

Supplementary Material for: Spatio-Temporal Cross-Covariance Functions under the Lagrangian Framework with Multiple Advections

1 Proof of Theorem 2

The derivation below follows the derivation in Schlather (2010). In order to easily follow the derivation, we color coded the terms that change from the left hand side of the equation to the right hand side. Let $\mathbf{T} = (t_1 \mathbf{I}_d \quad -t_2 \mathbf{I}_d)$ and $\mathbf{v}_{ij} = (\mathbf{v}_{ii}^\top \quad \mathbf{v}_{jj}^\top)^\top$. Operating on the exponentials, we have

$$\begin{aligned}
 & (\mathbf{s}_1 - \mathbf{s}_2 - t_1 \mathbf{v}_{ii} + t_2 \mathbf{v}_{jj})^\top (\mathbf{s}_1 - \mathbf{s}_2 - t_1 \mathbf{v}_{ii} + t_2 \mathbf{v}_{jj}) + (\mathbf{v}_{ij} - \mu_{\mathbf{v},ij})^\top \Sigma_{\mathbf{v},ij}^{-1} (\mathbf{v}_{ij} - \mu_{\mathbf{v},ij}) \\
 = & (\mathbf{h} - \mathbf{T} \mathbf{v}_{ij})^\top (\mathbf{h} - \mathbf{T} \mathbf{v}_{ij}) + (\mathbf{v}_{ij} - \mu_{\mathbf{v},ij})^\top \Sigma_{\mathbf{v},ij}^{-1} (\mathbf{v}_{ij} - \mu_{\mathbf{v},ij}) \\
 \Rightarrow & (\mathbf{h} - \mathbf{T} \mathbf{v}_{ij})^\top (\mathbf{h} - \mathbf{T} \mathbf{v}_{ij}) + (\mathbf{v}_{ij} - \mu_{\mathbf{v},ij})^\top \Sigma_{\mathbf{v},ij}^{-1} (\mathbf{v}_{ij} - \mu_{\mathbf{v},ij}) \\
 = & \mathbf{h}^\top \mathbf{h} - \mathbf{h}^\top \mathbf{T} \mathbf{v}_{ij} - \mathbf{v}_{ij}^\top \mathbf{T}^\top \mathbf{h} + \mathbf{v}_{ij}^\top \mathbf{T}^\top \mathbf{T} \mathbf{v}_{ij} + (\mathbf{v}_{ij} - \mu_{\mathbf{v},ij})^\top \Sigma_{\mathbf{v},ij}^{-1} (\mathbf{v}_{ij} - \mu_{\mathbf{v},ij}) \\
 \Rightarrow & \mathbf{h}^\top \mathbf{h} - \mathbf{h}^\top \mathbf{T} \mathbf{v}_{ij} - \mathbf{v}_{ij}^\top \mathbf{T}^\top \mathbf{h} + \mathbf{v}_{ij}^\top \mathbf{T}^\top \mathbf{T} \mathbf{v}_{ij} + \mathbf{v}_{ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mathbf{v}_{ij} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij} - \mathbf{v}_{ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij} - \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mathbf{v}_{ij} \\
 \Rightarrow & \mathbf{h}^\top \mathbf{h} - \mathbf{h}^\top \mathbf{T} \mathbf{v}_{ij} - \mathbf{v}_{ij}^\top \mathbf{T}^\top \mathbf{h} + \mathbf{v}_{ij}^\top \mathbf{T}^\top \mathbf{T} \mathbf{v}_{ij} + \mathbf{v}_{ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mathbf{v}_{ij} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij} - \mathbf{v}_{ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij} - \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mathbf{v}_{ij} \\
 = & \mathbf{h}^\top \mathbf{h} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij} + \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} - \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{h} + \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij}) - (\mathbf{h}^\top \mathbf{T} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} \\
 \Rightarrow & \mathbf{h}^\top \mathbf{h} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij} + \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} - \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{h} + \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij}) - (\mathbf{h}^\top \mathbf{T} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} \\
 = & \mathbf{h}^\top \mathbf{h} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij} + \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} - \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} (\mathbf{T}^\top \mathbf{h} + \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij}) \\
 & - (\mathbf{h}^\top \mathbf{T} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1}) (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} \\
 \Rightarrow & \mathbf{h}^\top \mathbf{h} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij} + \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} - \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} (\mathbf{T}^\top \mathbf{h} + \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij}) \\
 & - (\mathbf{h}^\top \mathbf{T} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1}) (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} \\
 = & \mathbf{h}^\top \mathbf{h} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij} + \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} - \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} (\mathbf{T}^\top \mathbf{h} + \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij}) \\
 & - (\mathbf{h}^\top \mathbf{T} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1}) (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} \\
 \Rightarrow & \mathbf{h}^\top \mathbf{h} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij} + \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij} - \mathbf{v}_{ij}^\top (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} (\mathbf{T}^\top \mathbf{h} + \Sigma_{\mathbf{v},ij}^{-1} \mu_{\mathbf{v},ij}) \\
 & - (\mathbf{h}^\top \mathbf{T} + \mu_{\mathbf{v},ij}^\top \Sigma_{\mathbf{v},ij}^{-1}) (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} (\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1}) \mathbf{v}_{ij}
 \end{aligned}$$

