Supplementary Material for "A Linear Mixed Model Formulation for Spatio-Temporal Random Processes with Computational Advances for the Product, Sum, and Product-Sum Covariance Functions"

Michael Dumelle[a,*], Jay M. Ver Hoef[b], Claudio Fuentes[a], Alix Gitelman[a]

[a]*Department of Statistics, Oregon State University, Corvallis, OR 97331*
[b]*Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA Fisheries, Seattle, WA 98115*

Contents

*Corresponding author
Email address: dumellem@oregonstate.edu (Michael Dumelle)

1. Proof of Strict Positive Definiteness of the Sum-With-Error Covariance Function

Myers and Journel (1990) show that even when spatial and temporal covariance functions are strictly positive definite, the sum covariance function may only be positive definite. The sum covariance matrix can be written as

$$\mathbf{\Sigma}_{\text{SUM}} = \sigma_\delta^2 \mathbf{Z}_s \mathbf{R}_s \mathbf{Z}_s' + \sigma_\gamma^2 \mathbf{Z}_s \mathbf{Z}_s' + \sigma_\tau^2 \mathbf{Z}_t \mathbf{R}_t \mathbf{Z}_t' + \sigma_\eta^2 \mathbf{Z}_t \mathbf{Z}_t'.$$

The sum-with-error covariance matrix can be written as

$$\mathbf{\Sigma}_{\text{SWE}} = \sigma_\delta^2 \mathbf{Z}_s \mathbf{R}_s \mathbf{Z}_s' + \sigma_\gamma^2 \mathbf{Z}_s \mathbf{Z}_s' + \sigma_\tau^2 \mathbf{Z}_t \mathbf{R}_t \mathbf{Z}_t' + \sigma_\eta^2 \mathbf{Z}_t \mathbf{Z}_t' + \sigma_\nu^2 \mathbf{I}_{st} = \mathbf{\Sigma}_{\text{SUM}} + \sigma_\nu^2 \mathbf{I}_{st}.,$$

Suppose $\sigma_\nu^2 > 0$. The matrix $\sigma_\nu^2 \mathbf{I}_{st}$ is positive definite because for a sample size n, $\sigma_\nu^2 \mathbf{I}_{st}$'s characteristic polynomial equals $(\sigma_\nu^2 - \lambda)^n$. The roots of this characteristic polynomial all equal σ_ν^2, so the eigenvalues of $\sigma_\nu^2 \mathbf{I}_{st}$ are all positive and $\sigma_\nu^2 \mathbf{I}_{st}$ is positive definite. So then for a vector of real numbers $\mathbf{x}$, it follows that

$$\begin{aligned} \mathbf{x}' \mathbf{\Sigma}_{\text{SWE}} \mathbf{x} &= \mathbf{x}' (\mathbf{\Sigma}_{\text{SUM}} + \sigma_\nu^2 \mathbf{I}_{st}) \mathbf{x} \\ &= \mathbf{x}' \mathbf{\Sigma}_{\text{SUM}} \mathbf{x} + \mathbf{x}' \sigma_\nu^2 \mathbf{I}_{st} \mathbf{x} \\ &> 0 \end{aligned}$$

because $\mathbf{x}' \mathbf{\Sigma}_{\text{SUM}} \mathbf{x} \geqslant 0$ ($\mathbf{\Sigma}_{\text{SUM}}$ is positive semi-definite) and $\mathbf{x}' \sigma_\nu^2 \mathbf{I}_{st} \mathbf{x} > 0$ ($\sigma_\nu^2 \mathbf{I}_{st}$ is positive definite). Thus the sum-with-error covariance matrix is positive definite, implying the sum-with-error covariance function is strictly positive definite.

2. Limiting Behavior of Covariance Functions and Semivariograms in the Spatio-Temporal LMM

Suppose $\text{Cov}(\mathbf{h})$ is a covariance function that depends on a distance $\mathbf{h}$. We define some notation:

$$\text{Cov}(\mathbf{0}) = \text{Cov}(\mathbf{h})|\mathbf{h} = 0, \tag{1}$$

$$\text{Cov}(\mathbf{0}^+) = \lim_{\mathbf{h}\to\mathbf{0}^+} \text{Cov}(\mathbf{h}), \text{ and} \tag{2}$$

$$\text{Cov}(\boldsymbol{\infty}) = \lim_{\mathbf{h}\to\boldsymbol{\infty}} \text{Cov}(\mathbf{h}). \tag{3}$$

