Spatio-Temporal Cross-Covariance Functions under the Lagrangian Framework with Multiple Advections

Mary Lai O. Salvaña, Amanda Lenzi, and Marc G. Genton[1]

February 9, 2021

Abstract

When analyzing the spatio-temporal dependence in most environmental and earth sciences variables such as pollutant concentrations at different levels of the atmosphere, a special property is observed: the covariances and cross-covariances are stronger in certain directions. This property is attributed to the presence of natural forces, such as wind, which cause the transport and dispersion of these variables. This spatio-temporal dynamics prompts the integration of the Lagrangian reference frame to any Gaussian spatio-temporal geostatistical model. Under this modeling framework, a whole new class was birthed and is known as the class of spatio-temporal covariance functions under the Lagrangian framework, with several developments already established in the univariate setting, in both stationary and nonstationary formulations, but less so in the multivariate case. Despite the many advances in this modeling approach, efforts have yet to be directed to probing the case for the use of multiple advections, especially when several variables are involved. Accounting for multiple advections makes the Lagrangian framework a more viable approach in modeling realistic multivariate transport scenarios. In this work, we establish a class of Lagrangian spatio-temporal cross-covariance functions with multiple advections, study its properties, and demonstrate its use on a bivariate pollutant dataset of particulate matter in Saudi Arabia.

Some key words: cross-covariance function; Lagrangian framework; multiple advections; multivariate random field; spatio-temporal; transport effect.

1 Introduction

Many environmental and earth sciences datasets record several variables at certain locations over certain periods of time. These datasets contain rich spatio-temporal information which can be used to enhance predictions. In Geostatistics, these data are used to calibrate the parameters of spatio-temporal cross-covariance functions which are valid functions that describe the spatio-temporal relationships within each variable and between any two variables; see Genton and Kleiber (2015), Alegría et al. (2019), and Salvaña and Genton (2020) for recent reviews of the available models in the literature. In multivariate spatio-temporal Gaussian geostatistical modeling, we work with a spatio-temporal process $\mathbf{Y}(\mathbf{s}, t) = \{Y_1(\mathbf{s}, t), \ldots, Y_p(\mathbf{s}, t)\}^\top$, $(\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}$, such that at each spatial location $\mathbf{s} \in \mathbb{R}^d, d \geq 1$, and at each time $t \in \mathbb{R}$, there are p variables. A common assumption on $\mathbf{Y}(\mathbf{s}, t)$ is that it can be decomposed into a sum of a deterministic and a random component, i.e.,

$$\mathbf{Y}(\mathbf{s}, t) = \boldsymbol{\mu}(\mathbf{s}, t) + \mathbf{Z}(\mathbf{s}, t), \tag{1}$$

where $\boldsymbol{\mu}(\mathbf{s}, t)$ is a mean function and $\mathbf{Z}(\mathbf{s}, t)$ a zero mean multivariate spatio-temporal Gaussian process. When $\mathbf{Z}(\mathbf{s}, t)$ is second-order stationary, it is completely characterized by its spatio-temporal matrix-valued stationary cross-covariance function $\mathbf{C}(\mathbf{h}, u) = \{C_{ij}(\mathbf{h}, u)\}_{i,j=1}^p$ on $\mathbb{R}^d \times \mathbb{R}$, with entries defined as follows:

$$C_{ij}(\mathbf{h}, u) = \text{cov}\left\{ Z_i(\mathbf{s}, t), Z_j(\mathbf{s} + \mathbf{h}, t + u) \right\}, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \tag{2}$$

for $i, j = 1, \ldots, p$. Often, C_{ij} is termed the marginal covariance function when $i = j$, and it is called the cross-covariance function whenever $i \neq j$. From (2), it is easy to check that $C_{ij}(\mathbf{h}, u) = C_{ji}(-\mathbf{h}, -u)$, for any $(\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$ and $i, j = 1, \ldots, p$. However, it is not always the case that $C_{ij}(\mathbf{h}, u) = C_{ij}(\mathbf{h}, -u) = C_{ij}(-\mathbf{h}, u) = C_{ij}(-\mathbf{h}, -u)$ for $i, j = 1, \ldots, p$. This equality involving different combinations of the signs of the spatial and temporal lags is referred to as the full spatio-

temporal symmetry property. A consequence of the full spatio-temporal symmetry property is that the marginal and cross-covariances between observations at $(\mathbf{s}, t)$ and $(\mathbf{s}+\mathbf{h}, t+u)$ is equal to the marginal and cross-covariances between observations at $(\mathbf{s}, t+u)$ and $(\mathbf{s}+\mathbf{h}, t)$. This property is highly restrictive and most often not exhibited by environmental and earth sciences data. For instance, with wind that blows from West to East, airborne substances can be transported such that the measurement of its concentration at a reference site will be highly correlated to those, taken after some time, at the sites to its East. Because of the presence of a transport medium, the same degree of correlation may not be expressed between the concentration measurements at the aforementioned reference site and those, taken after some time, at the sites to its West. This behavior is recognized as spatio-temporal asymmetry in the marginal and cross-covariance structures (Gneiting et al., 2007; Huang et al., 2020). Whenever such asymmetry is detected, more appropriate spatio-temporal asymmetric cross-covariance functions must be used such as the class of spatio-temporal asymmetric models proposed in Stein (2005a), suitable for $p = 1$, and the latent dimensions model in Apanasovich and Genton (2010), catering to $p > 1$.

Within the category of models that captures spatio-temporal asymmetric dependence, a rich subclass was established dedicated to transported purely spatial random fields. The models under this subclass are termed spatio-temporal cross-covariance functions under the Lagrangian framework and the pioneering work in the univariate setting is attributed to Cox and Isham (1988). Under this modeling framework, consider a purely spatial random field with a stationary covariance function $C^S(\mathbf{h})$ and suppose that this entire random field moves forward in time, with a random advection velocity $\mathbf{V} \in \mathbb{R}^d$. This means that the observations from all spatial locations are being advected or transported by one and the same random advection velocity at each time. The resulting Lagrangian spatio-temporal covariance function has the form:

$$C(\mathbf{h}, u) = \mathrm{E}_{\mathbf{V}} \left\{ C^S(\mathbf{h} - \mathbf{V}u) \right\}, \tag{3}$$

where the expectation is taken with respect to $\mathbf{V}$. Through $\mathbf{V}$, one can transform a purely

spatial random field into a spatio-temporal random field. Moreover, by such a transformation to the spatial coordinates that incorporates information regarding a transport phenomenon, a spatio-temporal covariance function can be derived from a purely spatial covariance function. Depending on the distribution assumed by $\mathbf{V}$, the model in (3) may not obtain an explicit form, however, numerical solutions can be easily obtained. The simplest form of (3) may be derived when $\mathbf{V}$ is chosen to be constant, i.e., $\mathbf{V} = \mathbf{v}$, and the model is then termed the frozen field model. Several authors have used the frozen field model to analyze wind speeds (Gneiting et al., 2007; Ezzat et al., 2018) and solar irradiance (Lonij et al., 2013; Inoue et al., 2012; Shinozaki et al., 2016). Intuitively, the frozen field model is an unrealistic assumption and a highly idealized model of the transport phenomenon. Wind, for example, may or may not blow at any time of the day, and when it blows, the wind speeds and directions are rarely identical. Therefore, models that allow for variability of the transport, i.e., the non-frozen field models, are preferred. An explicit form of (3) exists when $\mathbf{V} \sim \mathcal{N}_d(\boldsymbol{\mu}_{\mathbf{V}}, \boldsymbol{\Sigma}_{\mathbf{V}})$ and $C^S(\mathbf{h})$ is a normal scale-mixture model (Schlather, 2010). The derived non-frozen field model has the form:

$$C(\mathbf{h}, u) = \frac{1}{\sqrt{|\mathbf{I}_d + \boldsymbol{\Sigma}_{\mathbf{V}} u^2|}} C^S\{(\mathbf{h} - \boldsymbol{\mu}_{\mathbf{V}} u)^\top \left(\mathbf{I}_d + \boldsymbol{\Sigma}_{\mathbf{V}} u^2\right)^{-1} (\mathbf{h} - \boldsymbol{\mu}_{\mathbf{V}} u)\}. \tag{4}$$

Several versions of the Lagrangian spatio-temporal covariance functions have appeared since the seminal work of Cox and Isham (1988). Park and Fuentes (2006) adapted the modeling framework for axial symmetry in time, axial symmetry in space, and diagonal symmetry in space. Porcu et al. (2006) explored some anisotropic extensions and Christakos (2017) introduced an acceleration component. Salvaña and Genton (2020) proposed the multivariate extension such that the model in (3) remains valid when the underlying purely spatial covariance function is a matrix-valued nonstationary cross-covariance function $\mathbf{C}^S(\mathbf{s}_1, \mathbf{s}_2)$ on $\mathbb{R}^d$, $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$. Their spatio-temporal matrix-valued nonstationary extension has the form:

$$\mathbf{C}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = \mathrm{E}_{\mathbf{V}}\{\mathbf{C}^S(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2)\}, \tag{5}$$

3

for $t_1, t_2 \in \mathbb{R}$. The above model includes matrix-valued stationary cross-covariance functions, i.e., $\mathbf{C}^S(\mathbf{h})$, as the underlying purely spatial cross-covariance functions. Their umbrella theorem relies on a single $\mathbf{V}$ which implies that every component of $\mathbf{Z}(\mathbf{s}, t) \in \mathbb{R}^p$ is transported by the same advection velocity. However, different variables may experience different transport patterns which render the model in (5) inadequate. To deal with this issue, they proposed a Lagrangian spatio-temporal cross-covariance function that is a linear combination of uncorrelated univariate Lagrangian spatio-temporal covariance functions, each depending on different advections. Their proposal is a good first step to addressing this multiple advections problem. When the marginal and cross-advections are introduced, i.e., $\mathbf{V}_{ij} \in \mathbb{R}^d, i, j = 1, \ldots, p$, several questions arise regarding the validity of the extended model, including which values of $\mathbf{V}_{ij}$, $i \neq j$, will preserve the positive definiteness of the cross-covariance matrix resulting from (5). In this work, we aim at answering such a fundamental question and providing a comprehensive treatment to the Lagrangian spatio-temporal cross-covariance functions with multiple advections, with a main focus on underlying purely spatial cross-covariance functions that are stationary.

The remainder of the paper is organized as follows. Section 2 presents the proposed extension of (5) with multiple advections and introduces some examples. Section 3 details the estimation procedure. Section 4 investigates the consequences of neglecting multiple advections in multivariate Lagrangian spatio-temporal modeling. Section 5 compares the performance of the proposed models with other benchmark models in the literature using a bivariate pollutant dataset of particulate matter in Saudi Arabia. The conclusion is presented in Section 6 and proofs are collected in the Appendix.

2 Lagrangian Framework with Multiple Advections

The validity of Lagrangian spatio-temporal cross-covariance functions with different advection for every variable can be established by considering a zero-mean multivariate spatio-temporal

random field

$$\mathbf{Z}(\mathbf{s}, t) = \left\{ \tilde{Z}_1(\mathbf{s} - \mathbf{V}_{11}t), \ldots, \tilde{Z}_p(\mathbf{s} - \mathbf{V}_{pp}t) \right\}^\top, \tag{6}$$

such that $\tilde{\mathbf{Z}}(\mathbf{s}) = \left\{ \tilde{Z}_1(\mathbf{s}), \ldots, \tilde{Z}_p(\mathbf{s}) \right\}^\top$ is a zero-mean multivariate purely spatial random field and every component of $\tilde{\mathbf{Z}}$ is transported by different random advections $\mathbf{V}_{ii} \in \mathbb{R}^d, i = 1, \ldots, p$. The resulting matrix-valued spatio-temporal cross-covariance function of the process in (6) is given in the following theorem.

Theorem 1 *Let* $\mathbf{V}_{11}, \mathbf{V}_{22}, \ldots, \mathbf{V}_{pp}$ *be random vectors on* $\mathbb{R}^d$. *If* $\mathbf{C}^S(\mathbf{h})$ *is a valid purely spatial matrix-valued stationary cross-covariance function on* $\mathbb{R}^d$ *then*

$$\mathbf{C}(\mathbf{h}; t_1, t_2) = \mathrm{E}_{\boldsymbol{\mathcal{V}}}[\{C_{ij}^S(\mathbf{h} - \mathbf{V}_{ii}t_1 + \mathbf{V}_{jj}t_2)\}_{i,j=1}^p], \tag{7}$$

where the expectation is taken with respect to the joint distribution of $\boldsymbol{\mathcal{V}} = (\mathbf{V}_{11}^\top, \mathbf{V}_{22}^\top, \ldots, \mathbf{V}_{pp}^\top)^\top$, *is a valid matrix-valued spatio-temporal cross-covariance function on* $\mathbb{R}^d \times \mathbb{R}$ *provided that the expectation exists.*

When $\mathbf{V}_{ii} = \mathbf{V}$, for all i, the above model reduces to the single advection case. Moreover, the model in (7) can be rewritten to resemble the form in (3) such that the temporal lag u appears, i.e.,

$$\mathbf{C}(\mathbf{h}; t_1, t_2) = \mathrm{E}_{\boldsymbol{\mathcal{V}}}[C_{ij}^S\{\mathbf{h} - \overline{\mathbf{V}}_{ij}u + (\mathbf{V}_{jj} - \mathbf{V}_{ii})m\}_{i,j=1}^p], \tag{8}$$

where $\overline{\mathbf{V}}_{ij} = \frac{\mathbf{V}_{ii} + \mathbf{V}_{jj}}{2}$, $i, j = 1, \ldots, p$, and $m = \frac{t_1 + t_2}{2}$. It can be seen that for a stationary $\mathbf{C}^S(\mathbf{h})$ and for $i \neq j$, nonstationarity in time is introduced by the Lagrangian shift in the cross-covariances. However, when $i = j$, the term that depends on the midpoint between t_1 and t_2 disappears. Hence, the marginal covariances remain stationary in time.

An explicit form of (7) can also be derived similar to (4) and is given in the following theorem.

