Spatial Statistics 37 (2020) 100411

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Nonstationary cross-covariance functions for multivariate spatio-temporal random fields ☆

Mary Lai O. Salvaña *, Marc G. Genton

Statistics Program, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

ARTICLE INFO

Article history:
Received 13 October 2019
Received in revised form 11 January 2020
Accepted 16 January 2020
Available online 25 January 2020

Keywords:
Cross-covariance function
Lagrangian framework
Multivariate random field
Nonstationarity
Space–time
Transport effect

ABSTRACT

In multivariate spatio-temporal analysis, we are faced with the formidable challenge of specifying a valid spatio-temporal cross-covariance function, either directly or through the construction of processes. This task is difficult as these functions should yield positive definite covariance matrices. In recent years, we have seen a flourishing of methods and theories on constructing spatio-temporal cross-covariance functions satisfying the positive definiteness requirement. A subset of those techniques produced spatio-temporal cross-covariance functions possessing the additional feature of nonstationarity. Here we provide a review of the state-of-the-art methods and technical progress regarding model construction. In addition, we introduce a rich class of multivariate spatio-temporal asymmetric nonstationary models stemming from the Lagrangian framework. We demonstrate the capabilities of the proposed models on a bivariate reanalysis climate model output dataset previously analyzed using purely spatial models. Furthermore, we carry out a cross-validation study to examine the advantages of using spatio-temporal models over purely spatial models. Finally, we outline future research directions and open problems.

1. Introduction

The importance of multivariate spatio-temporal geostatistical models is far-reaching and interdisciplinary in nature. Theoretical developments in the field of geostatistics have provided

☆ This research was supported by the King Abdullah University of Science and Technology (KAUST), Saudi Arabia.

* Corresponding author.
E-mail addresses: marylai.salvana@kaust.edu.sa (M.L.O. Salvaña), marc.genton@kaust.edu.sa (M.G. Genton).

https://doi.org/10.1016/j.spasta.2020.100411
2211-6753/

frameworks and tools for several important applications. Traditionally, geostatistical models were used to produce high-quality mappings of heavy metal contamination, ore deposits, pollutants, sand distribution in shelf seas, temperature, vegetation, to name a few. They were, and still are, needed to accurately model predictor variables, such as precipitation, the importance of which cannot be overlooked as these predictors serve as key input conditions for running physical models, such as those in hydrology and geomorphology. Constructions of full-coverage high-resolution maps of forest composition are made possible by augmenting the collected remotely sensed light detection and ranging (LiDAR) data with geostatistical modeling techniques (Taylor-Rodriguez et al., 2019). Today, as the field experiences phenomenal research progress, more advanced models are becoming the norm. Specifically, the geostatistics community witnessed research advances in the field of multivariate spatio-temporal modeling. The rationale for turning significant research efforts in this direction is rooted in the fact that data are almost always multivariate, spatial, and temporal.

Consider a multivariate spatio-temporal random field $\mathbf{Z}(\mathbf{s},t) = \left\{Z_1(\mathbf{s},t),\ldots,Z_p(\mathbf{s},t)\right\}^\top$, such that there are p variables at each spatio-temporal location $(\mathbf{s},t) \in \mathbb{R}^d \times \mathbb{R}$, $d \geq 1$. Operating under the Gaussian multivariate random field assumption, the stochastic behavior of $\mathbf{Z}(\mathbf{s},t)$ is fully described by its mean vector $\boldsymbol{\mu}(\mathbf{s},t) = \mathrm{E}\left\{\mathbf{Z}(\mathbf{s},t)\right\}$ and (nonstationary) cross-covariance structure $\mathbf{C}(\mathbf{s}_1,\mathbf{s}_2;t_1,t_2) = \left\{C_{ij}(\mathbf{s}_1,\mathbf{s}_2;t_1,t_2)\right\}_{i,j=1}^p$, where $C_{ij}(\mathbf{s}_1,\mathbf{s}_2;t_1,t_2) = \mathrm{cov}\left\{Z_i(\mathbf{s}_1,t_1),Z_j(\mathbf{s}_2,t_2)\right\}$, $\mathbf{s}_1,\mathbf{s}_2 \in \mathbb{R}^d$, $t_1,t_2 \in \mathbb{R}$, for $i,j = 1,\ldots,p$. Because of this property, multivariate spatio-temporal geostatistics is heavily focused on specifying appropriate spatio-temporal cross-covariance functions, C_{ij}. The resulting spatio-temporal covariance matrix has to be positive definite, i.e., for any $n \in \mathbb{Z}^+$, for any finite set of points $(\mathbf{s}_1,t_1),\ldots,(\mathbf{s}_n,t_n)$, and for any vector $\boldsymbol{\lambda} \in \mathbb{R}^{np}$, we have $\boldsymbol{\lambda}^\top \boldsymbol{\Sigma} \boldsymbol{\lambda} \geq 0$, where $\boldsymbol{\Sigma}$ is an $np \times np$ matrix with $n \times n$ block elements of $p \times p$ matrices $\mathbf{C}(\mathbf{s}_l,\mathbf{s}_r;t_l,t_r)$, $l,r = 1,\ldots,n$, with n indicating the number of spatio-temporal locations. A valid cross-covariance function ensures the $(np \times np)$-dimensional covariance matrix of the np-dimensional vector $\left\{\mathbf{Z}(\mathbf{s}_1,t_1)^\top,\ldots,\mathbf{Z}(\mathbf{s}_n,t_n)^\top\right\}^\top$ to be positive definite.