[illegible]

$$\begin{aligned}
&\Rightarrow \frac{1}{(2\pi)^d |\Sigma_{\mathbf{v},ij}|^{1/2}} \frac{(2\pi)^{d/2}}{|2(\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})|^{1/2}} \exp \left[-(\mathbf{h} - \mathbf{T}\mu_{\mathbf{v},ij})^\top \{\mathbf{I}_d - \mathbf{T}(\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} \mathbf{T}^\top\} (\mathbf{h} - \mathbf{T}\mu_{\mathbf{v},ij}) \right] \\
&= \frac{1}{|\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T} + \mathbf{I}_{2d}|^{1/2}} \exp \left[-(\mathbf{h} - \mathbf{T}\mu_{\mathbf{v},ij})^\top \{\mathbf{I}_d - \mathbf{T}(\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} \mathbf{T}^\top\} (\mathbf{h} - \mathbf{T}\mu_{\mathbf{v},ij}) \right]
\end{aligned}$$

Note that the expression

$$\begin{aligned}
(\mathbf{T}\mu_{\mathbf{v},ij})^\top [\mathbf{I}_d - \mathbf{T}\{\mathbf{T}^\top \mathbf{T} + (\Sigma_{\mathbf{v},ij})^{-1}\}^{-1} \mathbf{T}^\top] (\mathbf{T}\mu_{\mathbf{v},ij}) &= \mu_{\mathbf{v},ij}^\top \{\mathbf{T}^\top \mathbf{T} - \mathbf{T}^\top \mathbf{T}(\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} \mathbf{T}^\top \mathbf{T}\} \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top [\mathbf{T}^\top \mathbf{T} - \mathbf{T}^\top \mathbf{T}\{(\mathbf{T}^\top \mathbf{T})^{-1} \mathbf{T}^\top \mathbf{T} + (\mathbf{T}^\top \mathbf{T})^{-1} \Sigma_{\mathbf{v},ij}^{-1}\}^{-1}] \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top [\mathbf{T}^\top \mathbf{T} - \mathbf{T}^\top \mathbf{T}\{\mathbf{I}_{2d} + (\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T})^{-1}\}^{-1}] \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top \{\mathbf{T}^\top \mathbf{T} - \mathbf{T}^\top \mathbf{T}(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T} + \mathbf{I}_{2d})^{-1} \Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T}\} \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top [\mathbf{T}^\top \mathbf{T} - \mathbf{T}^\top \mathbf{T}\{(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T})^{-1} - (\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T})^{-1}(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T} + \mathbf{I}_{2d})^{-1}\} \\
&\quad \Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T}] \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top [\mathbf{T}^\top \mathbf{T} - \mathbf{T}^\top \mathbf{T}(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T})^{-1} \Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T} \\
&\quad + \mathbf{T}^\top \mathbf{T}(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T})^{-1}(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T} + \mathbf{I}_{2d})^{-1} \Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T}] \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top \{\mathbf{T}^\top \mathbf{T}(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T})^{-1}(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T} + \mathbf{I}_{2d})^{-1} \Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T}\} \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top [\mathbf{T}^\top \mathbf{T}(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T})^{-1}\{\mathbf{I}_{2d} + (\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T})^{-1}\}^{-1}] \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top \{\mathbf{T}^\top \mathbf{T}(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T} + \mathbf{I}_{2d})^{-1}\} \mu_{\mathbf{v},ij},
\end{aligned}$$