Theorem 2 *For* $p > 2$, *let* $\boldsymbol{\mathcal{V}} = (\mathbf{V}_{11}^\top, \mathbf{V}_{22}^\top, \ldots, \mathbf{V}_{pp}^\top)^\top \sim \mathcal{N}_{pd}(\boldsymbol{\mu}_{\boldsymbol{\mathcal{V}}}, \boldsymbol{\Sigma}_{\boldsymbol{\mathcal{V}}})$. *If* $\mathbf{C}^S(\mathbf{h})$ *is a matrix-*

valued normal scale-mixture cross-covariance function, then

$$C_{ii}(\mathbf{h}, u) = \frac{C_{ii}^S\{(\mathbf{h} - \mathbf{e}_{(di-1):(di)}^\top \boldsymbol{\mu_\nu} u)^\top (\mathbf{I}_d + \mathbf{e}_{(di-1):(di)}^\top \boldsymbol{\Sigma_\nu} u^2)^{-1}(\mathbf{h} - \mathbf{e}_{(di-1):(di)}^\top \boldsymbol{\mu_\nu} u)\}}{|\mathbf{I}_d + \mathbf{e}_{(di-1):(di)}^\top \boldsymbol{\Sigma_\nu} u^2|^{1/2}}, \quad (9)$$

where $\mathbf{e}_{(di-1):(di)}$ is the sub-matrix of $\mathbf{I}_{pd}$, comprised of its $(di-1)$-th and (di)-th rows, for $i = 1, \ldots, p$, and

$$C_{ij}(\mathbf{h}; t_1, t_2) = \frac{C_{ij}^S((\mathbf{h} - \mathbf{T}\tilde{\mathbf{e}}^\top \boldsymbol{\mu_\nu})^\top [\mathbf{I}_d + \mathbf{T}\{\mathbf{T}^\top \mathbf{T} + (\tilde{\mathbf{e}}^\top \boldsymbol{\Sigma_\nu})^{-1}\}^{-1}\mathbf{T}^\top](\mathbf{h} - \mathbf{T}\tilde{\mathbf{e}}^\top \boldsymbol{\mu_\nu}))}{|\mathbf{I}_{2d} + (\tilde{\mathbf{e}}^\top \boldsymbol{\Sigma_\nu})\mathbf{T}^\top \mathbf{T}|^{1/2}}, \quad (10)$$

where $\mathbf{T} = (t_1 \mathbf{I}_d \quad -t_2 \mathbf{I}_d)$, $\tilde{\mathbf{e}} = \mathbf{e}_{\{(di-1):(di),(dj-1):(dj)\}}$, such that $\mathbf{e}_{\{(di-1):(di),(dj-1):(dj)\}}$ is the sub-matrix of $\mathbf{I}_{pd}$ comprised of its $(di-1)$-th, (di)-th, $(dj-1)$-th, and (dj)-th rows, for $i, j = 1, \ldots, p$, $i \neq j$.

Several properties of non-frozen Lagrangian spatio-temporal cross-covariance functions with multiple advections can be identified based on the forms given in Theorem 2. First, the spatial lag at which maximum value occurs is at $\mathbf{e}_{(di-1):(di)}^\top \boldsymbol{\mu_\nu} u$, for the marginals, and at $\mathbf{T}\tilde{\mathbf{e}}^\top \boldsymbol{\mu_\nu}$, for the cross-covariances. While the spatial lag $\mathbf{e}_{(di-1):(di)}^\top \boldsymbol{\mu_\nu} u$ is the same for any t_1 and t_2 such that $u = t_1 - t_2$, the spatial lag $\mathbf{T}\tilde{\mathbf{e}}^\top \boldsymbol{\mu_\nu}$ is different for different values of t_1 and t_2. A significant implication of this is that for any variable, regardless of its exact spatial and temporal location, it is highly dependent with itself that is situated at a location $\mathbf{e}_{(di-1):(di)}^\top \boldsymbol{\mu_\nu} u$ away and u time steps away. Similar dynamics cannot be observed regarding the dependence between two variables with different advections. Second, there is purely spatial variable asymmetry whenever $t_1 = t_2$ (or $u = 0$) but $t_1, t_2 \neq 0$. Purely spatial variable asymmetry is a property of the purely spatial marginal cross-covariance function wherein $C_{ij}(\mathbf{h}; t, t) \neq C_{ji}(\mathbf{h}; t, t)$, for any t such that $t = t_1 = t_2$ (Li and Zhang, 2011; Huang et al., 2020). A consequence of this property is that the maximum purely spatial cross-covariance does not occur at spatial lag $\mathbf{0}$. As aforementioned, the maximum value of (10) occurs at $\mathbf{T}\tilde{\mathbf{e}}^\top \boldsymbol{\mu_\nu}$, which is a nonzero vector unless $t_1 = t_2 = 0$. Furthermore, the spatial lag $\mathbf{T}\tilde{\mathbf{e}}^\top \boldsymbol{\mu_\nu}$ is time-varying, .i.e., it changes as the values of t_1 and t_2 change.
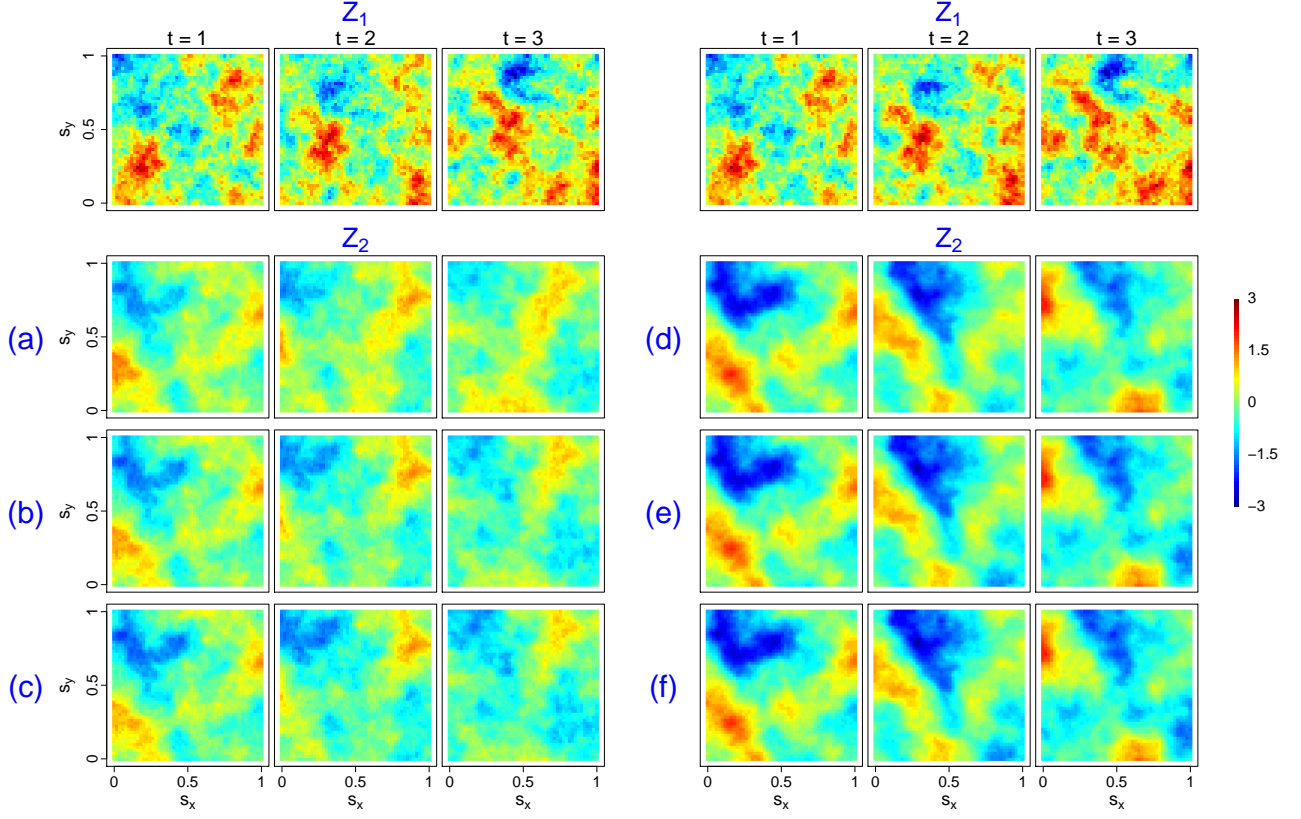
Figure 1: Simulated realizations of the Lagrangian spatio-temporal parsimonious Matérn cross-covariance function for $p = 2$ and $d = 2$, on a 50×50 regular grid in the unit square $[0,1]^2$ with purely spatial parameters, namely, $\nu_{11} = 0.5$, $\nu_{22} = 1.5$, $a = 0.23$, $\rho = 0.5$, $\sigma_{11}^2 = \sigma_{22}^2 = 1$. The plots on the left hand side show the bivariate spatio-temporal random fields simulated from (9) and (10), where a common $Z_1(\mathbf{s}, t)$ is simulated for every configuration shown in Panels (a)-(c). The different realizations of $Z_2(\mathbf{s}, t)$ under varying degrees of dependence between $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$, namely, (a) $\mathbf{V}_{11} = 0.9\mathbf{V}_{22}$, (b) $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ are independent, and (c) $\mathbf{V}_{11} = -0.9\mathbf{V}_{22}$ are displayed in Panels (a)-(c). The plots on the right hand side show the bivariate spatio-temporal random fields simulated from (13) with $d' = 1$, $\mathbf{s}'_{11} = 0.2$, $\mathbf{s}'_{22} = 0$. Similarly, a common $Z_1(\mathbf{s}, t)$ is simulated for different realizations of $Z_2(\mathbf{s}, t)$ in Panels (d)-(f), where (d) $\text{cov}(V'_{11}, V'_{22}) = 0.9$, (e) $\text{cov}(V'_{11}, V'_{22}) = 0$, and (f) $\text{cov}(V'_{11}, V'_{22}) = -0.9$.

Figure 1 shows simulated Lagrangian spatio-temporal bivariate random fields from (9) and (10) on a 50×50 regular grid in the unit square $[0,1]^2$ for $p = 2$ and $d = 2$, with the parsimonious Matérn cross-covariance function as $\mathbf{C}(\mathbf{h})$ (Gneiting et al., 2010). Panels (a)-(c) correspond to three different joint distributions of $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$, namely, (a) $\mathbf{V}_{11} = 0.9\mathbf{V}_{22}$, (b) $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ are independent, and (c) $\mathbf{V}_{11} = -0.9\mathbf{V}_{22}$. The purely spatial parameters were chosen such that the practical spatial range of the variable with a less smooth field is equal to 0.7, i.e., $C_{11}(\mathbf{h}, 0)/C_{11}(\mathbf{0}, 0) \approx 0.05$ when $\|\mathbf{h}\| = 0.7$. The marginal mean advection parameters are as
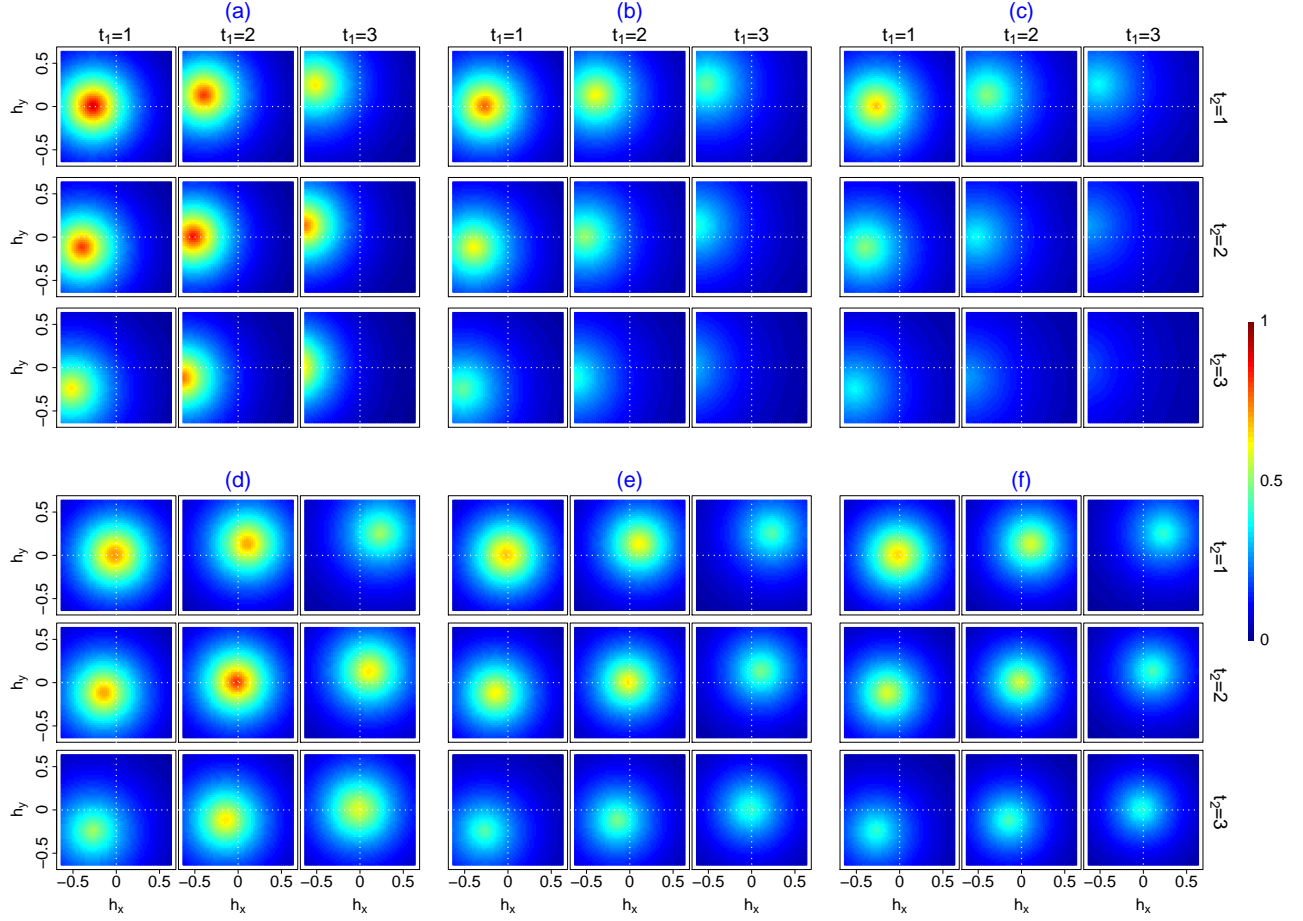
Figure 2: Heatmaps of values of $C_{12}(\mathbf{h}; t_1, t_2)$, in Panels (a)-(c), and $C_{12}\{(\mathbf{h}, \mathbf{h}'_{12}); t_1, t_2\}$, in Panels (d)-(f), corresponding to the simulated realizations in Figure 1. The values of $C_{21}(\mathbf{h}; t_1, t_2)$ and $C_{21}\{(\mathbf{h}, \mathbf{h}'_{21}); t_1, t_2\}$ can also be retrieved accordingly. The spatio-temporal dependence profiles differ depending on the joint distribution of the advection velocities.

follows: $\mathrm{E}(\mathbf{V}_{11}) = \boldsymbol{\mu}_{\mathbf{V}_{11}} = (0.1, 0.1)^\top$, $\mathrm{E}(\mathbf{V}_{22}) = \boldsymbol{\mu}_{\mathbf{V}_{22}} = (-0.1, 0.1)^\top$, $\mathrm{var}(\mathbf{V}_{22}) = 0.1 \times \mathbf{I}_2$ in (a)-(c), and $\mathrm{var}(\mathbf{V}_{11}) = 0.1 \times \mathbf{I}_2$ in (b), where $\mathbf{I}_d$ is the $d \times d$ dimensional identity matrix.

