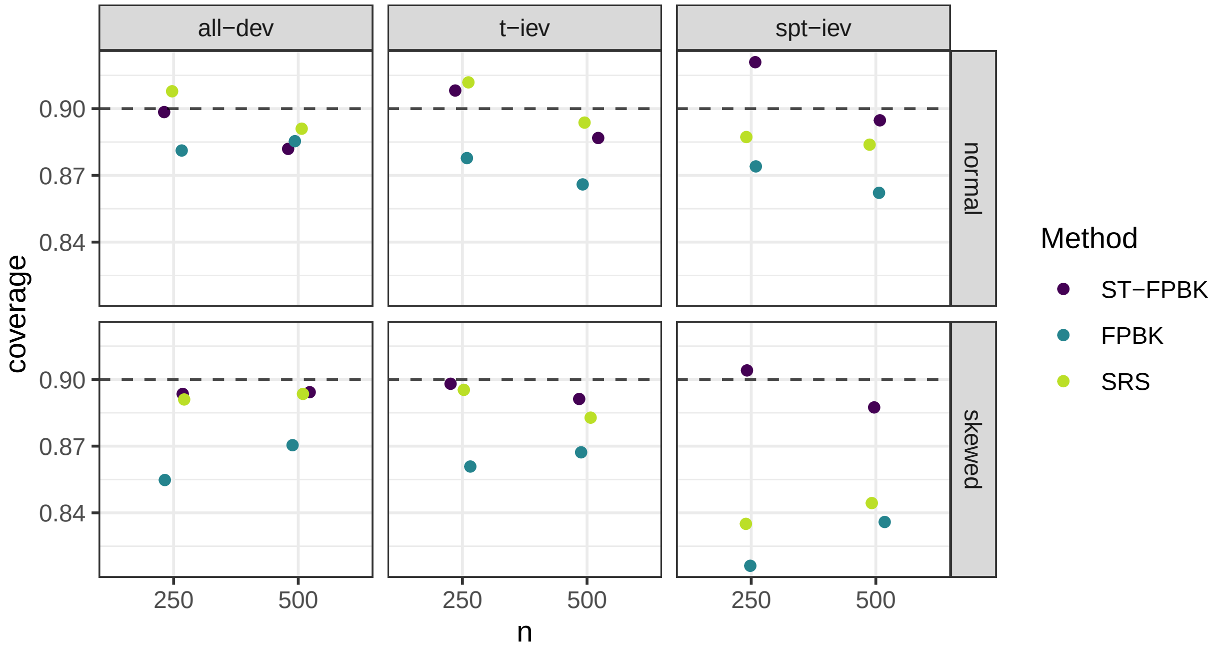

Figure 8. Prediction interval coverage for all simulation settings, where the prediction intervals are normal-based and the nominal level is 0.90. The ST-FPBK predictor has close to appropriate coverage in all settings tested.

5. Discussion

We see in the moose application in Section 3 that there is substantial reduction in the standard error of the predictor for the total moose abundance in 2020 when incorporating data from surveys in previous years. In the simulation study in 4, we find that the ST-FPBK predictor has lower rMSPE than the FPBK predictor from a purely spatial model and an SRS design-based estimator in many settings. The ST-FPBK predictor is less beneficial when the temporal independent error variance contributes a large proportion to the overall variance. Additionally, the ST-FPBK predictor maintains appropriate interval coverage in all settings tested, even when the covariance for the errors is mis-specified.

An additional possible benefit of using the ST-FPBK predictor compared to a purely spatial FPBK predictor is the potential for forecasting abundance before a survey is completed. In Figure 5, we see the forecasted prediction for abundance in the year 2021. While there is a (presumed) loss in precision by constructing a prediction for a year that has no observed samples, the prediction could still be useful to wildlife managers for decision-making before a survey from that year is completed and analyzed. Constructing a prediction for years or time points at which a survey is not completed can be applied to other contexts as well, including temporal interpolation (e.g., the year 2016 in Figure 5).

The ability to predict the abundance (or other quantity) in time points that were not surveyed also allows biologists to investigate how much efficiency is lost from, for example, sampling every other year instead of every year. These types of surveys are often expensive, so perhaps the drop in efficiency from sampling every other year is worth the cost of completing those surveys annually.

We would also like to give our perception of the benefits and drawbacks of our approach with that of Schmidt et al. (2022), who use a hierarchical Bayesian model with spatial radial basis functions that are estimated per year and with time as a trend component in the fixed effects. The benefits of our approach include a faster fitting time, as there is no need to construct and implement the time-consuming Markov chain Monte Carlo sampler. Therefore, our approach is easier to assess in a simulation study, which would be too time-prohibitive for the Bayesian model. Biometricians could also use simulation with our approach to answer various questions given proposed values of covariance parameters like how much efficiency would drop if a survey was only conducted every other year. We argue that our approach is simpler overall for a practitioner to use and could be integrated more readily with the current GSPE software. Finally, our approach allows for temporal interpolation and forecasting while the estimation of the spatial radial basis functions in Schmidt et al. (2022) for each time point do not allow for inference outside of the time points observed.