We use similar notation for semivariograms, $\gamma(\mathbf{h})$. The covariance matrix of $\mathbf{y}$ in the spatio-temporal LMM, denoted by $\boldsymbol{\Sigma}$, is

$$\boldsymbol{\Sigma} = \sigma_\delta^2 \mathbf{Z}_s \mathbf{R}_s \mathbf{Z}_s' + \sigma_\gamma^2 \mathbf{Z}_s \mathbf{Z}_s' + \sigma_\tau^2 \mathbf{Z}_t \mathbf{R}_t \mathbf{Z}_t' + \sigma_\eta^2 \mathbf{Z}_t \mathbf{Z}_t + \sigma_\omega^2 \mathbf{R}_{st} + \sigma_\nu^2 \mathbf{I}_{st}. \tag{4}$$

Assuming second-order stationarity in space and in time, there is a special relationship between covariance functions and semivariograms:

$$\text{Cov}(\mathbf{h}_s, \text{h}_t) = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{h}_s, \text{h}_t). \tag{5}$$

Using equations (1), (2), (3), and (5), we can derive representations of the variance parameters in equation (4) by evaluating the covariance function and semivariogram in several cases:

$$\begin{array}{lccl}
\text{Cov}(\mathbf{0}, 0) & = & \sigma_\gamma^2 + \sigma_\delta^2 + \sigma_\eta^2 + \sigma_\tau^2 + \sigma_\omega^2 + \sigma_\nu^2 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{0}, 0), \\
\text{Cov}(\mathbf{0}^+, 0) & = & \sigma_\delta^2 + \sigma_\eta^2 + \sigma_\tau^2 + \sigma_\omega^2 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{0}^+, 0), \\
\text{Cov}(\boldsymbol{\infty}, 0) & = & \sigma_\eta^2 + \sigma_\tau^2 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma((\boldsymbol{\infty}, 0), \\
\text{Cov}(\mathbf{0}, 0^+) & = & \sigma_\gamma^2 + \sigma_\delta^2 + \sigma_\tau^2 + \sigma_\omega^2 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{0}, 0^+), \\
\text{Cov}(\mathbf{0}, \infty) & = & \sigma_\gamma^2 + \sigma_\delta^2 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{0}, \infty), \\
\text{Cov}(\mathbf{0}^+, 0^+) & = & \sigma_\delta^2 + \sigma_\tau^2 + \sigma_\omega^2 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{0}^+, 0^+), \\
\text{Cov}(\boldsymbol{\infty}, 0^+) & = & \sigma_\tau^2 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\boldsymbol{\infty}, 0^+), \\
\text{Cov}(\mathbf{0}^+, \infty) & = & \sigma_\delta^2 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{0}^+, \infty), \text{ and} \\
\text{Cov}(\boldsymbol{\infty}, \infty) & = & 0 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\boldsymbol{\infty}, \infty).
\end{array} \tag{6}$$

There are multiple ways to solve for each variance parameter in (6). The spatial-only variance parameters, σ_δ^2 and σ_γ^2, and the temporal-only variance parameters, σ_τ^2 and σ_η^2, can each be represented by a linear combination of no more than two covariances functions or semivariograms:

$$\begin{array}{lccl}
\text{Cov}(\mathbf{0}^+, \infty) & = & \sigma_\delta^2 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{0}^+, \infty), \\
\text{Cov}(\mathbf{0}, \infty) - \text{Cov}(\mathbf{0}^+, \infty) & = & \sigma_\gamma^2 & = \gamma(\mathbf{0}^+, \infty) - \gamma(\mathbf{0}, \infty), \\
\text{Cov}(\boldsymbol{\infty}, 0^+) & = & \sigma_\tau^2 & = \gamma(\boldsymbol{\infty}, \infty) - \gamma(\boldsymbol{\infty}, 0^+), \text{ and} \\
\text{Cov}(\infty, \mathbf{0}) - \text{Cov}(\boldsymbol{\infty}, 0^+) & = & \sigma_\eta^2 & = \gamma(\boldsymbol{\infty}, 0^+) - \gamma(\infty, \mathbf{0}).
\end{array}$$

These variance parameters are identified in Figure 1 for a product-sum LMM with spherical spatial and temporal covariance functions. The variance parameters σ_ω^2 and σ_ν^2 can each be represented