The three scenarios described above imply different strengths of spatio-temporal dependence. For ease of comparison, we simulate the same $Z_1(\mathbf{s}, t)$ for every configuration and contrast different simulated $Z_2(\mathbf{s}, t)$. It can be seen in Panels (a)-(c) that while the direction of transport of $Z_2(\mathbf{s}, t)$ is to the North West in all three cases, the spatio-temporal random fields are substantially different, with the different scenarios in the joint distribution of $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ having visible consequences in the values of Z_2 as time progresses.

The comparisons can also be done by directly examining the values of $\mathbf{C}(\mathbf{h}; t_1, t_2)$ at some

spatial lag $\mathbf{h}$ and temporal locations t_1 and t_2. Since $\mathrm{cov}(\mathbf{V}_{11}, \mathbf{V}_{22})$ does not affect the marginals, the corresponding values of C_{11} and C_{22} are the same in all three scenarios. The difference lies in the values of C_{12} and C_{21}. Panels (a)-(c) in Figure 2 visualize the corresponding spatio-temporal cross-covariance structure of the simulations in Panels (a)-(c) in Figure 1 at different $(\mathbf{h}; t_1, t_2)$ combinations via heatmaps. The plots in Figure 2 narrate how strong the dependence is between Z_1 and Z_2 taken at two spatial locations that are $\mathbf{h}$ units apart and Z_1 is either behind, ahead, or at the same time as Z_2. Figure 2(a) can be read as follows: the first plot in the first row of the 3×3 panel indicates that when Z_1 and Z_2 are both taken at the same temporal location, i.e., $t_1 = t_2 = 1$, then Z_1 has the highest degree of dependence with Z_2 taken at a spatial location that is $\boldsymbol{\mu}_{\mathbf{V}_{11}} - \boldsymbol{\mu}_{\mathbf{V}_{22}}$ units away from Z_1. Next, the second plot in the first row shows that if Z_2 is taken at $t_2 = 1$ and Z_1 is taken at a future time $t_1 = t_2 + 1$, then Z_1 is most highly dependent with the Z_2 at a site that is $2\boldsymbol{\mu}_{\mathbf{V}_{11}} - \boldsymbol{\mu}_{\mathbf{V}_{22}}$ units away. The other plots can also be interpreted in the same fashion.

The difference between scenarios (a)-(c) is more pronounced in Figure 2 than in Figure 1. Although the direction of advection is the same in all three, the strength of dependence is remarkably different. Strong cross-covariance is sustained when $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ are highly positively correlated, as in (a). On the other hand, the cross-covariance weakens rapidly when $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ are highly negatively correlated, as in (c). The nonstationarity in time is revealed when looking at the heatmaps along the diagonal with values of the maximum cross-covariance decreasing as time goes farther away from zero. Furthermore, the location at which the maximum cross-covariance occurs changes and moves away from $\mathbf{h} = \mathbf{0}$. The single advection model is highly restrictive and does not allow for this nonstationarity in time of the cross-covariance. This implies that failing to acknowledge multiple advections can lead to overestimation or underestimation of the cross-covariances.

In the models so far, the cross-advections $\mathbf{V}_{ij}$, $i \neq j$, cannot be prescribed independently of the temporal locations, i.e., the transport behavior experienced in the cross-covariance remains

a function of the marginal advections $\mathbf{V}_{ii}$ and the temporal locations. An interesting problem is to find a form for $\mathbf{V}_{ij}$ such that

$$\mathbf{C}(\mathbf{h}, u) = \mathrm{E}_{\boldsymbol{\mathcal{V}}}[\{C_{ij}^S(\mathbf{h} - \mathbf{V}_{ij}u)\}_{i,j=1}^p], \tag{11}$$

is valid. Here $\boldsymbol{\mathcal{V}}$ is the vector of marginal and cross-advections. Suppose from (11) that we build the covariance matrix $\boldsymbol{\Sigma}$ as follows:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{C}_{11,\mathbf{V}_{11}} & \mathbf{C}_{12,\mathbf{V}_{12}} & \cdots & \mathbf{C}_{1p,\mathbf{V}_{1p}} \\ \mathbf{C}_{21,\mathbf{V}_{21}} & \mathbf{C}_{22,\mathbf{V}_{22}} & \cdots & \mathbf{C}_{2p,\mathbf{V}_{2p}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{p1,\mathbf{V}_{p1}} & \mathbf{C}_{p2,\mathbf{V}_{p2}} & \cdots & \mathbf{C}_{pp,\mathbf{V}_{pp}} \end{bmatrix} \in \mathbb{R}^{np \times np}, \tag{12}$$

where $\mathbf{C}_{ij,\mathbf{V}_{ij}} = \left(\mathrm{E}_{\boldsymbol{\mathcal{V}}} \left[C_{ij}^S \{ \mathbf{s}_l - \mathbf{s}_r - \mathbf{V}_{ij}(t_l - t_r) \} \right] \right)_{l,r=1}^n \in \mathbb{R}^{n \times n}$, for $i, j = 1, \ldots, p$. For the model in (11) to be valid, $\boldsymbol{\Sigma}$ has to be positive definite. By Theorem 1 in Ip and Li (2015), $\boldsymbol{\Sigma}$ is positive definite if and only if the $(np \times np)$ matrix $\mathbf{K}$ with entries $\mathbf{K} = \left(\mathbf{C}_{ii,\mathbf{V}_{ii}}^{-1/2} \mathbf{C}_{ij,\mathbf{V}_{ij}} \mathbf{C}_{jj,\mathbf{V}_{jj}}^{-1/2} \right)_{i,j=1}^p \in \mathbb{R}^{np \times np}$ is positive definite. Here $\mathbf{A}^{1/2}$ is the square root of a square matrix $\mathbf{A}$ such that $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$. This result gives us more control regarding the transport or advection behavior in the cross-covariances. However, it remains a challenge to more precisely characterize $\mathbf{V}_{ij}$ such that $\mathbf{K}$ is indeed positive definite.

Manipulating the transport behavior in the cross-covariances can also be done by defining new dimensions in space or time and allowing variable specific advections in those extra dimensions. The following theorem establishes a Lagrangian spatio-temporal cross-covariance model that augments the spatial dimensions.

Theorem 3 *Let $\mathbf{V}_{11}, \mathbf{V}_{22}, \ldots, \mathbf{V}_{pp}$ be random vectors on $\mathbb{R}^d$ and $\mathbf{V}'_{11}, \mathbf{V}'_{22}, \ldots, \mathbf{V}'_{pp}$ be random vectors on $\mathbb{R}^{d'}$. If $C_{ij}^S(\mathbf{h}, \mathbf{h}')$ is a valid purely spatial stationary cross-covariance function on $\mathbb{R}^{d+d'}$, then*

$$C_{ij}\{(\mathbf{h}, \mathbf{h}'_{ij}); t_1, t_2\} = \mathrm{E}_{\tilde{\boldsymbol{\mathcal{V}}}}\{C_{ij}^S(\mathbf{h} - \mathbf{V}_{ii}t_1 + \mathbf{V}_{jj}t_2, \mathbf{h}'_{ij} - \mathbf{V}'_{ii}t_1 + \mathbf{V}'_{jj}t_2)\}, \tag{13}$$

where $\mathbf{h}'_{ij} = \mathbf{s}'_{ii} - \mathbf{s}'_{jj}$, for $\mathbf{s}'_{ii}, \ldots, \mathbf{s}'_{pp} \in \mathbb{R}^{d'}$, and the expectation is taken with respect to the joint distribution of $\tilde{\boldsymbol{\mathcal{V}}} = \{(\mathbf{V}_{11}^{\top}, \mathbf{V}_{11}'^{\top}), (\mathbf{V}_{22}^{\top}, \mathbf{V}_{22}'^{\top}), \ldots, (\mathbf{V}_{pp}^{\top}, \mathbf{V}_{pp}'^{\top})\}^{\top}$, is a valid spatio-temporal cross-covariance function on $\mathbb{R}^{d+d'} \times \mathbb{R}$ provided that the expectation exists.

When $d' = 1$, $s'_{ii} \in \mathbb{R}$ can possibly be the altitude or the location in the z-axis at which variable i was taken and $V'_{ii} \in \mathbb{R}$ is the component of the advection velocity along that axis, $i = 1, \ldots, p$. Augmenting the temporal dimensions can also be done similarly. However, introducing a vector of temporal locations brings an unnatural physical interpretation to the Lagrangian transport phenomenon. Even in the classical univariate model of Cox and Isham (1988) in (3), the form of the Lagrangian shift when the temporal argument becomes a vector is nontrivial and has not yet been explored anywhere. A Lagrangian shift of the form $\mathbf{h} - \ddot{\mathbf{V}}\mathbf{u}$, where $\mathbf{h} \in \mathbb{R}^d$, $\mathbf{u} \in \mathbb{R}^{d''}$, and $\ddot{\mathbf{V}} \in \mathbb{R}^{d \times d''}$ is the advection velocity matrix, can be pursued when faced with multiple dimensions in time. The columns of $\ddot{\mathbf{V}}$ indicate the component of the transport velocity in every dimension of time. However, due to a lack of useful physical interpretation of Lagrangian models with an advection velocity matrix, we discuss only (13) and pursue the idea of advection velocity matrix in another work.

We simulate from (13) using the closed forms in (9) and (10) such that Z_1 is taken at $(\mathbf{s}^{\top}, 0.2)^{\top}$ and Z_2 at $(\mathbf{s}^{\top}, 0)^{\top}$, i.e., Z_1 and Z_2 have the same locations in the xy-axis but are separated 0.2 units away in the z-axis. Moreover, we set the advection of Z_1 and Z_2 in the xy-axis to be the same, i.e., $\boldsymbol{\mu}_{\mathbf{V}_{11}} = \boldsymbol{\mu}_{\mathbf{V}_{22}} = (0.1, 0.1)^{\top}$, but we augment them with different vertical components. In particular, we set V'_{11} and V'_{22} such that Z_1 gets transported downwards, while Z_2 gets transported upwards, i.e., $\mathrm{E}(V'_{11}) = -0.05$ and $\mathrm{E}(V'_{22}) = 0.05$. Furthermore, we consider three different strengths of dependence between the vertical components, namely (a) $\mathrm{cov}(V'_{11}, V'_{22}) = 0.9$, (b) $\mathrm{cov}(V'_{11}, V'_{22}) = 0$, and (c) $\mathrm{cov}(V'_{11}, V'_{22}) = -0.9$. The realizations and the corresponding heatmaps of the resulting bivariate Lagrangian spatio-temporal random field are shown in Panels (d)-(f) in Figures 1 and 2, respectively. From the figures, it can be seen that more sophisticated spatio-temporal dependencies spring from adding dimensions in

space. For instance, in Figure 2(d), the maximum cross-covariance occurs at $t_1 = t_2 = 2$ and not at $t_1 = t_2 = 1$, which is the case in Figure 2(a). This is expected because while Z_1 and Z_2 are colocated in $\mathbb{R}^2$, they are actually 0.2 units apart in $\mathbb{R}^3$. As Z_1 travels downwards and Z_2 travels upwards, they decrease the spatial separation between them, which increases their spatial dependence. However, this behavior is not manifested in Figure 2(e)-(f). This is because independent or negatively highly correlated advections bring about greater reduction in the maximum purely spatial cross-covariance. That is, Z_1 and Z_2 may be colocated in $\mathbb{R}^3$ after some time but variability from the mean advections distorts the transported purely spatial random fields, thereby reducing the maximum purely spatial cross-covariance attainable.

Another way to introduce multiple advections in the Lagrangian framework while remaining stationary in time is by using latent uncorrelated univariate transported purely spatial random fields, each influenced by different advection velocities. Salvaña and Genton (2020) suggested such models with latent transported purely spatial random fields that are second-order nonstationary. Their proposed model of course remains valid when the latent transported purely spatial random fields are second-order stationary. We formalize such models in the following theorem.

Theorem 4 *Let* $\mathbf{V}_r$, $r = 1, \ldots, R$, *be random vectors on* $\mathbb{R}^d$. *If* $\rho_r(\mathbf{h})$ *is a valid univariate stationary correlation function on* $\mathbb{R}^d$, *then*

$$\mathbf{C}(\mathbf{h}, u) = \sum_{r=1}^{R} \mathrm{E}_{\mathbf{V}_r} \left\{ \rho_r(\mathbf{h} - \mathbf{V}_r u) \right\} \mathbf{T}_r \tag{14}$$

is a valid spatio-temporal matrix-valued stationary cross-covariance function on $\mathbb{R}^d \times \mathbb{R}$, *for any* $1 \leq R \leq p$ *and* $\mathbf{T}_r$, $r = 1, \ldots, R$, *are positive semi-definite matrices.*

The model in (14) is the resulting Lagrangian spatio-temporal cross-covariance function of the following multivariate spatio-temporal process:

$$\mathbf{Z}(\mathbf{s}, t) = \mathbf{A}\mathbf{W}(\mathbf{s}, t) = \mathbf{A}[W_1(\mathbf{s} - \mathbf{V}_1 t), W_2(\mathbf{s} - \mathbf{V}_2 t), \ldots, W_R(\mathbf{s} - \mathbf{V}_R t)]^\top, \tag{15}$$

where $\mathbf{A}$ is $p \times R$ matrix and the components of $\mathbf{W}(\mathbf{s}, t) \in \mathbb{R}^R$ are independent but not identi-

cally distributed. Each component W_r has a univariate Lagrangian spatio-temporal stationary correlation function $\rho_r(\mathbf{h} - \mathbf{V}_r u)$, $r = 1, \ldots, R$. Here, $\mathbf{T}_r = \mathbf{a}_r \mathbf{a}_r^\top$, where $\mathbf{a}_r$ is the rth column of $\mathbf{A}$. Moreover, when $\mathbf{V}_1 = \mathbf{V}_2 = \cdots = \mathbf{V}_R = \mathbf{V}$, we return to the single advection velocity vector case and retrieve the Lagrangian spatio-temporal version of the linear model of coregionalization (LMC); see Gelfand et al. (2002) and Wackernagel (2003) for the discussion of such class of purely spatial cross-covariance functions.