The Bayesian approach by Schmidt et al. (2022), however, offers features that would be harder to implement in our approach. Their method allows for incorporation of more levels in the Bayesian hierarchical model, including allowing for imperfect detection of animals from a separate detectability survey. Additionally, the Bayesian hierarchical model can use a Poisson or negative binomial model for the counts. Therefore, an appropriate prediction interval for the response on one particular site could be constructed. On the other hand, for our approach, we rely on the central limit theorem for dependent data to form a prediction interval for the total, which would not apply for a prediction interval for the response on just one site.

We have developed a finite population block kriging predictor for spatio-temporal data, which adjusts the variance of the predictor to be appropriate for sampling from a finite population. The resulting predictor is generally at least as good as the predictor form a purely spatial model, and, is often much better. Monitoring programs that use regularly scheduled surveys should consider incorporating data from past surveys to improve precision in the predictor for the most current survey.

Future work in this area includes developing a frequentist model for which imperfect detection of units through time is incorporated into the predictor or how best to select sites to sample for future surveys given proposed values for the spatio-temporal covariance parameters. Additionally, for moose surveys in particular, updating the GSPE software to include analysis for spatio-temporal data could be useful for practitioners. Though we recognize that doing so would be a substantial undertaking, the `R` package that we provide could be a useful starting point for the integration.

Appendix

Table A1. root-mean-squared-prediction-error (rMSPE) for the ST-FPBK predictor, the FPBK predictor, and the SRS estimator for each of the 12 simulation settings. In all settings, the rMSPE for the ST-FPBK predictor is approximately equal to or lower than the rMSPE for the other two methods.

Simulation Setting			rMSPE		
scenario	n	Response Type	ST-FPBK	FPBK	SRS
spt-iev	250	normal	14.58	25.06	24.64
t-iev	250	normal	10.31	10.53	10.76
all-dev	250	normal	11.18	14.97	17.23
spt-iev	500	normal	10.84	14.91	14.71
t-iev	500	normal	6.12	6.22	6.50
all-dev	500	normal	6.09	8.43	10.24
spt-iev	250	skewed	15.49	26.14	24.56
t-iev	250	skewed	11.89	12.10	12.05
all-dev	250	skewed	14.28	18.35	19.46
spt-iev	500	skewed	11.22	15.38	15.04
t-iev	500	skewed	7.34	7.32	7.82
all-dev	500	skewed	7.55	9.89	11.45

Table A2. Bias (Realized Current Total - Predicted Current Total) for the ST-FPBK predictor, the FPBK predictor, and the SRS estimator for each of the 12 simulation settings. In all settings, all methods appear fairly unbiased.

Simulation Setting			Bias		
scenario	n	Response Type	ST-FPBK	FPBK	SRS
spt-iev	250	normal	0.73	1.38	1.55
t-iev	250	normal	0.44	0.39	0.47
all-dev	250	normal	0.48	0.27	0.45
spt-iev	500	normal	0.46	0.60	0.67
t-iev	500	normal	0.15	0.14	0.07
all-dev	500	normal	0.04	0.07	0.04
spt-iev	250	skewed	0.36	0.56	1.48
t-iev	250	skewed	0.33	0.22	0.41
all-dev	250	skewed	-0.07	-0.85	-0.49
spt-iev	500	skewed	0.32	0.29	0.66
t-iev	500	skewed	0.24	0.15	0.08
all-dev	500	skewed	-0.10	-0.39	-0.37

Table A3. Prediction interval coverage for the ST-FPBK predictor, the FPBK predictor, and the SRS for each of the 12 simulation settings. All intervals are normal-based and have a nominal coverage level of 0.90.

Simulation Setting			Coverage		
scenario	n	Response Type	ST-FPBK	FPBK	SRS
spt-iev	250	normal	0.92	0.87	0.89
t-iev	250	normal	0.91	0.88	0.91
all-dev	250	normal	0.90	0.88	0.91
spt-iev	500	normal	0.90	0.86	0.88
t-iev	500	normal	0.89	0.87	0.89
all-dev	500	normal	0.88	0.89	0.89
spt-iev	250	skewed	0.90	0.82	0.84
t-iev	250	skewed	0.90	0.86	0.90
all-dev	250	skewed	0.89	0.86	0.89
spt-iev	500	skewed	0.89	0.84	0.84
t-iev	500	skewed	0.89	0.87	0.88
all-dev	500	skewed	0.89	0.87	0.89

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.

Brus, Dick J. 2021. "Statistical approaches for spatial sample survey: Persistent misconceptions and new developments." *European Journal of Soil Science* 72 (2): 686–703.