where we used the Sherman Woodbury matrix inverse formula, is equal to the 2nd and 5th term in (****), i.e.,

$$\begin{aligned}
\mu_{\mathbf{v},ij}^\top \{\Sigma_{\mathbf{v},ij}^{-1} - (\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij} + \mathbf{I}_{2d})^{-1} \Sigma_{\mathbf{v},ij}^{-1}\} \mu_{\mathbf{v},ij} &= \mu_{\mathbf{v},ij}^\top [\Sigma_{\mathbf{v},ij}^{-1} - \{\mathbf{I}_{2d} - \mathbf{I}_{2d} \mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij} (\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij} + \mathbf{I}_{2d})^{-1}\} \Sigma_{\mathbf{v},ij}^{-1}] \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top [\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij} (\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij} + \mathbf{I}_{2d})^{-1} \Sigma_{\mathbf{v},ij}^{-1}] \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top \{\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij} (\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij} + \Sigma_{\mathbf{v},ij})^{-1}\} \mu_{\mathbf{v},ij} \\
&= \mu_{\mathbf{v},ij}^\top \{\mathbf{T}^\top \mathbf{T}(\Sigma_{\mathbf{v},ij} \mathbf{T}^\top \mathbf{T} + \mathbf{I}_{2d})^{-1}\} \mu_{\mathbf{v},ij}.
\end{aligned}$$

Also

$$\begin{aligned}
(\mathbf{T}\mu_{\mathbf{v},ij})^\top \{\mathbf{I}_d + \mathbf{T}(\mathbf{T}^\top \mathbf{T} + \Sigma^{-1})^{-1} \mathbf{T}^\top\} \mathbf{h} &= \mu_{\mathbf{v},ij}^\top \mathbf{T}^\top \{\mathbf{I}_d + \mathbf{T}(\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} \mathbf{T}^\top\} \mathbf{h} \\
&= \mu_{\mathbf{v},ij}^\top \mathbf{T}^\top \mathbf{h} - \mu_{\mathbf{v},ij}^\top \mathbf{T}^\top \mathbf{T}(\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1} \mathbf{T}^\top \mathbf{h} \\
&= \mu_{\mathbf{v},ij}^\top \{\mathbf{I}_{2d} - \mathbf{T}^\top \mathbf{T}(\mathbf{T}^\top \mathbf{T} + \Sigma_{\mathbf{v},ij}^{-1})^{-1}\} \mathbf{T}^\top \mathbf{h} \\
&= \mu_{\mathbf{v},ij}^\top [\mathbf{I}_{2d} - \{\mathbf{I}_{2d} + (\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij})^{-1}\}^{-1}] \mathbf{T}^\top \mathbf{h} \\
&= \mu_{\mathbf{v},ij}^\top (\mathbf{I}_{2d} - [\mathbf{I}_{2d} - (\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij})^{-1}\{\mathbf{I}_{2d} + (\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij})^{-1}\}^{-1}]) \mathbf{T}^\top \mathbf{h} \\
&= \mu_{\mathbf{v},ij}^\top [\mathbf{I}_{2d} - \mathbf{I}_{2d} + (\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij})^{-1}\{\mathbf{I}_{2d} + (\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij})^{-1}\}^{-1}] \mathbf{T}^\top \mathbf{h} \\
&= \mu_{\mathbf{v},ij}^\top (\mathbf{T}^\top \mathbf{T} \Sigma_{\mathbf{v},ij} + \mathbf{I}_{2d})^{-1} \mathbf{T}^\top \mathbf{h},
\end{aligned}$$

which is the 3rd and 4th term in (****).