3 Estimation

In this section, we outline a viable estimation procedure involving Lagrangian spatio-temporal cross-covariance functions with multiple advections. Suppose $\mathbf{Y} = \{\mathbf{Y}(\mathbf{s}_1, t_1)^\top, \ldots, \mathbf{Y}(\mathbf{s}_n, t_n)^\top\}^\top \in \mathbb{R}^{np}$ is an np-vector of multivariate spatio-temporal observations such that n is the total number of spatio-temporal locations and p is the number of variables. Assume that the mean function in (1) can be characterized as a linear combination of some covariates $X_1, X_2, \ldots, X_M$. Denote by $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \ldots, \boldsymbol{\beta}_p^\top)^\top \in \mathbb{R}^{Mp}$ the vector of mean parameters, where $\boldsymbol{\beta}_i = (\beta_{1,i}, \ldots, \beta_{M,i})^\top \in \mathbb{R}^M$, for $i = 1, \ldots, p$, and $\mathbf{X} = \{\mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}_1, t_1)^\top, \mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}_2, t_2)^\top, \ldots, \mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}_n, t_n)^\top\}^\top \in \mathbb{R}^{np \times Mp}$, where $\mathbf{X}(\mathbf{s}, t) = \{X_1(\mathbf{s}, t), \ldots, X_M(\mathbf{s}, t)\} \in \mathbb{R}^M$. The model in (1) becomes $\mathbf{Y}(\mathbf{s}, t) = \{\mathbf{I}_p \otimes \mathbf{X}(\mathbf{s}, t)^\top\} \boldsymbol{\beta} + \mathbf{Z}(\mathbf{s}, t)$. Furthermore, denote by $\boldsymbol{\Sigma}(\boldsymbol{\Theta})$ the $np \times np$ covariance matrix, parameterized by $\boldsymbol{\Theta} \in \mathbb{R}^q$ such that $\boldsymbol{\Sigma}(\boldsymbol{\Theta}) = [\{C_{ij}(\mathbf{s}_l - \mathbf{s}_r; t_l, t_r|\boldsymbol{\Theta})\}_{i,j=1}^p]_{l,r=1}^n$. The mean parameters, $\boldsymbol{\beta}$, and the cross-covariance parameters, $\boldsymbol{\Theta}$, are estimated via restricted maximum likelihood estimation (REML) which proceeds by maximizing, through an iterative procedure (Cressie and Lahiri, 1996; Shor et al., 2019, Equations 4 and 5),

$$
\begin{aligned}
l_{\mathrm{REML}}(\boldsymbol{\Theta}, \boldsymbol{\beta}; \mathbf{Y}) &= l(\boldsymbol{\Theta}, \boldsymbol{\beta}; \mathbf{Y}) + \frac{M}{2} \log(2\pi) + \frac{1}{2} \log |\mathbf{X}^\top \mathbf{X}| - \frac{1}{2} \log |\mathbf{X}^\top \boldsymbol{\Sigma}(\boldsymbol{\Theta})^{-1} \mathbf{X}|, \\
l(\boldsymbol{\Theta}, \boldsymbol{\beta}; \mathbf{Y}) &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}(\boldsymbol{\Theta})| - \frac{1}{2} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^\top \boldsymbol{\Sigma}(\boldsymbol{\Theta})^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}).
\end{aligned}
\tag{16}
$$

The iteration procedure begins with an initialization of $\boldsymbol{\beta}$ which we set to the ordinary least squares estimate (OLS), $\hat{\boldsymbol{\beta}}_{\mathrm{OLS}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y}$, and which we plug-in to the likelihood equations

13

above wherever $\boldsymbol{\beta}$ appears. Next, we estimate $\boldsymbol{\Theta}$ in a multi-step fashion. Splitting the estimation problem into several parts has been routinely employed when groups of parameters in the cross-covariance function can be estimated sequentially (Apanasovich and Genton, 2010; Bourotte et al., 2016; Qadir et al., 2021). Furthermore, it has been established that under some fairly general conditions, the multi-step MLE yields consistent estimators of the parameters in the last step (Murphy and Topel, 2002; Zhelonkin et al., 2012; Greene, 2014). Based on the properties of the Lagrangian spatio-temporal cross-covariance functions with multiple advections, maximizing (16) with respect to $\boldsymbol{\Theta}$ can be done in two steps. Suppose $\boldsymbol{\Theta} = \left\{ (\boldsymbol{\theta}_{\mathcal{M}}^S)^\top, (\boldsymbol{\theta}_{\mathcal{C}}^S)^\top, (\boldsymbol{\theta}^{ST})^\top \right\}^\top$, where $\boldsymbol{\theta}_{\mathcal{M}}^S$ is the vector of marginal purely spatial parameters, $\boldsymbol{\theta}_{\mathcal{C}}^S$ is the vector of purely spatial cross-covariance parameters, and $\boldsymbol{\theta}^{ST}$ is the vector of advection velocity parameters. The dimensions of each vector depend on the underlying purely spatial cross-covariance function and the assumed distribution of the advection velocity vectors. The elements of $\boldsymbol{\Theta}$ are estimated sequentially as follows:

1. Since the purely spatial marginal covariance functions can be derived, independently of the cross-covariance and advection parameters, the first step involves finding the vector $\hat{\boldsymbol{\theta}}_{\mathcal{M}}^S$ that maximizes the marginal purely spatial version of (16).

2. Since the cross-covariance functions are nonstationary in time, the cross-covariance and advection parameters need to be jointly estimated. That is, embedding the previously-found MLEs, $\hat{\boldsymbol{\theta}}_{\mathcal{M}}^S$, find $\hat{\boldsymbol{\theta}}_{\mathcal{C}}^S$ and $\hat{\boldsymbol{\theta}}^{ST}$ that maximize (16). Note that to ensure $\boldsymbol{\Sigma}_{\boldsymbol{\mathcal{V}}}$ remains positive definite, its entries are parameterized via its Cholesky decomposition.

Once $\hat{\boldsymbol{\Theta}}$ is obtained, we solve for $\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$, where $\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$ is the vector of estimates of the regression coefficients via generalized least squares (GLS) of the form $\hat{\boldsymbol{\beta}}_{\mathrm{GLS}} = \{\mathbf{X}^\top \boldsymbol{\Sigma}(\boldsymbol{\Theta})^{-1} \mathbf{X}\}^{-1} \mathbf{X}^\top \boldsymbol{\Sigma}(\boldsymbol{\Theta})^{-1} \mathbf{Y}$ and loop again through the above multi-step estimation of $\boldsymbol{\Theta}$. The procedure is terminated when a stopping criterion is reached.

When T is large and there is negligible dependence between observations that are very distant

in the temporal sense, according to Stein (2005b), (16) can be approximated as:

$$l_{\text{REML}}(\boldsymbol{\Theta}, \boldsymbol{\beta}; \mathbf{Y}) \approx l_{\text{REML}}(\boldsymbol{\Theta}, \boldsymbol{\beta}; \mathbf{Y}_{1,t^*}) + \sum_{j=t^*+1}^{T} l_{\text{REML}}(\boldsymbol{\Theta}, \boldsymbol{\beta}; \mathbf{Y}_j | \mathbf{Y}_{j-t^*, j-1}), \qquad (17)$$

where $\mathbf{Y}_{a,b} = (\mathbf{Y}_a^\top, \ldots, \mathbf{Y}_b^\top)^\top \in \mathbb{R}^{Npt^*}$, $\mathbf{Y}_t = \{\mathbf{Y}(\mathbf{s}_1, t)^\top, \ldots, \mathbf{Y}(\mathbf{s}_N, t)^\top\}^\top \in \mathbb{R}^{Np}$, for $a < b$, and t^* specifies the number of consecutive temporal locations included in the conditional distribution. Here $l_{\text{REML}}(\boldsymbol{\Theta}, \boldsymbol{\beta}; \mathbf{Y}_j | \mathbf{Y}_{j-t^*, j-1})$ is the log-likelihood function based only on the vector of spatio-temporal measurements $\mathbf{Y}_{j-t^*, j-1} = (\mathbf{Y}_{j-t^*}^\top, \ldots, \mathbf{Y}_{j-1}^\top)^\top$. In the subsequent sections, we do not perform any approximations since the T in our real data application is small and thus, full REML computations are feasible.

4 Simulation Study

Under the single advection setting and for $p = 1$, Salvaña and Genton (2021) showed that when the components of $\mathbf{V} \sim \mathcal{N}_d(\boldsymbol{\mu}_\mathbf{V}, \boldsymbol{\Sigma}_\mathbf{V})$ have zero mean, the same variance, and are uncorrelated, i.e., $\boldsymbol{\mu}_\mathbf{V} = \mathbf{0}$ and $\boldsymbol{\Sigma}_\mathbf{V} = \sigma_\mathbf{V}^2 \mathbf{I}_d$, for any $\sigma_\mathbf{V}^2 > 0$, the univariate Lagrangian spatio-temporal models with normal scale-mixture C^S reduce to univariate non-Lagrangian spatio-temporal isotropic covariance functions belonging to the Gneiting class (Gneiting, 2002). A breakdown on any of the above-mentioned restrictions dichotomizes univariate Lagrangian spatio-temporal models from their non-Lagrangian counterparts. Their simulation studies can be adopted for $p > 1$ and similar conclusions can be drawn.

When faced with a multivariate Lagrangian spatio-temporal random field with multiple advections, one can either fit multivariate Lagrangian spatio-temporal models with multiple advections, as in (8), or marginally fit univariate Lagrangian spatio-temporal models, as in (3), each with different advections. While it has been shown in the literature that multivariate modeling generally yields lower prediction errors as the presence of the other variables essentially increases the sample size of one variable (Genton and Kleiber, 2015; Zhang and Cai, 2015; Salvaña et al., 2020), it remains to be explored how the dependence between any two advection velocities affects

the accuracy of predictions. To answer this inquiry, we perform experiments to identify scenarios where multivariate Lagrangian spatio-temporal models with multiple advections are favorable over multiple univariate Lagrangian spatio-temporal ones.

Another objective of this section is to show the consequences of using a bivariate Lagrangian spatio-temporal covariance function with single advection to model a bivariate Lagrangian spatio-temporal random field with advections $\mathbf{V}_{11} \sim \mathcal{N}_d(\boldsymbol{\mu}_{\mathbf{V}_{11}}, \boldsymbol{\Sigma}_{\mathbf{V}_{11}})$ and $\mathbf{V}_{22} \sim \mathcal{N}_d(\boldsymbol{\mu}_{\mathbf{V}_{22}}, \boldsymbol{\Sigma}_{\mathbf{V}_{22}})$ such that $\boldsymbol{\mu}_{\mathbf{V}_{11}} = \boldsymbol{\mu}_{\mathbf{V}_{22}}$ and $\boldsymbol{\Sigma}_{\mathbf{V}_{11}} = \boldsymbol{\Sigma}_{\mathbf{V}_{22}}$ but $\mathbf{V}_{11} \neq \mathbf{V}_{22}$. Such bivariate random fields appear to be driven by single advection when in fact they are not. Again, the dependence between $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ introduces some interesting dynamics that may or may not be useful in modeling or prediction. In this section, we aim to expose such consequences.

4.1   Design

All the simulation studies are framed under the assumption that $d = 2$ and $p = 2$. Consider the following Lagrangian spatio-temporal models:

- M1: univariate Lagrangian spatio-temporal model in (4), where C^S is the Matérn covariance function, with purely spatial parameters σ, a, ν;

- M2: bivariate Lagrangian spatio-temporal model with single advection, i.e.,

$$C_{ij}(\mathbf{h}, u) = \frac{\rho \sigma_{ii} \sigma_{jj}}{\sqrt{|\mathbf{I}_d + \boldsymbol{\Sigma}_{\mathbf{V}} u^2|}} \mathcal{M}\{(\mathbf{h} - \boldsymbol{\mu}_{\mathbf{V}} u)^\top (\mathbf{I}_d + \boldsymbol{\Sigma}_{\mathbf{V}} u^2)^{-1} (\mathbf{h} - \boldsymbol{\mu}_{\mathbf{V}} u); a, \nu_{ij}\}, \quad i, j = 1, 2,$$

  where $\mathcal{M}(\mathbf{h}; a, \nu)$ is the univariate Matérn correlation with spatial scale and smoothness parameters a and ν, respectively; and

- M3: bivariate Lagrangian spatio-temporal model with multiple advections in (9) and (10), where C_{ij}^S is the parsimonious Matérn cross-covariance function, with purely spatial parameters $\rho, \sigma_{ij}, a, \nu_{ij}, i, j = 1, 2$.

We simulate 100 sample zero-mean bivariate spatio-temporal Gaussian random fields, $Z_1(\mathbf{s}, t)$ and $Z_2(\mathbf{s}, t)$, with purely spatial parameters as in Figure 1, containing $N = 529$ spatial observations,

on a 23×23 grid in the unit square, at time $t = 0, 1, \ldots, 5$. This number of spatio-temporal locations from which the data is generated is chosen to mimic the setup in the real spatio-temporal data to be analyzed in Section 5. In the following experiments, we remove the observations at $t = 5$ and use the remaining observations to fit the models. Upon obtaining the parameter estimates, we predict the previously removed values using simple kriging for the univariate model and simple cokriging for the bivariate models and report the error of the predictions measured by the Mean Square Error (MSE), $\text{MSE} = \frac{1}{pN} \sum_{i=1}^{p} \sum_{l=1}^{N} \{\hat{Z}_i(\mathbf{s}_l, 5) - Z_i(\mathbf{s}_l, 5)\}^2$.