Cressie, Noel. 2015. *Statistics for spatial data - revised edition.* John Wiley & Sons.

Cressie, Noel, and Soumendra Nath Lahiri. 1993. "The asymptotic distribution of REML estimators." *Journal of multivariate analysis* 45 (2): 217–233.

De Cesare, Luigi, DE Myers, and D Posa. 2001. "Product-sum covariance for space-time modeling: an environmental application." *Environmetrics: The official journal of the International Environmetrics Society* 12 (1): 11–23.

De Iaco, S, M Palma, and D Posa. 2015. "Spatio-temporal geostatistical modeling for French fertility predictions." *Spatial Statistics* 14: 546–562.

De Iaco, Sandra, Donald E Myers, and Donato Posa. 2001. "Space–time analysis using a general product–sum model." *Statistics & Probability Letters* 52 (1): 21–28.

DeLong, Robert A. 2006. *Geospatial population estimator software user's guide.* Alaska Department of Fish and Game, Division of Wildlife Conservation.

Dumelle, Michael, Matt Higham, Jay M Ver Hoef, Anthony R Olsen, and Lisa Madsen. 2022. "A comparison of design-based and model-based approaches for finite population spatial sampling and inference." *Methods in Ecology and Evolution* 13 (9): 2018–2029.

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.

Harville, David A. 1977. "Maximum likelihood approaches to variance component estimation and to related problems." *Journal of the American statistical association* 72 (358): 320–338.

Heyde, CC. 1994. "A quasi-likelihood approach to the REML estimating equations." *Statistics & Probability Letters* 21 (5): 381–384.

Higham, Matt, Jay Ver Hoef, Bryce Frank, and Michael Dumelle. 2021. *sptotal: Predicting Totals and Weighted Sums from Spatial Data.* R package version 1.0.0, `https://highamm.github.io/sptotal/index.html`.

Kellie, Kalin A, and Robert A DeLong. 2006. *Geospatial survey operations manual.* Technical Report. Alaska Department of Fish and Game.

Kellie, Kalin Ann, Kassidy E Colson, and Joel H Reynolds. 2019. *Challenges to Monitoring Moose in Alaska.* Alaska Department of Fish and Game, Division of Wildlife Conservation Juneau . . . .

Lemos, Ricardo T, and Bruno Sansó. 2009. "A spatio-temporal model for mean, anomaly, and trend fields of North Atlantic sea surface temperature." *Journal of the American Statistical Association* 104 (485): 5–18.

Lohr, Sharon L. 2021. *Sampling: design and analysis.* Chapman and Hall/CRC.

Martínez-Beneito, Miguel A, Antonio López-Quilez, and Paloma Botella-Rocamora. 2008. "An autoregressive approach to spatio-temporal disease mapping." *Statistics in medicine* 27 (15): 2874–2889.

Patterson, H Desmond, and Robin Thompson. 1971. "Recovery of inter-block information when block sizes are unequal." *Biometrika* 58 (3): 545–554.

Peters, Wibke, Mark Hebblewhite, Kirby G Smith, Shevenell M Webb, Nathan Webb, Mike Russell, Curtis Stambaugh, and Robert B Anderson. 2014. "Contrasting aerial moose population estimation methods and evaluating sightability in west-central Alberta, Canada." *Wildlife Society Bulletin* 38 (3): 639–649.

Sahu, Sujit K, and Dankmar Böhning. 2022. "Bayesian spatio-temporal joint disease mapping of Covid-19 cases and deaths in local authorities of England." *Spatial Statistics* 49: 100519.

Schabenberger, Oliver, and Carol A Gotway. 2017. *Statistical methods for spatial data analysis: Texts in statistical science.* Chapman and Hall/CRC.

Schmidt, Joshua H, Matthew D Cameron, Kyle Joly, Jordan M Pruszenski, Joel H Reynolds, and Mathew S Sorum. 2022. "Bayesian spatial modeling of moose count data: increasing estimator efficiency and exploring ecological hypotheses." *The Journal of Wildlife Management* e22220.

Smith, Tony E. 1980. "A central limit theorem for spatial samples." *Geographical Analysis* 12 (4): 299–324.

Ver Hoef, Jay M. 2008. "Spatial methods for plot-based sampling of wildlife populations." *Environmental and Ecological Statistics* 15 (1): 3–13.

Wickham, Hadley. 2016. "Data analysis." In *ggplot2*, 189–201. Springer.

Wikle, Christopher K, Andrew Zammit-Mangion, and Noel Cressie. 2019. *Spatio-temporal Statistics with R.* Chapman and Hall/CRC.

Xu, Jiaqi, and Hong Shu. 2015. "Spatio-temporal kriging based on the product-sum model: some computational aspects." *Earth Science Informatics* 8 (3): 639–648.