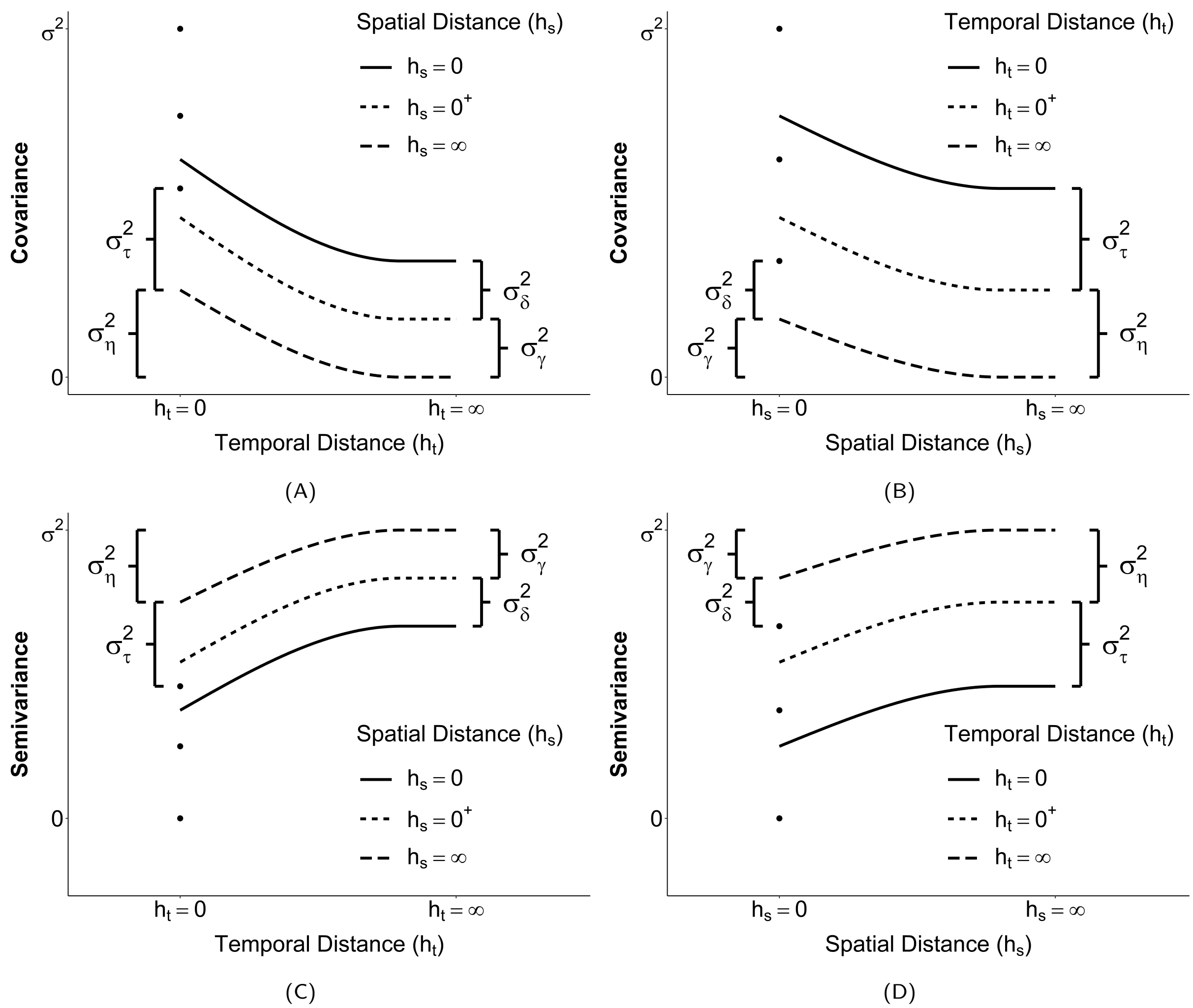

Figure 1: Covariance function and semivariogram behavior in the spatio-temporal LMM. In (A), the covariance function is viewed with temporal distance on the x-axis and spatial distance using line types. In (B), the covariance function is viewed with spatial distance on the x-axis and temporal distance using line types. In (C), the semivariogram is viewed with temporal distance on the x-axis and spatial distance using line types. In (D), the semivariogram is viewed with spatial distance on the x-axis and temporal distance using line types. Distances of 0^+ indicates a right limit as the distance approaches zero. The variance parameters σ_δ^2 (spatial dependent variance), σ_γ^2 (spatial independent variance), σ_τ^2 (temporal dependent variance, and σ_η^2 (temporal independent variance) are identified using brackets. The σ^2 parameter denotes the sum of all variance components (the overall variance).