In order to understand the effect of the dependence between the two advection velocities, $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$, on the accuracy of predictions, we simulate from M3 under different assumptions on the joint distribution of $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$, namely, $\boldsymbol{\mu}_{\mathbf{V}_{11}} = (0.1, 0.1)^\top$, $\boldsymbol{\mu}_{\mathbf{V}_{22}} = (-0.1, 0.1)^\top$ and (a) $\mathbf{V}_{11} = 0.9\mathbf{V}_{22}$, (b) $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ are independent, and (c) $\mathbf{V}_{11} = -0.9\mathbf{V}_{22}$, for several values of the spatial cross-correlation parameter, i.e., $\rho = \pm 0.3, \pm 0.6, \pm 0.9$. It is worthwhile to note that configuration (a) is the closest to the single advection model, in which case $\mathbf{V}_{11} = \mathbf{V}_{22}$, while (c) is the farthest. On the simulated values, we fit the true model, M3, and a simpler alternative model, M1.

To reveal the consequences of using a single advection model instead of a multiple advections model, we simulate from M3 with $\boldsymbol{\mu}_{\mathbf{V}_{11}} = \boldsymbol{\mu}_{\mathbf{V}_{22}} = (0.1, 0.1)^\top$ with different $\text{cov}(\mathbf{V}_{11}, \mathbf{V}_{22})$, namely, (d) $\text{cov}(\mathbf{V}_{11}, \mathbf{V}_{22}) = k \left(\begin{smallmatrix} 1 & 0.9 \\ 0.9 & 1 \end{smallmatrix} \right) \otimes \mathbf{I}_2$, (e) $\text{cov}(\mathbf{V}_{11}, \mathbf{V}_{22}) = k\mathbf{I}_4$, (f) $\text{cov}(\mathbf{V}_{11}, \mathbf{V}_{22}) = k \left(\begin{smallmatrix} 1 & -0.9 \\ -0.9 & 1 \end{smallmatrix} \right) \otimes \mathbf{I}_2$, for $k = 0.001, 0.1, 1$, and $\rho = \pm 0.3, \pm 0.6, \pm 0.9$. Scenarios (d) and (f) represent the highly positive and negative dependence between the corresponding components of $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$, respectively, while (e) establishes that $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ are independent. The parameter k acts as the scale parameter in time, with higher values of k implying faster drop in correlation in the temporal domain. Both the true model, M3, and a simpler model, M2, are fitted to the aforementioned simulated random fields.
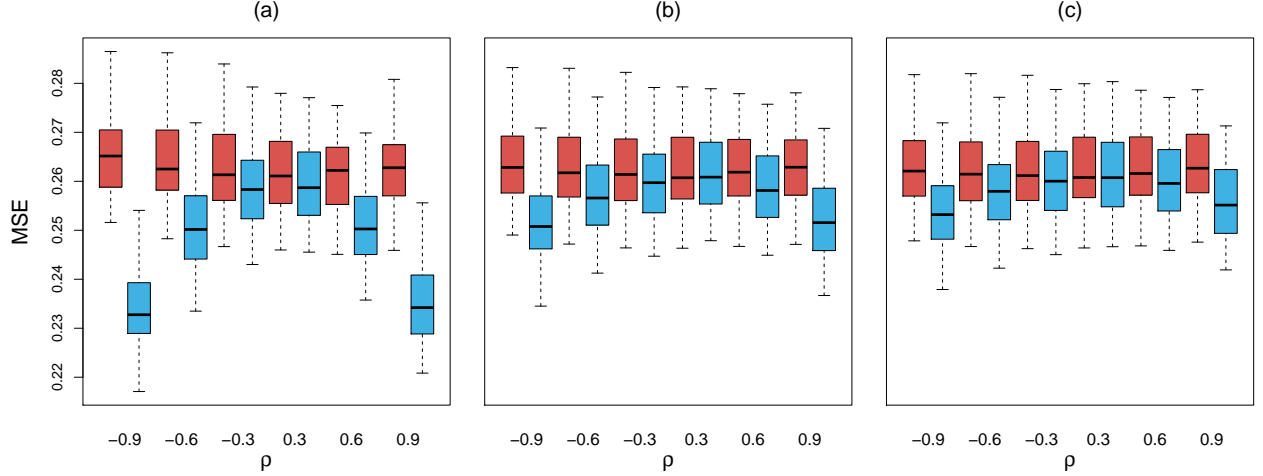
Figure 3: Boxplots of the MSEs under different assumptions on the joint distribution of $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$, namely, (a) $\mathbf{V}_{11} = 0.9\mathbf{V}_{22}$, (b) $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ are independent, and (c) $\mathbf{V}_{11} = -0.9\mathbf{V}_{22}$, when M1 (red) and M3 (blue) are fitted to data generated from M3 with different values of ρ.

4.2 Results and Analysis

In the first set of experiments, we compare the accuracy of the predictions between the two modeling paradigms, namely, marginally fitting multiple univariate Lagrangian spatio-temporal models with different advections (M1) and fitting a bivariate Lagrangian spatio-temporal model with multiple advections (M3), given that the true model is the latter. Figure 3 summarizes the boxplots of the MSEs under different strengths of dependence between $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ and different values of the colocated correlation ρ. It can be seen that the prediction performance of M1 is fairly similar regardless of the true value of ρ and the dependence between the different advections. Using a bivariate model (M3), on the other hand, improves predictions for nonzero ρ, with higher gains in accuracy as ρ gets farther away from 0. Furthermore, M3 yields more accurate predictions when the true model consists of a nonzero ρ with advection vectors that are highly dependent. When utilizing M1 over M3, one disregards a possible purely spatial variable dependence and possible dependence between the advections, which can help improve predictions.

In the second set of experiments, we study the effect of fitting a single advection model to a bivariate random field that appears to be simulated from a single advection model on two
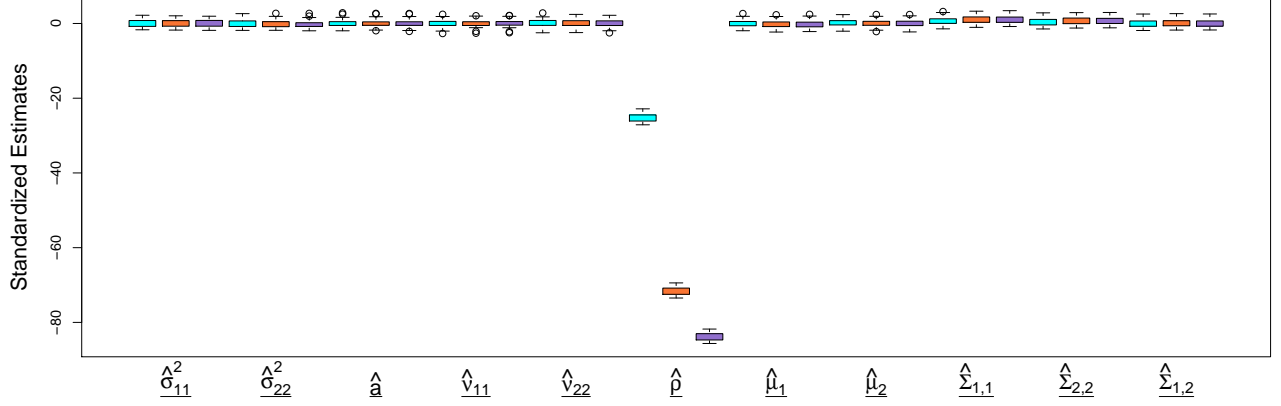
Figure 4: Boxplots of the centered and scaled MLEs of the parameters of M2 when it is fitted to data generated from M3 under scenarios (d), in cyan, (e), in orange, and (f), in purple, when $k = 0.001$ and $\rho = 0.6$.

fronts: 1) estimates of the common parameters; and 2) prediction performances of models M2 and M3. Figure 4 gives the summary of the centered and scaled MLEs of M2, denoted by $\underline{\hat{\boldsymbol{\Theta}}} = (\hat{\underline{\sigma}}^2_{11}, \hat{\underline{\sigma}}^2_{22}, \hat{\underline{a}}, \hat{\underline{\nu}}_{11}, \hat{\underline{\nu}}_{22}, \hat{\underline{\rho}}, \hat{\underline{\mu}}_1, \hat{\underline{\mu}}_2, \hat{\underline{\Sigma}}_{1,1}, \hat{\underline{\Sigma}}_{2,2}, \hat{\underline{\Sigma}}_{1,2})^\top$, where $\underline{\hat{\boldsymbol{\mu}}}_{\mathbf{V}} = (\hat{\underline{\mu}}_1, \hat{\underline{\mu}}_2)^\top$ and $\underline{\hat{\boldsymbol{\Sigma}}}_{\mathbf{V}} = \begin{pmatrix} \hat{\underline{\Sigma}}_{1,1} & \hat{\underline{\Sigma}}_{1,2} \\ \hat{\underline{\Sigma}}_{1,2} & \hat{\underline{\Sigma}}_{2,2} \end{pmatrix}$.

Clearly, the most impacted parameter is ρ. As the dependence between $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ goes from highly positive to highly negative, $\hat{\rho} \to 0$ for any value of ρ that we considered. This is expected for the following reason. The cross-covariance function in M2 is stationary in space and time, which implies that the purely spatial colocated correlation is constant. M3, on the other hand, has a cross-covariance function that is nonstationary in time with a time-varying and decreasing purely spatial colocated correlation as t_1 and t_2 move away from 0. Furthermore, the more negatively dependent $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ are, the faster the decline in the purely spatial colocated correlation; see Section 2 and Figure 2. Thus, when a model that can only handle a constant purely spatial colocated correlation parameter is fitted to a bivariate random field that possesses decreasing purely spatial colocated correlation, the optimization routine is required to find the best compromise for $\hat{\rho}$.

Additionally, we scrutinize the errors, summarized via boxplots in Figure 5, when fitting M2 and M3 to data generated from M3. It can be seen that M2 does not predict as well as M3 in all
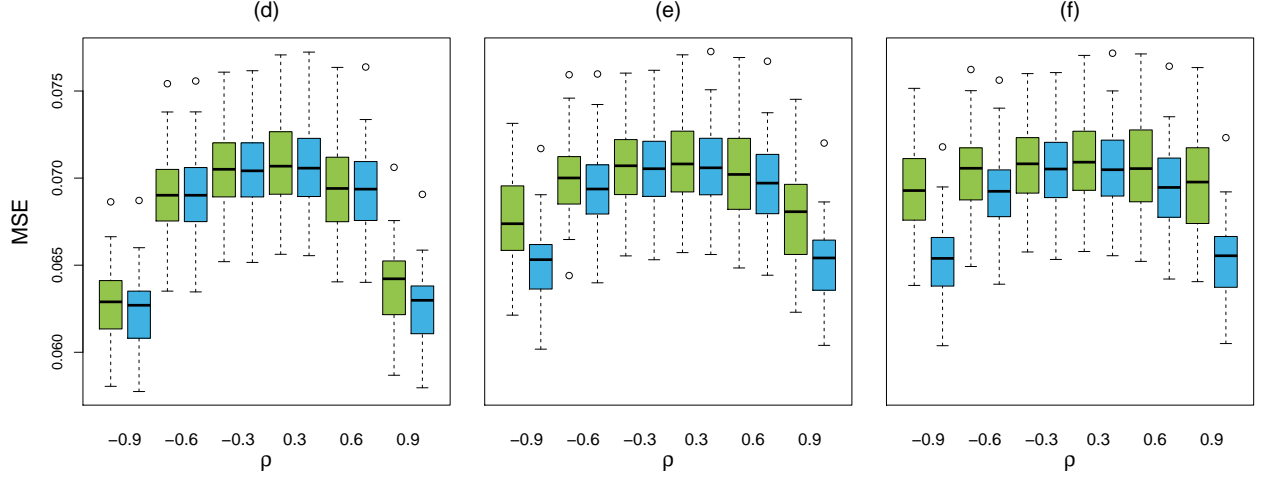
19

Figure 5: 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 $k = 0.001$.

cases but the difference in accuracy is the highest under scenario (f). This is because in scenario (f), M2 severely overvalues the cross-covariances compared to the ones obtained by M3, resulting in a grave misspecification of spatio-temporal dependence between Z_1 and Z_2. All in all, in both experiments, the results demonstrate that neglecting the multiple advection phenomenon leads to substantial losses in prediction accuracy.

5 Application to Particulate Matter Data

Wind is recognized as a major driver of pollutant transport in the atmosphere. The Modern-Era Retrospective Analysis for Research and Applications v.2 (MERRA-2) simulated particulate matter (PM) concentrations over Saudi Arabia plotted in Figure 6 demonstrate such transport behavior (Buchard et al., 2017). Any physical model on PM factors in mechanisms of transport such as the wind fields at different levels of the troposphere (Kallos et al., 2007), jet streams, pyroconvective events, and boundary layer turbulence (National Research Council, 2010). The transport behavior enables the propagation of suspended PM to locations far from the original source and this often results to higher PM concentrations on its path of transport.

The modeling of PM concentrations is an active field of research in computational fluid

dynamics (CFD). In that area of study, emphasis is placed on airflow motion and turbulence in determining PM concentrations at any location in space and time (Zhang and Chen, 2007; Katra et al., 2016). CFD involves tracking a large number of particles and employs highly specialized models, e.g., particle concentration equations coupled with momentum and turbulence equations, which need to be fed with various model input parameters (Knox, 1974). Lack of domain expert knowledge impedes adoption of these physically consistent models by non-experts in CFD.

The proliferation of PM measurements, which can be obtained from several sources such as ground stations, online databases (e.g., NASA Earthdata website), satellite remote sensing, lidar networks, and other outdoor monitoring systems, has launched a new wave of statistical modeling methodologies focused on modeling and predicting particulate matter concentrations (Shao et al., 2011). The list includes Bayesian models (Sahu et al., 2006; Calder, 2008), generalized additive models (Paciorek et al., 2009; Munir et al., 2013), geographically weighted regression (Van Donkelaar et al., 2016), stochastic partial differential equations (Cameletti et al., 2013), time series models (Goyal et al., 2006), and machine learning models (Mehdipour et al., 2018). These modeling approaches typically require only historical measurements of PM and other associated variables. In this work, we tackle the problem on hand by combining the strengths of spatio-temporal Gaussian geostatistical modeling and the concept of transport in CFD through the class of Lagrangian spatio-temporal models.