Without loss of generality, assume $\boldsymbol{\mu}(\mathbf{s},t) = \mathbf{0}$, for all $\mathbf{s}$ and t. A multivariate spatio-temporal random field is (weakly) stationary if its spatio-temporal cross-covariance function $C_{ij}(\mathbf{s}_1,\mathbf{s}_2;t_1,t_2)$ simplifies to $C_{ij}(\mathbf{h},u)$, where $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$ and $u = t_1 - t_2$, for $i,j = 1,\ldots,p$. This means that the cross-covariance between variables i and j depends only on the vector of their separation in space and time, and that this value does not change regardless of their locations in space and time. If the cross-covariance depends only on the magnitude of their separation, then $C_{ij}(\mathbf{s}_1,\mathbf{s}_2;t_1,t_2)$ further simplifies to $C_{ij}(\|\mathbf{h}\|,|u|)$, where $\|\mathbf{h}\| = \|\mathbf{s}_1 - \mathbf{s}_2\|$ and $|u| = |t_1 - t_2|$, for $i,j = 1,\ldots,p$. A multivariate spatio-temporal random field with such spatio-temporal cross-covariance structure is termed isotropic.

Multivariate geostatistics offers the tool of co-kriging which one can use in interpolating and predicting a variable, possibly undersampled, given its spatio-temporal relationship to other oversampled variables. In co-kriging, formally, the goal is to predict $\mathbf{Z}(\mathbf{s}_0,t_0)$ at an unobserved spatio-temporal location $\mathbf{s}_0 \in \mathbb{R}^d$ and $t_0 \in \mathbb{R}$, given $\mathbf{Z} = \left\{\mathbf{Z}(\mathbf{s}_1,t_1)^\top,\ldots,\mathbf{Z}(\mathbf{s}_n,t_n)^\top\right\}^\top$. Under the squared-error loss criterion, the simple co-kriging predictor of $\mathbf{Z}(\mathbf{s}_0,t_0)$ is the best linear unbiased predictor $\widehat{\mathbf{Z}}(\mathbf{s}_0,t_0) = \mathrm{E}\left\{\mathbf{Z}(\mathbf{s}_0,t_0)|\mathbf{Z}(\mathbf{s}_1,t_1),\ldots,\mathbf{Z}(\mathbf{s}_n,t_n)\right\}$ with the form $\widehat{\mathbf{Z}}(\mathbf{s}_0,t_0) = \boldsymbol{\Delta}_0^\top \boldsymbol{\Sigma}^{-1}\mathbf{Z}$, where $\boldsymbol{\Delta}_0 = \left\{\mathbf{C}(\mathbf{s}_0,\mathbf{s}_1;t_0,t_1),\mathbf{C}(\mathbf{s}_0,\mathbf{s}_2;t_0,t_2),\ldots,\mathbf{C}(\mathbf{s}_0,\mathbf{s}_n;t_0,t_n)\right\}^\top$.

The rest of the paper is organized as follows. An overview of the state-of-the-art advances in the field of multivariate spatial and spatio-temporal geostatistics is presented in Sections 2 and 3, respectively, focusing mainly on results that appeared after the review of Genton and Kleiber (2015). A specialized class of multivariate spatio-temporal models, based on transport, is described in Sections 4 and 5. In Section 6, the paper transitions from surveying recent works to the creation of new multivariate spatio-temporal nonstationary models. Section 7 illustrates the performance of the new models on a regional climate model output bivariate dataset. Section 8 outlines the theoretical challenges and practical considerations when the spatial locations are defined on a sphere and newly established models on the sphere are provided. Section 9 concludes with a discussion on new research avenues and remaining challenges.