by a linear combination of more than two covariance functions or semivariograms:

$$\begin{aligned}
\sigma_\omega^2 &= \mathrm{Cov}(\mathbf{0}^+, 0^+) - \mathrm{Cov}(\boldsymbol{\infty}, 0^+) - \mathrm{Cov}(\mathbf{0}^+, \infty) \\
&= \gamma(\boldsymbol{\infty}, 0^+) + \gamma(\mathbf{0}^+, \infty) - \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{0}^+, 0^+), \text{ and} \\
\sigma_\nu^2 &= \mathrm{Cov}(\mathbf{0}, 0) - [\mathrm{Cov}(\mathbf{0}^+, 0^+) + \mathrm{Cov}(\boldsymbol{\infty}, 0) + \mathrm{Cov}(\mathbf{0}, \infty) - \mathrm{Cov}(\boldsymbol{\infty}, 0^+) - \mathrm{Cov}(\mathbf{0}^+, \infty)] \\
&= \gamma(\mathbf{0}^+, 0^+) + \gamma(\mathbf{0}, \infty) + \gamma(\boldsymbol{\infty}, 0) - [\gamma(\mathbf{0}, 0) + \gamma(\boldsymbol{\infty}, 0^+) + \gamma(\mathbf{0}^+, \infty)] \\
&= \gamma(\mathbf{0}^+, 0^+) + \gamma(\mathbf{0}, \infty) + \gamma(\boldsymbol{\infty}, 0) - [\gamma(\boldsymbol{\infty}, 0^+) + \gamma(\mathbf{0}^+, \infty)].
\end{aligned}$$

There is a simpler representation for σ_ω^2 after solving for σ_δ^2 and σ_τ^2:

$$\begin{aligned}
\sigma_\omega^2 &= \mathrm{Cov}(\mathbf{0}^+, 0^+) - (\sigma_\tau^2 + \sigma_\delta^2) \\
&= \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{0}^+, 0^+) - (\sigma_\tau^2 + \sigma_\delta^2).
\end{aligned}$$

Similarly, there is a simpler representation for σ_ν^2 after solving for σ_ω^2, σ_δ^2, σ_γ^2, σ_τ^2, and σ_η^2:

$$\begin{aligned}
\sigma_\nu^2 &= \mathrm{Cov}(\mathbf{0}, 0) - (\sigma_\delta^2 + \sigma_\gamma^2 + \sigma_\tau^2 + \sigma_\eta^2 + \sigma_\omega^2) \\
&= \gamma(\boldsymbol{\infty}, \infty) - \gamma(\mathbf{0}, 0) - (\sigma_\delta^2 + \sigma_\gamma^2 + \sigma_\tau^2 + \sigma_\eta^2 + \sigma_\omega^2) \\
&= \gamma(\boldsymbol{\infty}, \infty) - (\sigma_\delta^2 + \sigma_\gamma^2 + \sigma_\tau^2 + \sigma_\eta^2 + \sigma_\omega^2).
\end{aligned}$$

3. Efficient Log Determinant Computation

3.1. Log Determinant Computations When $\{(\mathbf{s}_i, \mathrm{t}_j)\} = \mathbb{S} \times \mathbb{T}$

3.1.1. The Product LMM

The product (separable) LMM covariance matrix is

$$\mathbf{\Sigma} = \sigma_\omega^2(\mathbf{R}_t^* \otimes \mathbf{R}_s^*), \tag{7}$$

where $\mathbf{R}_t^* = (1 - \pi_t)\mathbf{R}_t + \pi_t\mathbf{I}_t$, $\mathbf{R}_s^* = (1 - \pi_s)\mathbf{R}_s + \pi_s\mathbf{I}_s$, and $0 \leqslant \pi_t, \pi_s \leqslant 1$. The log determinant of $\mathbf{\Sigma}$ in equation (7) is

$$\ln|\mathbf{\Sigma}| = ST\ln(\sigma_\omega^2) + S\ln|\mathbf{R}_t^*| + T\ln|\mathbf{R}_s^*|.$$