Sherman Woodbury formula,

$$\begin{aligned}
(A + B)^{-1} &= A^{-1} - A^{-1} B(A + B)^{-1} \\
(u^2 D + \mathbf{1}_{d \times d})^{-1} &= (u^2 D)^{-1} - (u^2 D)^{-1} \mathbf{1}_{d \times d} (u^2 D + \mathbf{1}_{d \times d})^{-1} = (u^2 D)^{-1} - (u^2 D)^{-1} (u^2 D + \mathbf{1}_{d \times d}) \\
(\mathbf{1}_{d \times d} + u^2 D)^{-1} &= \mathbf{1}_{d \times d} - \mathbf{1}_{d \times d} (u^2 D) (\mathbf{1}_{d \times d} + u^2 D)^{-1} = \mathbf{1}_{d \times d} - \{(u^2 D)^{-1} + \mathbf{1}_{d \times d}\}^{-1} \\
(\mathbf{1}_{d \times d} + u^2 D)^{-T} &= \mathbf{1}_{d \times d} - \{(u^2 D)^{-1} + \mathbf{1}_{d \times d}\}^{-T}
\end{aligned}$$

2 Estimation

2.1 Universal Kriging

Throughout this work, we assume the following model for our spatio-temporal processes:

$$\mathbf{Y}(\mathbf{s}, t) = \boldsymbol{\mu}(\mathbf{s}, t) + \mathbf{Z}(\mathbf{s}, t), \quad (\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}, \quad (1)$$

where $\boldsymbol{\mu}(\mathbf{s}, t)$ is the vector-valued spatio-temporal mean of multivariate random field that may be spatially varying and $\mathbf{Z}(\mathbf{s}, t)$ is a zero mean multivariate second-order stationary spatio-temporal random field. Assume that the mean function in (1) can be characterized as a linear combination of some covariates X_1, X_2, \dots, X_M . Denote by $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \dots, \boldsymbol{\beta}_p^\top)^\top \in \mathbb{R}^{Mp}$ the vector of mean parameters, where $\boldsymbol{\beta}_i = (\beta_{1,i}, \dots, \beta_{M,i})^\top \in \mathbb{R}^M$, for $i = 1, \dots, p$, and $\mathbf{X} = \{\mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}_1, t_1)^\top, \mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}_2, t_2)^\top, \dots, \mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}_n, t_n)^\top\}^\top \in \mathbb{R}^{np \times Mp}$, where $\mathbf{X}(\mathbf{s}, t) = \{X_1(\mathbf{s}, t), \dots, X_M(\mathbf{s}, t)\} \in \mathbb{R}^M$. The model in (1) becomes $\mathbf{Y}(\mathbf{s}, t) = \{\mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}, t)^\top\} \boldsymbol{\beta} + \mathbf{Z}(\mathbf{s}, t)$. The universal cokriging predictor $\hat{\mathbf{Z}}(\mathbf{s}_0, t_0)$ of $\mathbf{Z}(\mathbf{s}, t)$ at unsampled spatio-temporal location (\mathbf{s}_0, t_0) is

$$\hat{\mathbf{Y}}(\mathbf{s}_0, t_0) = [\mathbf{c} + \mathbf{X}\{\mathbf{X}^\top \boldsymbol{\Sigma}(\hat{\boldsymbol{\Theta}})^{-1} \mathbf{X}\}^{-1} \{\mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}_0, t_0)^\top - \mathbf{X}^\top \boldsymbol{\Sigma}(\hat{\boldsymbol{\Theta}})^{-1} \mathbf{c}\}]^\top \boldsymbol{\Sigma}^{-1} \mathbf{Y},$$

where $\mathbf{c} = ([\{C_{ij}(\mathbf{s}_l - \mathbf{s}_0, t_l, t_0; \hat{\boldsymbol{\Theta}})\}_{i,j=1}^n]^n)^\top \in \mathbb{R}^{np \times p}$. The corresponding cokriging variance is

$$\begin{aligned} \text{var}\{\hat{\mathbf{Y}}(\mathbf{s}_0, t_0)\} &= \text{trace}[\mathbf{C}(\mathbf{0}, 0; \hat{\boldsymbol{\Theta}}) - \mathbf{c}^\top \boldsymbol{\Sigma}(\hat{\boldsymbol{\Theta}})^{-1} \mathbf{c} \\ &\quad + \{\mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}_0, t_0)^\top - \mathbf{X}^\top \boldsymbol{\Sigma}(\hat{\boldsymbol{\Theta}})^{-1} \mathbf{c}\}^\top \{\mathbf{X}^\top \boldsymbol{\Sigma}(\hat{\boldsymbol{\Theta}})^{-1} \mathbf{X}\}^{-1} \{\mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}_0, t_0)^\top - \mathbf{X}^\top \boldsymbol{\Sigma}(\hat{\boldsymbol{\Theta}})^{-1} \mathbf{c}\}], \end{aligned}$$