2. Purely spatial models

Having established the setup of the multivariate spatio-temporal random field, we begin to highlight recent studies, starting with multivariate purely spatial models, and then proceeding with the spatio-temporal ones. For the purely spatial discussions, we use the same notations as above and simply suppress the temporal argument in the observations and set $u = 0$ for the temporal argument in the cross-covariance function.

2.1. Spatial stationary cross-covariance functions

Genton and Kleiber (2015) reviewed three general methods of constructing stationary cross-covariance functions from existing univariate stationary covariance functions: the linear model of coregionalization (LMC), convolution methods, and latent dimensions. The LMC considers the multivariate process as a linear combination of uncorrelated univariate spatial processes. One major shortcoming of this model is that it lacks flexibility as it bestows on all variables the smoothness of the roughest underlying univariate spatial process. The convolution methods, on the other hand, require convolving spatially-varying kernel functions. The resulting cross-covariance function may or may not have a closed form. Lastly, the latent dimensions approach works by representing the components of $\mathbf{Z}(\mathbf{s})$ as coordinates in a k-dimensional space, $1 \leq k \leq p$.

Another proposed model is the Matérn stationary cross-covariance function formulated by Gneiting et al. (2010). Additional work on the allowable parameter values for the stationary Matérn cross-covariance function was carried out by Apanasovich et al. (2012). Cressie et al. (2015) asserted the use of a multivariate spatial random effects model. Marcotte (2015) developed a non-linear model of coregionalization (N-LMC), addressing the aforementioned critical drawback of the LMC. The N-LMC allows for different sets of uncorrelated univariate spatial processes in the marginals and the cross-covariances. Cressie and Zammit-Mangion (2016) introduced the conditional approach to model multivariate spatial dependence, with variable asymmetry as an additional feature. They modeled the cross-covariance structure using univariate conditional covariance functions based on the partitions of $\mathbf{Z}(\mathbf{s})$. Ideally, the partitioning of $\mathbf{Z}(\mathbf{s})$ should reflect the causal relationship between the variables, but this may not be easy to define with many variables. Gnann et al. (2018) proposed a bivariate correlation model that resembles the LMC, with the form $C_{12}(\mathbf{h}) = \rho\sqrt{C_{11}(\mathbf{h})C_{22}(\mathbf{h})}$, where $-1 \leq \rho \leq 1$ is the usual colocated correlation coefficient. Although they did not provide a proof of its validity, the covariance matrices they obtained were positive definite. Marcotte (2019) issued some caution regarding the use of that model and provided four counterexamples, one for each exponential, squared exponential, Matérn, and spherical correlations, where the ordinary co-kriging variance turned out to be negative. Finally, not unrelated, is the work of Bevilacqua et al. (2015), where two criteria that compare the flexibility of two different cross-covariance functions were defined.

New univariate models, such as the modified Matérn of Laga and Kleiber (2017), require mention. The modified Matérn has a spectral density

$$f(\|\boldsymbol{\omega}\|) = \frac{(b^2 + \|\boldsymbol{\omega}\|)^{\xi}}{(a^2 + \|\boldsymbol{\omega}\|)^{\nu+d/2}}, \quad \boldsymbol{\omega} \in \mathbb{R}^d,$$

where $a, \nu > 0$, $b \geq 0$, and $\xi < \nu$. The last condition is in place to make sure that the process has finite variance. When $\xi = 0$, one obtains the classical Matérn spectral density. The model presented above is more flexible than the classical one as the maximum spectrum can occur at a non-zero frequency. When $d = 2$, they derived its resulting covariance as follows:

$$C(\mathbf{h}) = \frac{1}{2\pi}\frac{(-1)^{\nu}}{\nu!}\frac{\partial^{\nu}}{\partial(a^2)^{\nu}}\left\{(b^2 - a^2)^{\xi}\mathcal{K}_0\left(a\|\mathbf{h}\|\right)\right\}, \quad \mathbf{h} \in \mathbb{R}^d,$$

where $\mathcal{K}_0$ is a modified Bessel function of the second kind of order zero. The resulting random field of this model can exhibit strong periodicities, which the random fields from the Matérn model of Gneiting et al. (2010) do not possess.