3.1.2. The Sum-With-Error LMM

The sum-with-error LMM covariance matrix is

$$\mathbf{\Sigma} = \sigma_\delta^2\mathbf{Z}_s\mathbf{R}_s\mathbf{Z}_s' + \sigma_\gamma^2\mathbf{Z}_s\mathbf{Z}_s' + \sigma_\tau^2\mathbf{Z}_t\mathbf{R}_t\mathbf{Z}_t' + \sigma_\eta^2\mathbf{Z}_t\mathbf{Z}_t' + \sigma_\nu^2\mathbf{I}_{st}.$$

This can be rewritten using condensed notation as

$$\mathbf{\Sigma} = \mathbf{Z}_s\mathbf{\Sigma}_s\mathbf{Z}_s' + \mathbf{Z}_t\mathbf{\Sigma}_t\mathbf{Z}_t' + \sigma_\nu^2\mathbf{I}_{st},$$

where $\mathbf{\Sigma}_s = \sigma_\delta^2\mathbf{R}_s + \sigma_\gamma^2\mathbf{I}_s$, and $\mathbf{\Sigma}_t = \sigma_\tau^2\mathbf{R}_t + \sigma_\eta^2\mathbf{I}_t$. Our algorithm uses two applications of the Sherman-Morrison-Woodbury formula (Sherman, 1949; Sherman and Morrison, 1950; Woodbury, 1950). First, the log determinant of $(\mathbf{Z}_t\mathbf{\Sigma}_t\mathbf{Z}_t' + \sigma_\nu^2\mathbf{I}_{st})$, denoted by $\mathrm{SMW}_{\mathrm{LD}}(\ln|\sigma_\nu^2\mathbf{I}_{st}|, \mathbf{\Sigma}_t, \mathbf{Z}_t)$, can be expressed as

$$\mathrm{SMW}_{\mathrm{LD}}(\ln|\sigma_\nu^2\mathbf{I}_{st}|, \mathbf{\Sigma}_t, \mathbf{Z}_t) = ST\ln(\sigma_\nu^2) + \ln|\mathbf{\Sigma}_t| + \ln|\mathbf{\Sigma}_t^{-1} + \mathbf{Z}_t'\mathbf{Z}_t/\sigma_\nu^2|.$$

A second application of the Sherman-Morrison-Woodbury formula is used to compute the log determinant of $(\mathbf{Z}_s\mathbf{\Sigma}_s\mathbf{Z}_s' + \mathbf{Z}_t\mathbf{\Sigma}_t\mathbf{Z}_t' + \sigma_\nu^2\mathbf{I}_{st})$, which equals the log determinant of $\mathbf{\Sigma}$. The entire log determinant algorithm can be expressed compactly as

$$\ln|\mathbf{\Sigma}| = \mathrm{SMW}_{\mathrm{LD}}(\mathrm{SMW}_{\mathrm{LD}}(\ln|\sigma_\nu^2\mathbf{I}_{st}|, \mathbf{\Sigma}_t, \mathbf{Z}_t), \mathbf{\Sigma}_s, \mathbf{Z}_s).$$

3.1.3. The Product-Sum LMM

The product-sum LMM covariance matrix is

$$\mathbf{\Sigma} = \sigma_\delta^2\mathbf{Z}_s\mathbf{R}_s\mathbf{Z}_s' + \sigma_\gamma^2\mathbf{Z}_s\mathbf{Z}_s' + \sigma_\tau^2\mathbf{Z}_t\mathbf{R}_t\mathbf{Z}_t' + \sigma_\eta^2\mathbf{Z}_t\mathbf{Z}_t' + \sigma_\omega^2\mathbf{R}_t \otimes \mathbf{R}_s + \sigma_\nu^2\mathbf{I}_{st}.$$