5.1 Saudi Particulate Matter Dataset

The dataset used in the present study was obtained from NASA Earthdata and it contains measurements of PM (black carbon) concentrations, measured every 3 hours at two different pressure levels, 880 hPa ($\sim$ 1.2 km above sea level) and 985 hPa ($\sim$ 0.7 km above sea level), on a regular grid with pixel size $0.5° \times 0.625°$, with no missing observations. A logarithmic transformation is applied to obtain close to normally distributed measurements; see Paciorek et al. (2009), Sahu (2012), and Cameletti et al. (2013) for similar treatments. Additionally, the spatial coordinates
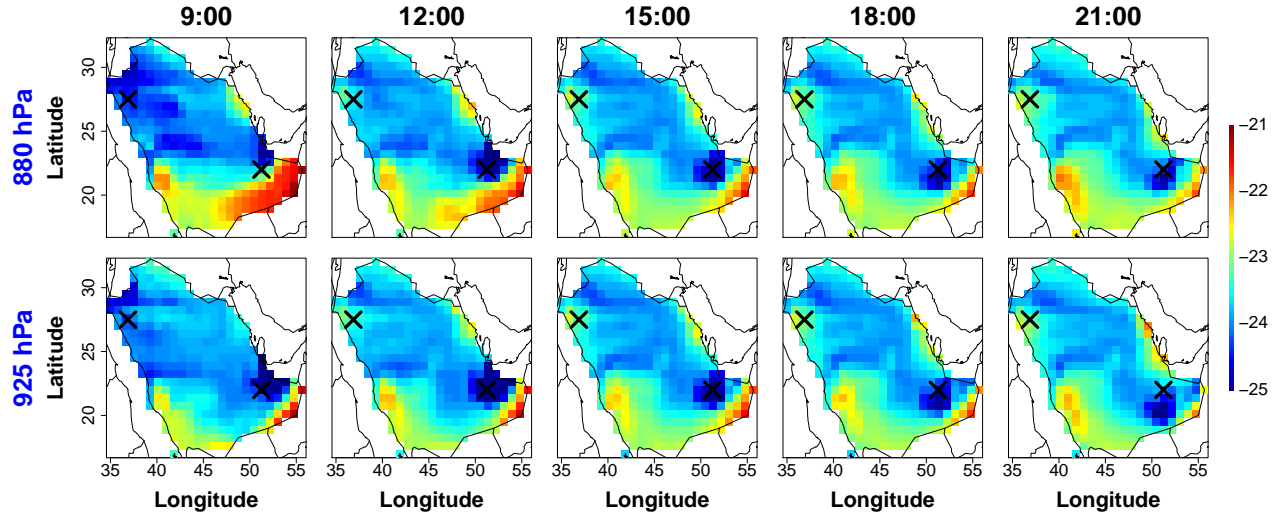
Figure 6: MERRA-2 simulated particulate matter concentrations on January 18, 2019, 9:00 to 21:00, at two pressure levels in log scale of $\mu g/m^3$. Two sites are marked with "×" to aid detection of transport behavior.

originally in degree latitude and longitude are converted to their corresponding Cartesian coordinates. This means that the distance between any two locations is measured in kilometers (km), with an overall minimum distance between any two stations equal to 16.9 km.

In this study, we consider only the five consecutive 3-hourly measurements starting from 9:00 to 21:00 on January 18, at 550 locations over Saudi Arabia, over the period 1980-2019. In this work, we assume each year as an independent replicate coming from the same underlying spatio-temporal Gaussian process. Figure 6 maps the log PM concentrations over the spatial and temporal domains under consideration for the year 2019. Using the reference locations as aid, the transport behavior can be detected, with a seemingly South East direction of movement at both pressure levels.

The spatio-temporal modeling is carried out following the linear model in (1), where $Y_1(\mathbf{s}, t)$ and $Y_2(\mathbf{s}, t)$ are the log PM measurements at 880 hPa and 925 hPa, respectively, at spatial location $\mathbf{s} \in \mathcal{D}$, $\mathcal{D}$ being the region inside the borders of Saudi Arabia, and temporal location $t \in \mathcal{T}$, $\mathcal{T} = \{0, \ldots, 4\}$. Moreover, we make the following assumptions on $\boldsymbol{\mu}(\mathbf{s}, t)$ and $\mathbf{Z}(\mathbf{s}, t)$. First, the mean function $\boldsymbol{\mu}(\mathbf{s}, t)$ can be modeled by a set of covariates comprising of the relative humidity and temperature at spatio-temporal location $(\mathbf{s}, t)$. Both covariates are accessible in the

same database where the log PM measurements were obtained. Second, the resulting residuals $\mathbf{Z}(\mathbf{s}, t)$ are assumed to be Lagrangian or transported.

NASA Earthdata also provides simulated wind vectors at the same spatial and temporal resolutions as the log PM concentrations at each pressure level. These simulated wind vectors were used as inputs by MERRA-2 to simulate the raw PM concentrations (Randles et al., 2017; Ukhov et al., 2020). In Figure 7, we plot the pairwise empirical bivariate distributions of the components of the wind vectors in m/s at the two pressure levels. A plausible estimate of $\boldsymbol{\mu_{\mathcal{V}}}$ is the empirical mean vector. When inspecting the second plot in the first row and the last plot in the third row, one may hypothesize that the bivariate spatio-temporal random field is from a single advection model since the bivariate distributions of the wind vectors of the two variables appear to be identical. However, as we have shown in Section 4, the appearance of a single advection phenomenon does not guarantee that the true model is a single advection model. Another observation is that the corresponding components of the advection vectors of the two variables are not equal but are highly positively dependent as shown in the third plot in the first row and the last plot in the second row, a scenario that we also considered and studied in Section 4. Thus, a multiple advections model may be the most suitable for this dataset.

5.2 Models

We consider six different spatio-temporal cross-covariance functions with bivariate parsimonious Matérn purely spatial margins, three of which were utilized in Section 4, namely, M1, M2, and M3, and the other three are as follows:

- M4: bivariate Lagrangian spatio-temporal model with variable specific advection in (13), with $d' = 1$;

- M5: bivariate Lagrangian LMC in (14); and

- M6: bivariate non-Lagrangian fully symmetric Gneiting-Matérn, i.e.,

$$C_{ij}(\mathbf{h}, u) = \frac{\rho \sigma_{ii} \sigma_{jj}}{\alpha |u|^{2\xi} + 1} \mathcal{M} \left\{ \frac{\mathbf{h}}{(\alpha |u|^{2\xi} + 1)^{b/2}}; a, \nu_{ij} \right\}, \quad i, j = 1, 2,$$
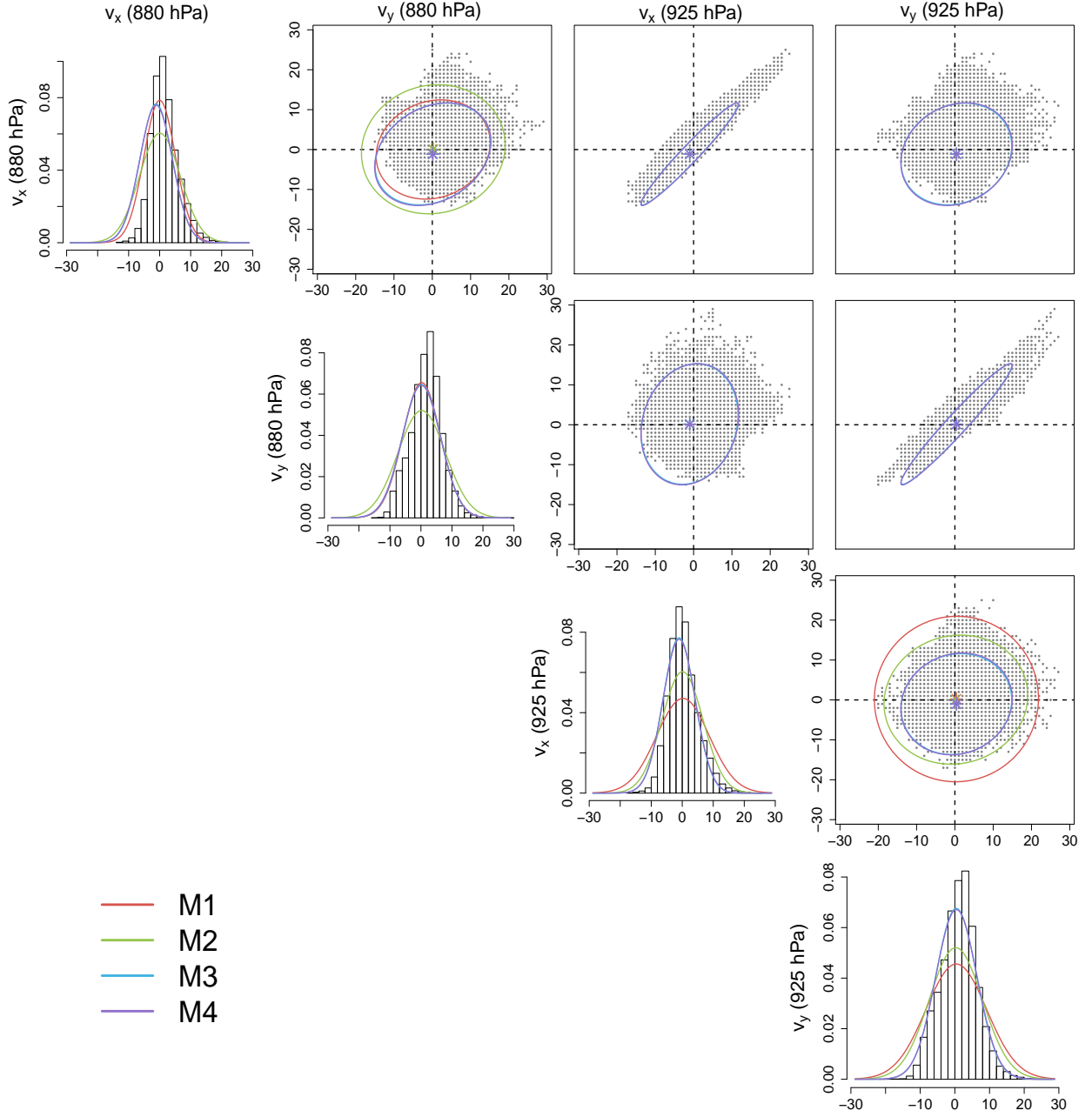
Figure 7: Empirical bivariate distributions of the MERRA-2 simulated wind vectors (in m/s) over the spatio-temporal domain under consideration. Here v_x (v_y) denotes the first (second) component or the component along the x-axis (y-axis) of the wind vector. Superimposed are $\hat{\boldsymbol{\mu}}_{\mathbf{V}}$ (in m/s) in M1, M2 (in red and green), $\hat{\boldsymbol{\mu}}_{\boldsymbol{\mathcal{V}}}$ (in m/s) in M3 (blue), and $\hat{\boldsymbol{\mu}}_{\tilde{\boldsymbol{\mathcal{V}}}}$ (in m/s) in M4 (purple), respectively. The lines in the bivariate plots enclose the 95% probability region. Only M3 and M4 are applicable in the 2×2 panels from the top right corner. Here M3 and M4 almost coincide, with M3 being the best performing model according to the BIC*.

where $\alpha > 0$ and $\xi \in (0, 1]$ describe the temporal range and smoothness, respectively (Bourotte et al., 2016). The parameter $b \in [0, 1]$, also called the "nonseparability parameter", represents the strength of the spatio-temporal interaction.

Preliminary data analyses suggest that there is an anisotropic behavior in the marginal and cross-covariance structures. Hence, instead of evaluating the models at $\mathbf{h}$, we use $\mathbf{Rh}$, where $\mathbf{R} = \begin{pmatrix} R_1 \cos R_3 & R_1 \sin R_3 \\ -R_2 \sin R_3 & R_2 \cos R_3 \end{pmatrix}$ is the anisotropy matrix parameterized by the anisotropy scale parameters, R_1 and R_2, and the anisotropy angle parameter, R_3; see Paciorek and Schervish (2006) and Hewer et al. (2017) for discussions on this modeling approach and more general parameterizations of $\mathbf{R}$.

5.3 Results and Discussions

Through the procedure described in Section 3, we fit the models using all 40 spatio-temporal random fields and predict the concentrations at all spatial locations at $t = 6$ for each year. The anisotropy parameters included in $\mathbf{R}$ are estimated in the first step alongside the marginal purely spatial parameters. Figure 8 shows the resulting map of predicted log concentrations under the different models and contrasts the predictions with the map of the true log concentrations for the year 2019. Table 1 reports the log-likelihood, MSE, and Akaike (AIC*) and Bayesian information criteria (BIC*), where $\text{AIC}^* = -2l_{\text{REML}}(\widehat{\boldsymbol{\Theta}}_1, \widehat{\boldsymbol{\Theta}}_2, \hat{\boldsymbol{\beta}}_{\text{GLS}}) + 2q$ and $\text{BIC}^* = -2l_{\text{REML}}(\widehat{\boldsymbol{\Theta}}_1, \widehat{\boldsymbol{\Theta}}_2, \hat{\boldsymbol{\beta}}_{\text{GLS}}) + q \log(n)$. Here $l_{\text{REML}}(\widehat{\boldsymbol{\Theta}}_1, \widehat{\boldsymbol{\Theta}}_2, \hat{\boldsymbol{\beta}}_{\text{GLS}})$ is the value of the log-likelihood function at the second estimation step with parameter estimates $\widehat{\boldsymbol{\Theta}}_2$ while fixing the parameters $\widehat{\boldsymbol{\Theta}}_1$ obtained at the first estimation step. The number of parameters in each model is also indicated in Table 1.