2.2. Spatial nonstationary cross-covariance functions

The next important advancement is the development of cross-covariance functions that lead to multivariate purely spatial second-order nonstationary behavior. These models made it possible to allow the purely spatial cross-covariance structure to depend on the spatial locations. The models in the previous section are only appropriate when applied to multivariate purely spatial stationary random fields. When stationarity in the second-order structure is untenable, one turns to multivariate purely spatial nonstationary models. Models that accommodate nonstationarity in the cross-covariance structure are more pertinent when studying large spatial domains, as the spatial second-order nonstationary behavior can be attributed to differing spatial features affecting the phenomena being investigated. The assumption of second-order stationarity is often reasonable when one studies a relatively small spatial domain or when one can substantiate that the spatial features in the whole domain are spatially invariable.

We now mention several studies that advanced the models available for studying multivariate purely spatial nonstationary behavior. Genton and Kleiber (2015) started the discussion on these models by highlighting the different nonstationary extensions of the LMC such as that of Gelfand et al. (2004) and Fouedjio (2018), and the nonstationary multivariate Matérn model of Kleiber and Nychka (2012). These studies are based on similar approaches, i.e., they use multivariate purely spatial stationary models as the building blocks of complicated nonstationary models.

Another way of building multivariate purely spatial nonstationary models is to start with valid univariate purely spatial nonstationary models and extend them to the multivariate case. Hence, it is not surprising that many advances in univariate purely spatial nonstationary models have been achieved. A survey of new approaches to building univariate nonstationary covariance functions was presented by Fouedjio (2017). In addition to the aforementioned survey are other papers on covariate-driven purely spatial nonstationary models (Risser, 2015; Risser and Calder, 2015), nonstationary convolution models (Fouedjio et al., 2016), space deformation approach (Sampson and Guttorp, 1992; Fouedjio et al., 2015; Kleiber, 2016), convolution models incorporating information regarding suspected potential sources, for instance pipes or reservoirs, with application on dosimetric data (Lajaunie et al., 2019), and a nonstationary Matérn model for data affected by boundaries, holes, or physical barriers (Bakka et al., 2019). Ton et al. (2018) merged principles from Fourier feature representations, Gaussian processes, and neural networks to create new nonstationary covariance functions. Multivariate extensions to these models are nontrivial and have yet to be proposed.

Dimension expansions and covariance functions built as a linear combination of several local basis functions are two other widely popular univariate nonstationary models with multivariate nonstationary versions yet to be seen in the literature. Briefly, we propose straightforward extensions of these two approaches in the multivariate arena. Consider the multivariate Karhunen–Loève expansion (Theorem 5.2.2 of Wang, 2008) of the multivariate purely spatial random field $\mathbf{Z}(\mathbf{s}) = \sum_{b=1}^{\infty} \left\{\xi_{b,1}\lambda_{b,1}\phi_{b,1}(\mathbf{s}), \ldots, \xi_{b,p}\lambda_{b,p}\phi_{b,p}(\mathbf{s})\right\}^{\top}$, such that $\boldsymbol{\xi} = (\xi_{b,1}, \ldots, \xi_{b,p})^{\top} \sim \mathcal{N}_p(\mathbf{0}, \mathbf{I}_{p\times p})$, $\lambda_{b,i} \in \mathbb{R}$, and $\phi_{b,i}(\mathbf{s})$, $i = 1, \ldots, p$, are the local basis functions. The resulting nonstationary cross-covariance function of this random field is

$$C_{ij}(\mathbf{s}_1, \mathbf{s}_2) = \sum_{b=1}^{\infty} \lambda_{b,i}\lambda_{b,j}\phi_{b,i}(\mathbf{s}_1)\phi_{b,j}(\mathbf{s}_2), \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d.$$