3 Simulation Study

Figures S2 and S3 present the boxplots when fitting M2 and M3 to data generated from M3 at other values of k , the parameter that controls the dependence in time, such that the dependence in time decreases as k increases. The main manuscript shows the results for $k = 0.001$, representing

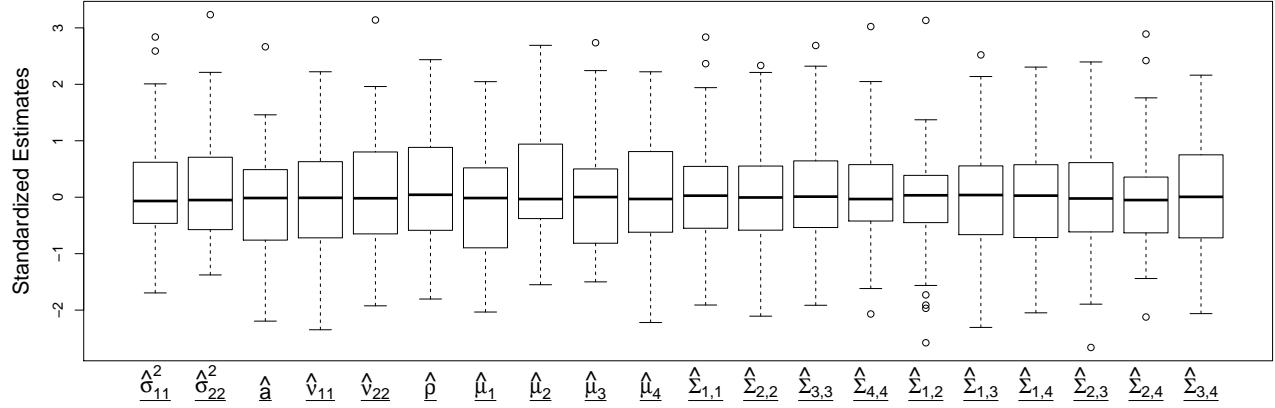


Figure S1: Boxplots of standardized estimated parameters of via multi-step MLE.

a long range / strong temporal dependence. The models with $k = 0.1$ and $k = 1$ represent the moderate and weak temporal dependence, respectively.

4 Application

4.1 Timeseries Plots

4.2 Testing for Stationarity of Spatio-Temporal Data

Since the data has a very short timeseries, we only test for stationarity in the spatial domain. Using the testing method in Jun and Genton (2012), we test the second-order stationarity of the residuals obtained using the generalized least squares regression with covariates, namely, relative humidity and temperature. The testing method requires the computation of the empirical spatial marginal and cross-covariances of two disjoint spatial domains. We perform the test two times. First, we divide Saudi Arabia vertically, i.e. $[35, 45] \times [17, 32]$ and $[45, 55] \times [17, 32]$, and test whether there is enough evidence to reject the assumption that the empirical spatial marginal and cross-covariances in the West and East of Saudi Arabia are the same. Next, we divide Saudi Arabia horizontally, i.e. $[35, 55] \times [17, 24.5]$ and $[35, 55] \times [24.5, 32]$, and perform the test.