The results show that models with multiple advections, i.e., M1, M3, and M4, obtain superior prediction performance compared to those with single (M2) or no advection at all (M6). Furthermore, all the Lagrangian spatio-temporal models, M1-M4, outperform the non-Lagrangian spatio-temporal benchmark model, M6, in prediction. This is due to the Lagrangian models being able to capture the transport behavior that is seen to have occurred when connecting
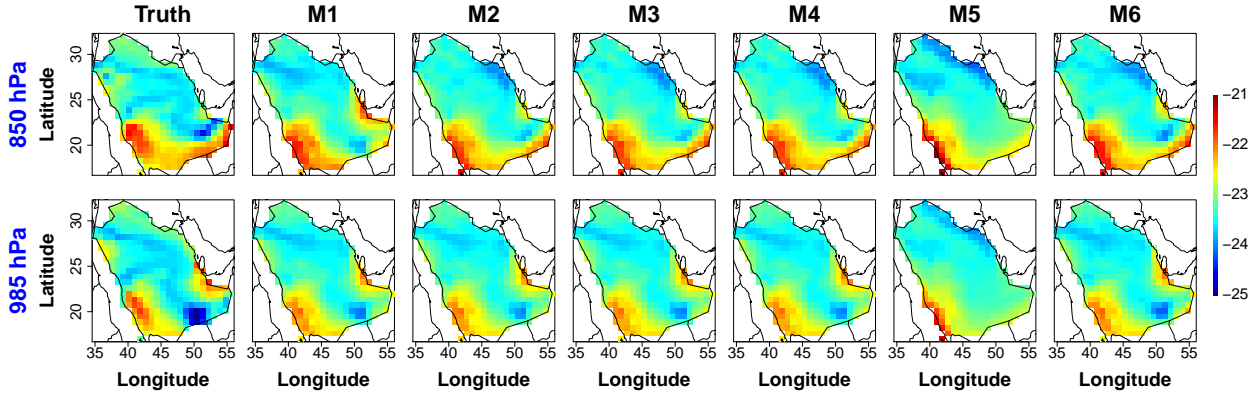
Figure 8: Map of predicted log residuals particulate matter concentrations on January 19, 2019, 0:00, under the six different models.

Table 1: In-sample (log-likelihood, AIC*, and BIC*) and out-of-sample (MSE) scores. The lower the values, the better. The best scores are given in bold.

| Model | In-Sample | | | Out-of-Sample | No. of | Computation |
	log-likelihood	AIC*	BIC*	MSE	parameters	Time (secs)
M1	$461,825$	$-923,614$	$-923,439$	0.0521	18	$10,857$
M2	$479,159$	$-958,290$	$-958,156$	0.0546	14	$13,446$
M3	$484,070$	$-968,094$	$\mathbf{-967,873}$	0.0516	23	$68,784$
M4	$\mathbf{484,150}$	$\mathbf{-968,210}$	$-967,777$	$\mathbf{0.0514}$	45	$227,444$
M5	$470,852$	$-941,658$	$-941,437$	0.1602	23	$15,059$
M6	$477,480$	$-954,936$	$-954,821$	0.0601	12	$8,391$

Figure 6 and the first column of Figure 8. The Lagrangian LMC model, M5, appears to be the worst performing model in terms of prediction. Even though M5 accommodates multiple advections, it does so in an unnatural way. Furthermore, the drawbacks possessed by the purely spatial LMC (Gneiting et al., 2010) is inherited by its Lagrangian spatio-temporal extension. M4 has the best performance among all the models in 3 out of 4 metrics, but at the cost of more parameters to be estimated and more computation time. The BIC* supports the usage of M3 as its performance is very close to M4 but with less parameters to estimate. Based on these results and the computational efficiency of the models, M3 emerges as the best model overall.

To visually validate the fit of the best performing model in terms of BIC*, we plot the empirical spatio-temporal marginal and cross-covariances of the resulting log residuals under M3 in Figure 9. The empirical plots of the resulting residuals reveal some anisotropy in both the marginal and cross-covariances. Furthermore the spatial lags at which the maximum values of
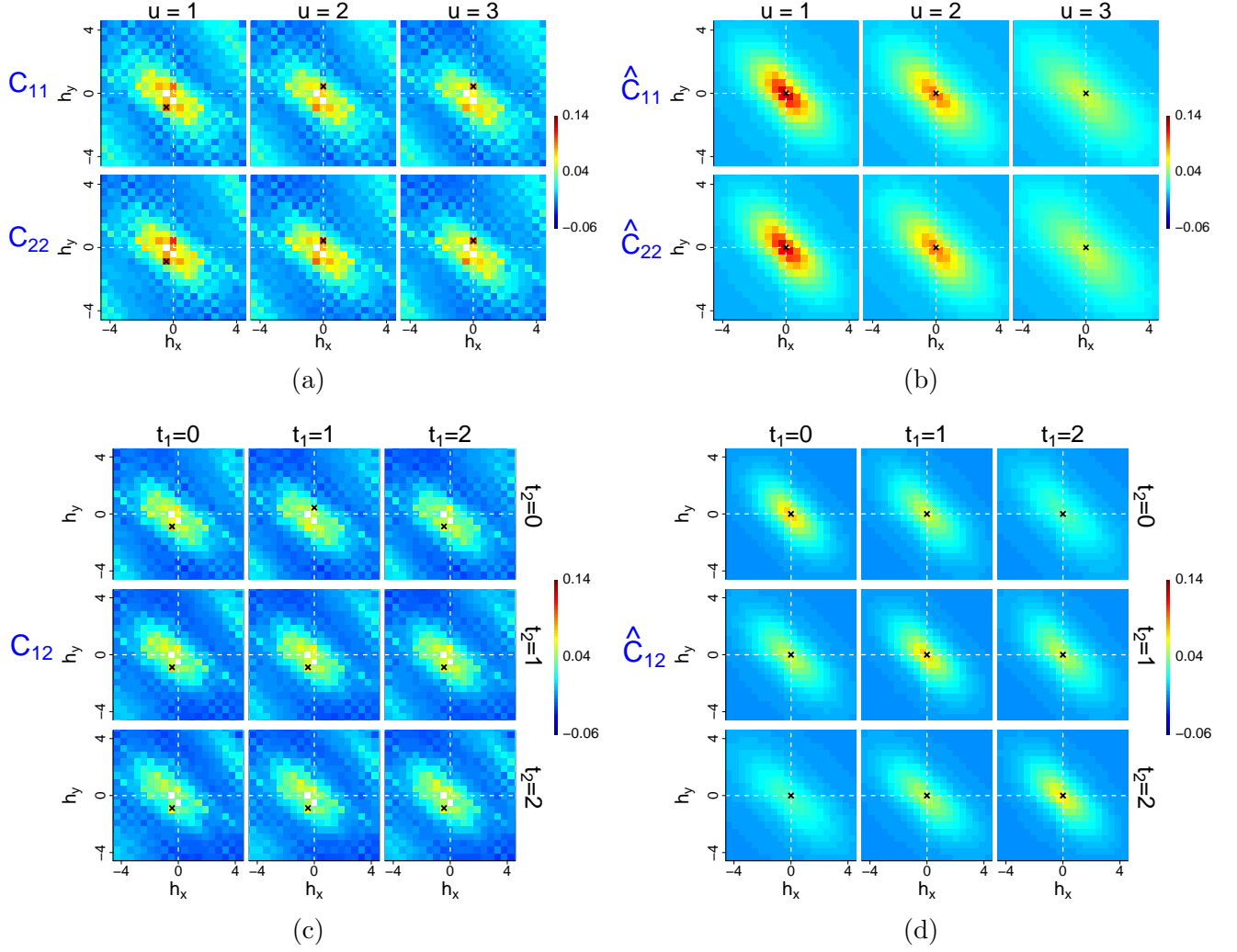
Figure 9: Empirical (C_{11} and C_{22}) and fitted ($\hat{C}_{11}$ and $\hat{C}_{22}$) marginals and empirical (C_{12}) and fitted ($\hat{C}_{12}$) cross-covariance values at every spatial and temporal lag are plotted as heatmaps. The spatial resolution is in $\times 10^2$ kilometers while the temporal resolution is 3 hours. The spatial lags at which the maximum value of the marginal and cross-covariance occur are marked with "$\times$".

the marginal and cross-covariances occur, i.e., locations marked with "$\times$", hover around $\mathbf{h} = \mathbf{0}$. These observations provide further support on including anisotropy parameters in the models and they also suggest that there may be no purely spatial variable asymmetry at any time. The fitted marginal and cross-covariances are juxtaposed beside their empirical counterparts. Clearly, M3 closely mirrors the spatio-temporal dependence profile of the real bivariate spatio-temporal dataset.

To further validate the in-sample and out-of-sample performance of the models, we boot-
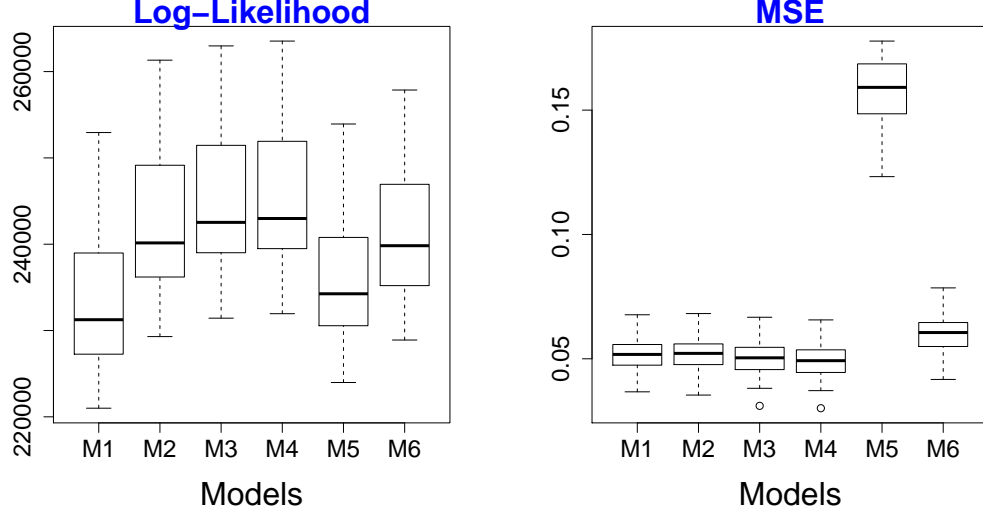
Figure 10: The log-likelihood and MSE values of the different models under the bootstrap study.
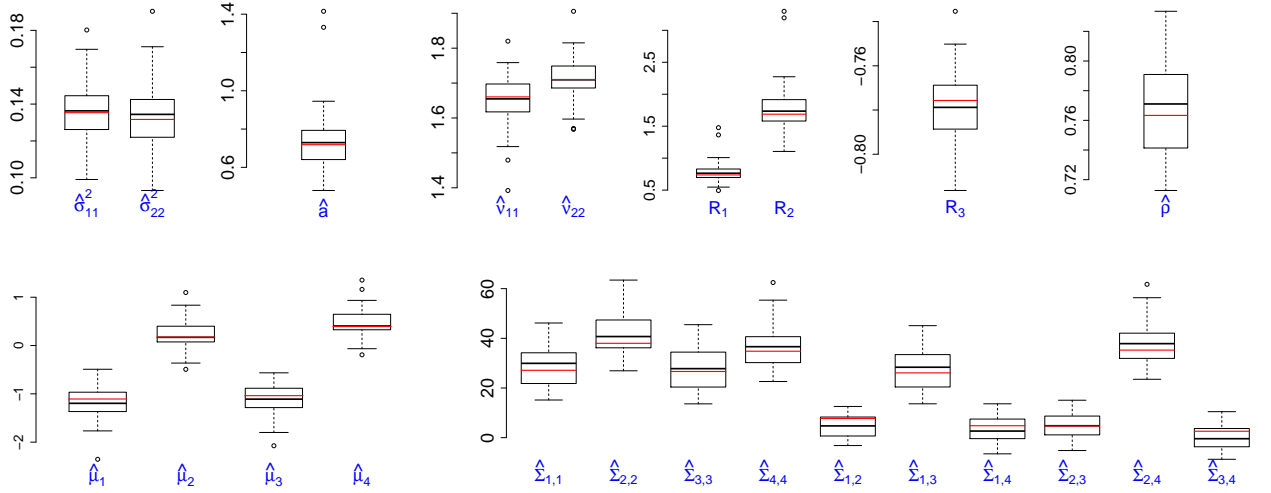

Figure 11: Boxplots of estimated parameters of M3. In red, the estimates obtained using all the 40 spatio-temporal random fields are plotted. Here $\mathbf{V}_{11} = \{V_x(880 \text{ hPa}), V_y(880 \text{ hPa})\}^\top$, $\mathbf{V}_{22} = \{V_x(925 \text{ hPa}), V_y(925 \text{ hPa})\}^\top$, $\hat{\boldsymbol{\mu}}_{\boldsymbol{\mathcal{V}}} = \{\text{E}(\mathbf{V}_{11})^\top, \text{E}(\mathbf{V}_{22})^\top\}^\top = (\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3, \hat{\mu}_4)^\top$ (in m/s), and $\hat{\boldsymbol{\Sigma}_{\boldsymbol{\mathcal{V}}}} = \begin{bmatrix} \text{cov}(\mathbf{V}_{11}, \mathbf{V}_{11}) & \text{cov}(\mathbf{V}_{11}, \mathbf{V}_{22}) \\ \text{cov}(\mathbf{V}_{11}, \mathbf{V}_{22}) & \text{cov}(\mathbf{V}_{22}, \mathbf{V}_{22}) \end{bmatrix}$ (in m^2/s^2) such that $\text{cov}(\mathbf{V}_{11}, \mathbf{V}_{11}) = \begin{pmatrix} \hat{\Sigma}_{1,1} & \hat{\Sigma}_{1,2} \\ \hat{\Sigma}_{1,2} & \hat{\Sigma}_{2,2} \end{pmatrix}$, $\text{cov}(\mathbf{V}_{11}, \mathbf{V}_{22}) = \begin{pmatrix} \hat{\Sigma}_{1,3} & \hat{\Sigma}_{1,4} \\ \hat{\Sigma}_{2,3} & \hat{\Sigma}_{2,4} \end{pmatrix}$, and $\text{cov}(\mathbf{V}_{22}, \mathbf{V}_{22}) = \begin{pmatrix} \hat{\Sigma}_{3,3} & \hat{\Sigma}_{3,4} \\ \hat{\Sigma}_{3,4} & \hat{\Sigma}_{4,4} \end{pmatrix}$.

strap through the 40 available spatio-temporal replicates, which we assume are independent and identically distributed, by performing 100 rounds of estimation and prediction using a subset of randomly sampled 30 spatio-temporal random fields. Figure 10 shows the boxplots of the log-likelihood and MSE values. Figure 11 displays the boxplots and, consequently, the uncertainties

of the estimated parameters of M3, such that highlighted by a red line are the parameters in the model fitting performed to obtain the results in Table 1. It can be seen that the estimated parameters are substantiated by the empirical marginal and cross-covariances shown in Figure 9.