This can be rewritten using condensed notation as

$$\mathbf{\Sigma} = \mathbf{Z}_s\mathbf{\Sigma}_s\mathbf{Z}_s' + \mathbf{Z}_t\mathbf{\Sigma}_t\mathbf{Z}_t' + \mathbf{\Sigma}_{st},$$

where $\mathbf{\Sigma}_s = \sigma_\delta^2\mathbf{R}_s + \sigma_\gamma^2\mathbf{I}_s$, $\mathbf{\Sigma}_t = \sigma_\tau^2\mathbf{R}_t + \sigma_\eta^2\mathbf{I}_t$, and $\mathbf{\Sigma}_{st} = \sigma_\omega^2\mathbf{R}_t \otimes \mathbf{R}_s + \sigma_\nu^2\mathbf{I}_{st}$. Our algorithm uses Stegle eigendecompositions (Stegle et al., 2011) and two applications of the Sherman-Morrison-Woodbury formula. Let $\mathbf{U}_s\mathbf{P}_s\mathbf{U}_s'$ be the eigendecomposition of $\mathbf{R}_s$ and $\mathbf{U}_t\mathbf{P}_t\mathbf{U}_t'$ be the eigendecomposition of $\mathbf{R}_t$. Then the log determinant of $\mathbf{\Sigma}_{st}$, denoted by $\mathrm{STE}_{\mathrm{LD}}(\mathbf{\Sigma}_{st})$, can be expressed as the sum of the logarithms of each diagonal element of $\mathbf{V} \equiv \sigma_\omega^2\mathbf{P}_t \otimes \mathbf{P}_s + \sigma_\nu^2\mathbf{I}_t \otimes \mathbf{I}_s$. Using notation, this is expressed as

$$\mathrm{STE}_{\mathrm{LD}}(\mathbf{\Sigma}_{st}) = \sum_i \ln(\mathbf{V}_{ii}).$$

Next the Sherman-Morrison-Woodbury formula is used to compute the log determinant of $(\mathbf{Z}_t\mathbf{\Sigma}_t\mathbf{Z}_t' + \mathbf{\Sigma}_{st})$. This log determinant, denoted by $\mathrm{SMW}_{\mathrm{LD}}(\mathrm{STE}_{\mathrm{LD}}(\mathbf{\Sigma}_{st}), \mathbf{\Sigma}_t, \mathbf{Z}_t)$, can be expressed as

$$\mathrm{SMW}_{\mathrm{LD}}(\mathrm{STE}_{\mathrm{LD}}(\mathbf{\Sigma}_{st}), \mathbf{\Sigma}_t, \mathbf{Z}_t) = \mathrm{STE}_{\mathrm{LD}}(\mathbf{\Sigma}_{st}) + \ln|\mathbf{\Sigma}_t| + \ln|\mathbf{\Sigma}_t^{-1} + \mathbf{Z}_t'\mathbf{\Sigma}_{st}^{-1}\mathbf{Z}_t|.$$

Finally a second application is used to compute the log determinant of $(|\mathbf{Z}_s\mathbf{\Sigma}_s\mathbf{Z}_s' + \mathbf{Z}_t\mathbf{\Sigma}_t\mathbf{Z}_t' + \mathbf{\Sigma}_{st})$, which equals the log determinant of $\mathbf{\Sigma}$. The entire log determinant algorithm can be expressed compactly as

$$\ln|\mathbf{\Sigma}| = \mathrm{SMW}_{\mathrm{LD}}(\mathrm{SMW}_{\mathrm{LD}}(\mathrm{STE}_{\mathrm{LD}}(\mathbf{\Sigma}_{st}), \mathbf{\Sigma}_t, \mathbf{Z}_t), \mathbf{\Sigma}_s, \mathbf{Z}_s).$$

3.2. Log Determinant Computations When $\{(\mathbf{s}_i, \mathrm{t}_j)\} \subset \mathbb{S} \times \mathbb{T}$

Recall the more general form of $\mathbf{y}$ that can always be partitioned by two components: $\mathbf{y}_o$, which contains the observable elements of $\mathbf{y}$, and $\mathbf{y}_u$, which contains the unobservable elements of $\mathbf{y}$. The covariance matrix of $\mathbf{y}$ can be written in block form:

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{oo} & \mathbf{\Sigma}_{ou} \\ \mathbf{\Sigma}_{uo} & \mathbf{\Sigma}_{uu} \end{bmatrix}, \tag{8}$$

where $\mathbf{\Sigma}_{oo} = \mathrm{Cov}(\mathbf{y}_o, \mathbf{y}_u)$, $\mathbf{\Sigma}_{ou} = \mathrm{Cov}(\mathbf{y}_u, \mathbf{y}_o)$, $\mathbf{\Sigma}_{uo} = \mathrm{Cov}(\mathbf{y}_u, \mathbf{y}_o)$, and $\mathbf{\Sigma}_{uu} = \mathrm{Cov}(\mathbf{y}_u, \mathbf{y}_u)$. Equation (8) has the following covariance matrix inverse:

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} \tilde{\mathbf{\Sigma}}_{oo} & \tilde{\mathbf{\Sigma}}_{ou} \\ \tilde{\mathbf{\Sigma}}_{uo} & \tilde{\mathbf{\Sigma}}_{uu} \end{bmatrix},$$

where the dimensions of the blocks in $\mathbf{\Sigma}^{-1}$ match the dimensions of the blocks in $\mathbf{\Sigma}$. Wolf (1978) showed

$$\mathbf{\Sigma}_{oo}^{-1} = \tilde{\mathbf{\Sigma}}_{oo} - \tilde{\mathbf{\Sigma}}_{ou}\tilde{\mathbf{\Sigma}}_{uu}^{-1}\tilde{\mathbf{\Sigma}}_{uo}.$$

Combining this result with the Schur complement of $\mathbf{\Sigma}^{-1}$ yields

$$\begin{aligned} \ln|\mathbf{\Sigma}^{-1}| &= \ln|\tilde{\mathbf{\Sigma}}_{uu}| + \ln|\tilde{\mathbf{\Sigma}}_{oo} - \tilde{\mathbf{\Sigma}}_{ou}\tilde{\mathbf{\Sigma}}_{uu}^{-1}\tilde{\mathbf{\Sigma}}_{uo}| \\ &= \ln|\tilde{\mathbf{\Sigma}}_{uu}| + \ln|\mathbf{\Sigma}_{oo}^{-1}|. \end{aligned}$$

This implies $\ln|\mathbf{\Sigma}_{oo}| = \ln|\mathbf{\Sigma}| + \ln|\tilde{\mathbf{\Sigma}}_{uu}|$, which can be easily found by computing $\ln|\mathbf{\Sigma}|$ using the previous algorithms for when $\{(\mathbf{s}_i, \mathrm{t}_j)\} = \mathbb{S} \times \mathbb{T}$ and then summing with $\ln|\tilde{\mathbf{\Sigma}}_{uu}|$.

4. Further Simulation Study and Data Analysis Details

In the main paper, we omitted mean fixed effect bias results for the simulation study, mean prediction bias for the simulation study and data analysis, and computational details for the additional simulation scenarios (S2, S3, and S4). These results are provided in Tables 1, 2, and 3, respectively. In the simulation study, initial values were perturbed around the true values using a Gaussian random variable with zero mean and variance equal to 10% of the parameter value (on the Gaussian scale), and the Nelder-Mead algorithm was used for optimization. In the data analysis, initial values and the covariance function types (exponential spatial and exponential temporal) were chosen after inspection of the empirical semivariogram (similar to Figures 1C and 1D), and the Nelder-Mead algorithm was used for optimization. In the simulations and the data anlaysis, empirical semivariograms were computed using 16 spatial distance bins (with a cutoff equaling half the maximum spatial distance) and 16 temporal distance bins (with a cutoff equaling half the maximum temporal distance), for a total of 256 spatio-temporal distance bins. The number of spatial and temporal distance classes were chosen to closely resemble the default arguments in **gstat**'s `variogramST()` function (Pebesma, 2004; Pebesma and Heuvelink, 2016).

Table 1: Mean bias of $\hat{\beta}_s, \hat{\beta}_t$, and $\hat{\beta}_{st}$ for all model and estimation method combinations ($\text{Model}_{\text{Method}}$) in all four simulation scenarios (S1, S2, S3, S4).

	S1			S2			S3			S4		
$\text{Model}_{\text{Method}}$	$\hat{\beta}_s$	$\hat{\beta}_t$	$\hat{\beta}_{st}$	$\hat{\beta}_s$	$\hat{\beta}_t$	$\hat{\beta}_{st}$	$\hat{\beta}_s$	$\hat{\beta}_t$	$\hat{\beta}_{st}$	$\hat{\beta}_s$	$\hat{\beta}_t$	$\hat{\beta}_{st}$
P_{REML}	.003	-.006	-.001	-.011	.004	-.002	.006	.000	-.001	-.006	-.004	-.001
P_{CWLS}	-.001	-.007	-.000	-.005	.000	-.002	.004	.001	-.001	-.008	-.002	-.001
SWE_{REML}	.006	-.006	.001	-.009	.003	-.001	-.002	-.001	-.001	-.004	-.004	-.001
SWE_{CWLS}	.006	-.004	.001	-.007	.007	-.001	-.002	-.002	-.001	-.008	-.003	.000
PS_{REML}	-.003	.007	.000	-.008	.004	-.001	.001	.002	-.001	-.007	-.003	-.001
PS_{CWLS}	-.001	.006	.001	-.008	.004	-.001	.004	-.001	-.001	-.005	-.002	-.005
IRE_{OLS}	.005	-.014	.001	.000	.012	-.001	.001	.010	-.003	.005	-.014	-.001