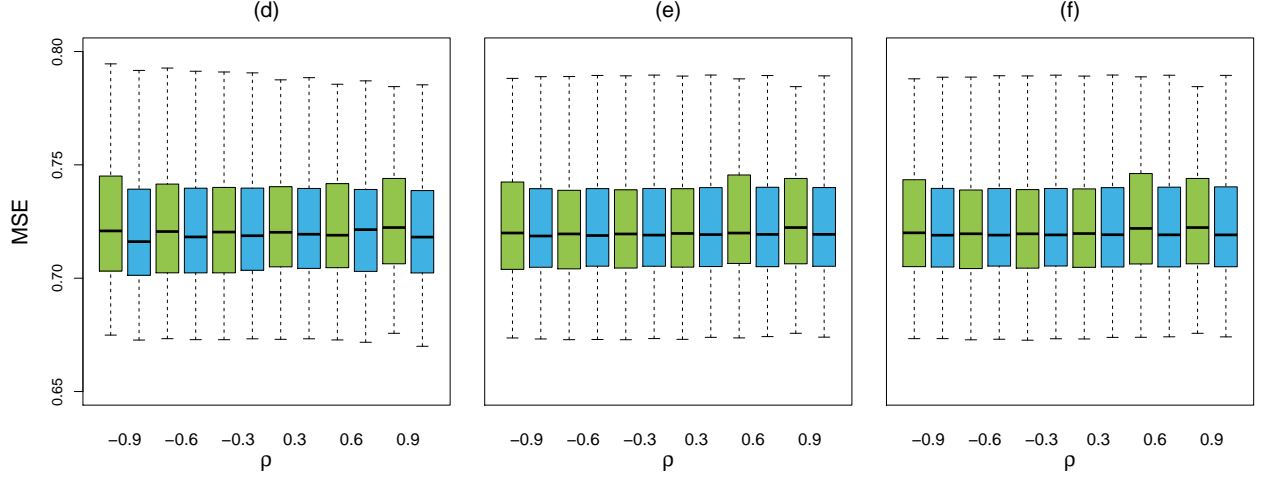


Figure S2: Boxplots of the MSEs under scenarios (d)-(f) when M2 (green) and M3 (blue) are fitted to data generated from M3 at different values of ρ at $k = 1$.

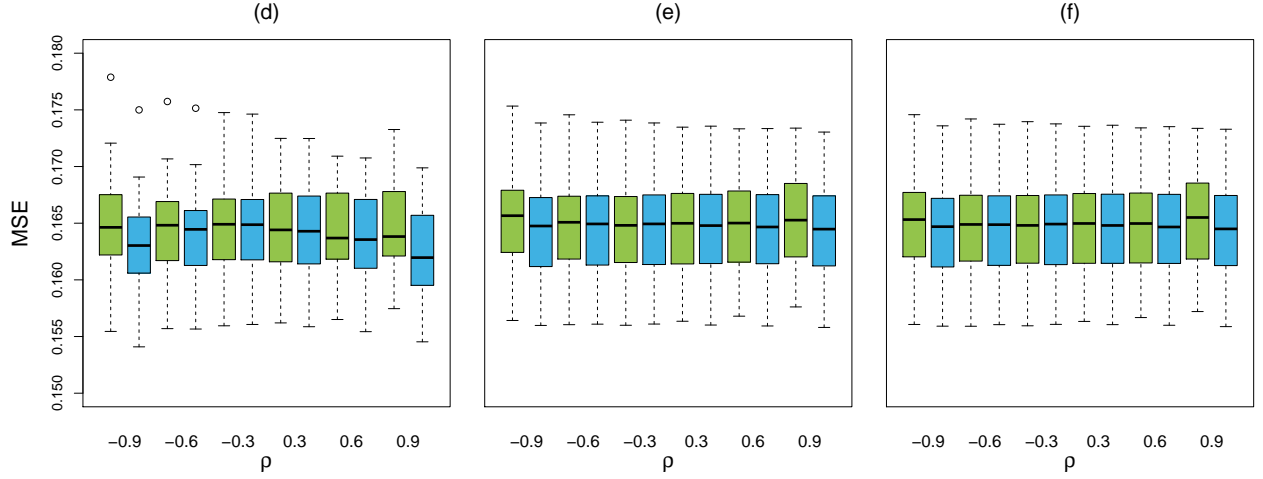


Figure S3: Boxplots of the MSEs under scenarios (d)-(f) when M2 (green) and M3 (blue) are fitted to data generated from M3 at different values of ρ at $k = 0.1$.