Additionally, the estimates associated with the transport phenomenon are validated against the real transport behavior observed in the spatio-temporal domain under study. In Figure 7, the estimated distribution of the advection velocity vectors in M1-M4 are superimposed on the real wind vectors. Visually, it can be seen that the advection parameters obtained by models M3 and M4 adequately captures the empirical distribution. Futhermore, M1 and M2 obtains higher values for the marginal variance parameters of the components of $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ compared to those in M3 and M4. This may be due to the unaccounted dependence between $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ when using models M1 and M2. Although the parameter estimates in Figure 11 show that $\hat{\mathbf{V}}_{11}$ and $\hat{\mathbf{V}}_{22}$ are almost identical, it is still paramount to consider and use a multiple advections model because from the real wind data, the wind vectors in 880 hPa and 925 hPa are not identical.

6   Conclusion

We successfully pursued the development of Lagrangian spatio-temporal cross-covariance functions with multiple advections. The proposed framework offers a suite of data-driven models which is more flexible and realistic. We also outlined an estimation procedure to get sensible estimates of all parameters. We showed through numerical experiments involving bivariate spatio-temporal random fields how failing to account for multiple advections produces poor predictions and how utilizing simpler univariate Lagrangian spatio-temporal models cannot capitalize on the purely spatial variable dependence. With wind recognized as the main driver of the transport behavior, the proposed models, along with other benchmark spatio-temporal models, were tested on PM concentrations over Saudi Arabia. Indeed, a model under our proposed class emerged as the best model and its estimate of the multivariate distribution of the advection velocities checks out with the prevailing behavior of wind in the spatio-temporal domain under consideration.

Appendix

Proof of Theorem 1 Let $\boldsymbol{\lambda}_l \in \mathbb{R}^p$, $l = 1, \ldots, n$. Then:

$$
\sum_{l=1}^{n} \sum_{r=1}^{n} \boldsymbol{\lambda}_l^\top \mathbf{C}(\mathbf{s}_l, \mathbf{s}_r; t_l, t_r)\boldsymbol{\lambda} = \sum_{l=1}^{n} \sum_{r=1}^{n} \boldsymbol{\lambda}_l^\top \mathrm{E}_{\boldsymbol{\mathcal{V}}} \left(\left[C_{ij}^S \left\{ (\mathbf{s}_l - \mathbf{V}_{ii}t_l) - (\mathbf{s}_r - \mathbf{V}_{jj}t_r) \right\} \right]_{i,j=1}^{p} \right) \boldsymbol{\lambda}_r
$$

$$
= \mathrm{E}_{\boldsymbol{\mathcal{V}}} \left[\sum_{l=1}^{n} \sum_{r=1}^{n} \boldsymbol{\lambda}_l^\top \left\{ C_{ij}^S (\mathbf{s}_l - \mathbf{s}_r - \mathbf{V}_{ii}t_l + \mathbf{V}_{jj}t_r) \right\}_{i,j=1}^{p} \boldsymbol{\lambda}_r \right] \geq 0
$$

for all $n \in \mathbb{Z}^+$ and $\{(\mathbf{s}_1, t_1), \ldots, (\mathbf{s}_n, t_n)\} \in \mathbb{R}^d \times \mathbb{R}$. The last inequality follows from the assumption that $\mathbf{C}^S$ is a valid purely spatial matrix-valued stationary cross-covariance function with variable asymmetry on $\mathbb{R}^d$; see Li and Zhang (2011), Genton and Kleiber (2015), and Qadir et al. (2021) for discussions on this class of cross-covariance functions. $\qquad\square$

Proof of Theorem 2 See Supplementary Materials.

Proof of Theorem 3 The validity is guaranteed by the dimension expansion approach in Bornn et al. (2012). $\qquad\square$

Proof of Theorem 4 The validity of (14) is established as it is the resulting spatio-temporal cross-covariance function of the process in (15). $\qquad\square$

Derivation of Equation (8)

$$
C_{ij} \left(\mathbf{h} - \mathbf{V}_{ii}t_l + \mathbf{V}_{jj}t_r \right) = C_{ij} \left\{ \mathbf{h} - \mathbf{V}_{ii} \left(m + \frac{u}{2} \right) + \mathbf{V}_{jj} \left(m - \frac{u}{2} \right) \right\} = C_{ij} \left\{ \mathbf{h} - \overline{\mathbf{V}}_{ij}u + (\mathbf{V}_{jj} - \mathbf{V}_{ii}) m \right\},
$$

where $u = t_l - t_r$, $m = \frac{t_l + t_r}{2}$, and $\overline{\mathbf{V}}_{ij} = \frac{\mathbf{V}_{ii} + \mathbf{V}_{jj}}{2}$. $\qquad\square$

Supplementary Materials and Codes

The Supplementary Materials and codes can be found in this website: https://github.com/ladybug-hash/multiple-advections

References

Alegría, A., Porcu, E., Furrer, R., and Mateu, J. (2019). Covariance functions for multivariate gaussian fields evolving temporally over planet earth. *Stochastic Environmental Research and Risk Assessment*, 33(8):1593–1608.

Apanasovich, T. V. and Genton, M. G. (2010). Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika*, 97(1):15–30.

Bornn, L., Shaddick, G., and Zidek, J. V. (2012). Modeling nonstationary processes through dimension expansion. *Journal of the American Statistical Association*, 107(497):281–289.

Bourotte, M., Allard, D., and Porcu, E. (2016). A flexible class of non-separable cross-covariance functions for multivariate space–time data. *Spatial Statistics*, 18:125–146.

Buchard, V., Randles, C., Da Silva, A., Darmenov, A., Colarco, P., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A., Ziemba, L., et al. (2017). The MERRA-2 aerosol reanalysis, 1980 onward. Part II: Evaluation and case studies. *Journal of Climate*, 30(17):6851–6872.

Calder, C. A. (2008). A dynamic process convolution approach to modeling ambient particulate matter concentrations. *Environmetrics*, 19(1):39–48.

Cameletti, M., Lindgren, F., Simpson, D., and Rue, H. (2013). Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Advances in Statistical Analysis*, 97(2):109–131.

Christakos, G. (2017). *Spatiotemporal Random Fields: Theory and Applications*. Elsevier.

Cox, D. and Isham, V. (1988). A simple spatial-temporal model of rainfall. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 415:317–328.

Cressie, N. and Lahiri, S. N. (1996). Asymptotics for REML estimation of spatial covariance parameters. *Journal of Statistical Planning and Inference*, 50(3):327–341.

Ezzat, A. A., Jun, M., and Ding, Y. (2018). Spatio-temporal asymmetry of local wind fields and its impact on short-term wind forecasting. *IEEE Transactions on Sustainable Energy*, 9(3):1437–1447.

Gelfand, A. E., Schmidt, A. M., and Sirmans, C. (2002). Multivariate spatial process models: conditional and unconditional Bayesian approaches using coregionalization. unpublished.

Genton, M. G. and Kleiber, W. (2015). Cross-covariance functions for multivariate geostatistics (with discussion). *Statistical Science*, 30(2):147–163.

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space–time data. *Journal of the American Statistical Association*, 97(458):590–600.

Gneiting, T., Genton, M. G., and Guttorp, P. (2007). Geostatistical space-time models, stationarity, separability, and full symmetry. In Finkenstaedt, B., Held, L., and Isham, V., editors, *Statistics of Spatio-Temporal Systems. Monographs in Statistics and Applied Probability*, chapter 4, pages 151–175. Chapman & Hall/CRC Press, Boca Raton, Florida.

Gneiting, T., Kleiber, W., and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105:1167–1177.

Goyal, P., Chan, A. T., and Jaiswal, N. (2006). Statistical models for the prediction of respirable suspended particulate matter in urban cities. *Atmospheric Environment*, 40(11):2068–2077.

Greene, W. H. (2014). *Econometric Analysis*. Pearson Education India.

Hewer, R., Friederichs, P., Hense, A., and Schlather, M. (2017). A Matérn-based multivariate gaussian random process for a consistent model of the horizontal wind components and related variables. *Journal of the Atmospheric Sciences*, 74(11):3833–3845.

Huang, H., Sun, Y., and Genton, M. G. (2020). Visualization of covariance structures for multivariate spatio-temporal random fields. *arXiv preprint arXiv:2008.03689*.

Inoue, T., Sasaki, T., and Washio, T. (2012). Spatio-temporal kriging of solar radiation incorporating direction and speed of cloud movement. In *The 26th Annual Conference of the Japanese Society for Artificial Intelligence*.

Ip, R. H. and Li, W. K. (2015). Time varying spatio-temporal covariance models. *Spatial Statistics*, 14:269–285.

Kallos, G., Astitha, M., Katsafados, P., and Spyrou, C. (2007). Long-range transport of anthropogenically and naturally produced particulate matter in the Mediterranean and North Atlantic: Current state of knowledge. *Journal of Applied Meteorology and Climatology*, 46(8):1230–1251.

Katra, I., Elperin, T., Fominykh, A., Krasovitov, B., and Yizhaq, H. (2016). Modeling of particulate matter transport in atmospheric boundary layer following dust emission from source areas. *Aeolian Research*, 20:147–156.

Knox, J. B. (1974). Numerical modeling of the transport diffusion and deposition of pollutants for regions and extended scales. *Journal of the Air Pollution Control Association*, 24(7):660–664.

Li, B. and Zhang, H. (2011). An approach to modeling asymmetric multivariate spatial covariance structures. *Journal of Multivariate Analysis*, 102(10):1445–1453.

Lonij, V. P., Brooks, A. E., Cronin, A. D., Leuthold, M., and Koch, K. (2013). Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors. *Solar Energy*, 97:58–66.

Mehdipour, V., Stevenson, D. S., Memarianfard, M., and Sihag, P. (2018). Comparing different methods for statistical modeling of particulate matter in Tehran, Iran. *Air Quality, Atmosphere & Health*, 11(10):1155–1165.

Munir, S., Habeebullah, T. M., Seroji, A. R., Morsy, E. A., Mohammed, A. M., Saud, W. A., Abdou, A. E., Awad, A. H., et al. (2013). Modeling particulate matter concentrations in Makkah, applying a statistical modeling approach. *Aerosol and Air Quality Research*, 13(3):901–910.

Murphy, K. M. and Topel, R. H. (2002). Estimation and inference in two-step econometric models. *Journal of Business & Economic Statistics*, 20(1):88–97.

National Research Council (2010). *Global sources of local pollution: an assessment of long-range transport of key air pollutants to and from the United States*. National Academies Press.

Paciorek, C. J. and Schervish, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, 17(5):483–506.

Paciorek, C. J., Yanosky, J. D., Puett, R. C., Laden, F., Suh, H. H., et al. (2009). Practical large-scale spatio-temporal modeling of particulate matter concentrations. *The Annals of Applied Statistics*, 3(1):370–397.

Park, M. S. and Fuentes, M. (2006). New classes of asymmetric spatial-temporal covariance models. Technical report, North Carolina State University. Dept. of Statistics.

Porcu, E., Gregori, P., and Mateu, J. (2006). Nonseparable stationary anisotropic space–time covariance functions. *Stochastic Environmental Research and Risk Assessment*, 21(2):113–122.

Qadir, G. A., Euán, C., and Sun, Y. (2021). Flexible modeling of variable asymmetries in cross-covariance functions for multivariate random fields. *Journal of Agricultural, Biological and Environmental Statistics*, 26(1):1–22.

Randles, C., Da Silva, A., Buchard, V., Colarco, P., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., et al. (2017). The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. *Journal of Climate*, 30(17):6823–6850.

Sahu, S. K. (2012). Hierarchical Bayesian models for space–time air pollution data. In *Handbook of Statistics*, volume 30, pages 477–495. Elsevier.

Sahu, S. K., Gelfand, A. E., and Holland, D. M. (2006). Spatio-temporal modeling of fine particulate matter. *Journal of Agricultural, Biological, and Environmental Statistics*, 11(1):61–86.

Salvaña, M. L. and Genton, M. G. (2020). Nonstationary cross-covariance functions for multivariate spatio-temporal random fields. *Spatial Statistics*, 36:100411.

Salvaña, M. L. and Genton, M. G. (2021). Lagrangian spatio-temporal nonstationary covariance functions. *Book Chapter in Festschrift for Prof. C. Thomas-Agnan*. To appear.

Salvaña, M. L., Abdulah, S., Huang, H., Ltaief, H., Sun, Y., Genton, M. G., and Keyes, D. E. (2020). High performance multivariate spatial modeling for geostatistical data on manycore systems. *arXiv preprint arXiv:2008.07437*.

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

Shao, Y., Wyrwoll, K.-H., Chappell, A., Huang, J., Lin, Z., McTainsh, G. H., Mikami, M., Tanaka, T. Y., Wang, X., and Yoon, S. (2011). Dust cycle: An emerging core theme in Earth system science. *Aeolian Research*, 2(4):181–204.

Shinozaki, K., Yamakawa, N., Sasaki, T., and Inoue, T. (2016). Areal solar irradiance estimated by sparsely distributed observations of solar radiation. *IEEE Transactions on Power Systems*, 31(1):35–42.

Shor, T., Kalka, I., Geiger, D., Erlich, Y., and Weissbrod, O. (2019). Estimating variance components in population scale family trees. *PLoS Genetics*, 15(5):e1008124.

Stein, M. L. (2005a). Space–time covariance functions. *Journal of the American Statistical Association*, 100(469):310–321.

Stein, M. L. (2005b). Statistical methods for regular monitoring data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(5):667–687.

Ukhov, A., Mostamandi, S., da Silva, A., Flemming, J., Alshehri, Y., Shevchenko, I., and Stenchikov, G. (2020). Assessment of natural and anthropogenic aerosol air pollution in the Middle East using MERRA-2, CAMS data assimilation products, and high-resolution WRF-Chem model simulations. *Atmospheric Chemistry and Physics*, 20(15):9281–9310.

Van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., Lyapustin, A., Sayer, A. M., and Winker, D. M. (2016). Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. *Environmental Science & Technology*, 50(7):3762–3772.

Wackernagel, H. (2003). *Multivariate Geostatistics: An Introduction with Applications*. Springer, Berlin, 3rd edition.

Zhang, H. and Cai, W. (2015). When doesn't cokriging outperform kriging? *Statistical Science*, 30(2):176–180.

Zhang, Z. and Chen, Q. (2007). Comparison of the Eulerian and Lagrangian methods for predicting particle transport in enclosed spaces. *Atmospheric Environment*, 41(25):5236–5248.

Zhelonkin, M., Genton, M. G., and Ronchetti, E. (2012). On the robustness of two-stage estimators. *Statistics & Probability Letters*, 82(4):726–732.