Table 2: Mean bias of $\hat{\beta}_s, \hat{\beta}_t$, and $\hat{\beta}_{st}$ for all model and estimation method combinations ($\text{Model}_{\text{Method}}$) in all four simulation scenarios (S1, S2, S3, S4) and the temperature data analysis (TDA).

$\text{Model}_{\text{Method}}$	S1	S2	S3	S4	TDA
P_{REML}	.044	-.070	-.061	-.008	-.173
P_{CWLS}	.053	-.058	-.057	-.029	.058
SWE_{REML}	.090	-.067	-.103	-.016	.432
SWE_{CWLS}	.093	-.063	-.106	-.018	-.358
PS_{REML}	.055	-.067	-.056	-.016	.445
PS_{CWLS}	.057	-.068	-.058	-.010	-.358
IRE_{OLS}	.044	-.086	-.054	.011	-.667

Table 3: The average empirical semivariogram calculation seconds (SVS), average covariance parameter estimation seconds (ES), average total empirical semivariogram calculation and covariance parameter estimation seconds (TS; the sum of SVS and ES), and average REML iterations (RI) for all models and estimation methods ($\mathrm{Model}_{\mathrm{Method}}$) in S2, S3, and S4.

	S2				S3				S4			
$\mathrm{Model}_{\mathrm{Method}}$	SVS	ES	TS	RI	SVS	ES	TS	RI	SVS	ES	TS	RI
$\mathrm{P}_{\mathrm{REML}}$	NA	4.94	4.94	85.32	NA	4.86	4.86	75.59.	NA	6.70	6.70	112.23
$\mathrm{P}_{\mathrm{CWLS}}$	0.32	0.07	0.39	NA	0.32	0.08	0.40	NA	0.32	0.06	0.38	NA
$\mathrm{SWE}_{\mathrm{REML}}$	NA	4.53	4.53	56.32	NA	5.45	5.45	65.81	NA	8.57	8.57	104.66
$\mathrm{SWE}_{\mathrm{CWLS}}$	0.32	0.14	0.46	NA	0.32	0.18	0.50	NA	0.32	0.15	0.47	NA
$\mathrm{PS}_{\mathrm{REML}}$	NA	14.09	14.09	75.60	NA	25.86	25.86	128.97	NA	27.63	27.63	143.53
$\mathrm{PS}_{\mathrm{CWLS}}$	0.32	.18	0.50	NA	0.32	0.24	0.56	NA	0.32	0.16	0.48	NA
$\mathrm{IRE}_{\mathrm{OLS}}$	NA	0.01	0.01	NA	NA	0.01	0.01	NA	NA	0.01	0.01	NA

References

Myers, D.E., Journel, A., 1990. Variograms with zonal anisotropies and noninvertible kriging systems. Mathematical Geology 22, 779–785.

Pebesma, E., Heuvelink, G., 2016. Spatio-temporal interpolation using gstat. RFID Journal 8, 204–218.

Pebesma, E.J., 2004. Multivariable geostatistics in s: the gstat package. Computers & Geosciences 30, 683–691.

Sherman, J., 1949. Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix. The Annals of Mathematical Statistics 20, 621.

Sherman, J., Morrison, W.J., 1950. Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. The Annals of Mathematical Statistics 21, 124–127.

Stegle, O., Lippert, C., Mooij, J.M., Lawrence, N.D., Borgwardt, K., 2011. Efficient inference in matrix-variate gaussian models with iid observation noise, in: Advances in neural information processing systems, pp. 630–638.

Wickham, H., 2016. ggplot2: elegant graphics for data analysis. Springer.

Wolf, H., 1978. The helmert block method, its origin and development, in: Proceedings of the Second International Symposium on Problems Related to the Redefinition of North American Geodetic Networks, pp. 319–326.

Woodbury, M.A., 1950. Inverting modified matrices. Statistical Research Group, Memorandum report No. 42 106.