We choose the following spatial lags at which we compute the p-values:

$$\Lambda_1 = \{\mathbf{h} = (0, 2), (2, 0), (2, 2)\},$$

$$\Lambda_2 = \{\mathbf{h} = (0, 2), (2, 0), (2, 2), (0, 4), (4, 0), (4, 4)\},$$

$$\Lambda_3 = \{\mathbf{h} = (0, 2), (2, 0), (2, 2), (0, 4), (4, 0), (4, 4), (0, 6), (6, 0), (6, 6)\}$$

in $\times 10^2$ km. Tables 1 and 2 present the p-values from the test for stationarity with the empirical marginal and cross-covariances calculated using spatial lag \mathbf{h} or anisotropy-corrected spatial lag

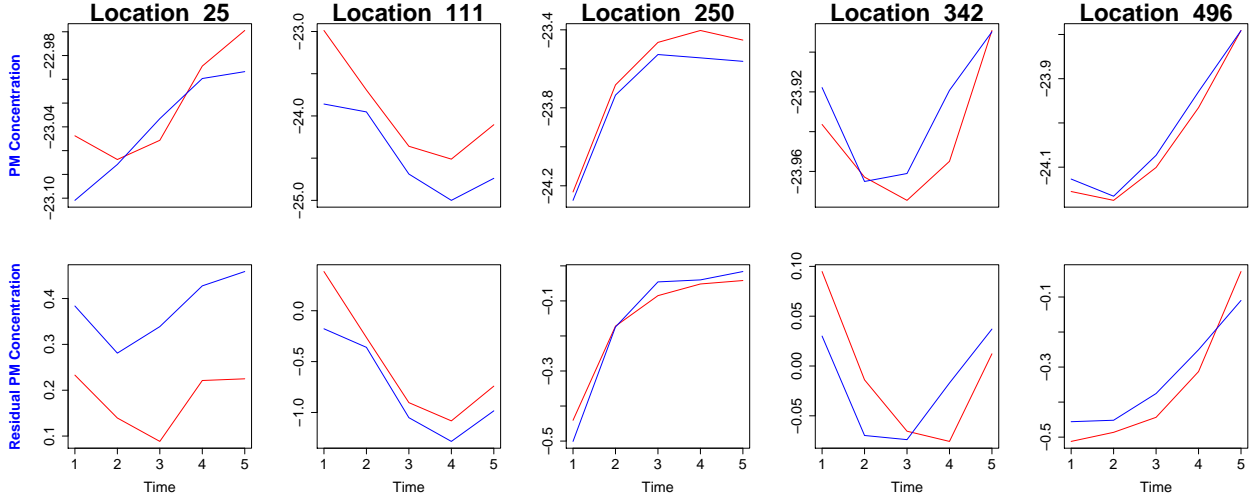


Figure S4: Timeseries plots of PM concentrations on January 18, 2019, 9:00 to 21:00, at 880 hPa (red) and 925 hPa (blue), in log scale of $\mu\text{g}/\text{m}^3$.

Rh. From the p-values, it can be seen that we need to reject the stationarity assumption using a significance level of $\alpha = 0.05$. Using the anisotropy corrected spatial lags, however, generally results to a failure in rejecting the stationarity assumption using a significance level of $\alpha = 0.05$

Table 1: P-values from the test for several combinations of spatial lags when dividing Saudi Arabia vertically and computing for the empirical C_{11} , C_{22} , and C_{12} .

	C_{11}			C_{22}			C_{12}		
	Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3
h	0.6081	0.1339	0.0286	0.6204	0.0451	0.0013	0.2636	0.0863	0.0136
Rh	0.9671	0.0881	0.0782	0.8838	0.0292	0.0348	0.9233	0.0928	0.0538

Table 2: P-values from the test for several combinations of spatial lags when dividing Saudi Arabia horizontally and computing for the empirical C_{11} , C_{22} , and C_{12} . Note that Λ_3 cannot be properly computed when dividing Saudi Arabia horizontally.

	C_{11}			C_{22}			C_{12}		
	Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3
h	0.1102	0.2939	-	0.0147	0.0362	-	0.0481	0.1266	-
Rh	0.1433	0.2349	-	0.0364	0.0517	-	0.0840	0.1801	-

References

Jun, M. and Genton, M. G. (2012). A test for stationarity of spatio-temporal random fields on planar and spherical domains. *Statistica Sinica*, 22:1737–1764.

Schlather, M. (2010). Some covariance models based on normal scale mixtures. *Bernoulli*, 16(3):780–797.