In a landmark research in nonstationary spatial modeling, Bornn et al. (2012) showed how reducing the spatial dimensions can cause a nonstationary behavior in the covariance structure. Following their construction approach, this time with $p > 1$, consider a multivariate purely spatial nonstationary random field $\mathbf{Z}(\mathbf{s}) = \left\{Z_1(\mathbf{s}), \ldots, Z_p(\mathbf{s})\right\}^{\top}$ such that $\mathbf{Z}(\mathbf{s}, \tilde{\boldsymbol{\eta}}) = \left\{Z_1(\mathbf{s}, \boldsymbol{\eta}_1), \ldots, Z_p(\mathbf{s}, \boldsymbol{\eta}_p)\right\}^{\top}$, where $\mathbf{s} \in \mathbb{R}^d$, $\tilde{\boldsymbol{\eta}} = \left\{\boldsymbol{\eta}_1^{\top}, \ldots, \boldsymbol{\eta}_p^{\top}\right\}^{\top} \in \mathbb{R}^{pd'}$, and $\boldsymbol{\eta}_i \in \mathbb{R}^{d'}$, $d' > 0$, is a multivariate purely spatial stationary random field, for $i = 1, \ldots, p$. The components i, j of $\mathbf{Z}(\mathbf{s}, \tilde{\boldsymbol{\eta}})$, taken at spatial locations $(\mathbf{s}_1, \boldsymbol{\eta}_{1i})$ and $(\mathbf{s}_2, \boldsymbol{\eta}_{2j})$, after accounting for latent spatial dimensions, have a stationary cross-covariance $C_{ij}\left\{(\mathbf{s}_1, \boldsymbol{\eta}_{1i}) - (\mathbf{s}_2, \boldsymbol{\eta}_{2j})\right\}$. The relationship between the stationary cross-covariance of the purely spatial stationary random field $\mathbf{Z}(\mathbf{s}, \tilde{\boldsymbol{\eta}})$, and that of the purely spatial nonstationary random field $\mathbf{Z}(\mathbf{s})$ cannot be explicitly characterized, except when the cross-covariance function, C_{ij}, is taken as the squared exponential stationary covariance function. In that case, $C_{ij}(\mathbf{s}_1 - \mathbf{s}_2) = C_{ij}\left\{(\mathbf{s}_1, \boldsymbol{\eta}_{1i}) - (\mathbf{s}_2, \boldsymbol{\eta}_{2j})\right\}/C_{ij}(\boldsymbol{\eta}_{1i} - \boldsymbol{\eta}_{2j})$. This dimension expansion approach allows one to use existing computationally more tractable multivariate stationary covariance functions on $\mathbb{R}^{d+d'}$ for data exhibiting second-order nonstationary behavior on $\mathbb{R}^d$.

Fuglstad et al. (2015) issued a caveat on using purely spatial nonstationary models hastily. They emphasized that a random field with a seemingly nonstationary second-order structure may possess that behavior due to an unaccounted nonstationarity in the mean structure. Spurious nonstationarity results in a misleading covariance structure. It is in the modeler's interest to correctly identify the source of the nonstationarity because while nonstationarity in the mean structure is cheap, nonstationarity in the covariance structure is not.

3. Spatio-temporal models

Combining spatial models with temporal information can tremendously improve modeling capabilities. The utility of spatio-temporal cross-covariance functions is predicated on the idea that closer objects tend to behave similarly than those that are distant to each other (Tobler, 1970). Recently, this principle was recognized to hold when distance is taken with respect to the objects' locations in time. This sparked enormous interest in building spatio-temporal models. Further, these models were constructed around the need to characterize the behavior and interaction of multiple variables as they evolve in space and time. Nevertheless, it is essential to clarify that the focus of spatio-temporal geostatistics is typically not on the "how" of evolution; it is on describing the spatio-temporal mechanisms of an underlying process that may have generated the data.

Often, spatio-temporal datasets are modeled in the purely spatial context. When there are missing data, as temporal information is sometimes limited, one usually resorts to collapsing a spatio-temporal dataset to a spatial one by taking the spatial location-wise arithmetic mean. This is a perfectly legitimate approach as long as the scientific question to be answered is purely spatial in nature. But when the question is spatio-temporal, purely spatial models are insufficient.

3.1. Spatio-temporal stationary cross-covariance functions

The common genesis of many established spatio-temporal stationary cross-covariance functions is either a purely spatial stationary cross-covariance function or a univariate spatio-temporal stationary covariance function. A hybrid of these two approaches, the spatio-temporal (space–time) separable stationary cross-covariance function is arguably the easiest way to build multivariate spatio-temporal stationary models. Given a purely spatial stationary cross-covariance, $C_{ij}(\mathbf{h})$, and a univariate purely temporal stationary covariance, $C_T(u)$, then their product $C_{ij}(\mathbf{h}, u) = C_{ij}(\mathbf{h})C_T(u)$, $\mathbf{h} \in \mathbb{R}^d, u \in \mathbb{R}$, is a valid spatio-temporal (space–time) separable stationary cross-covariance function. However, multivariate spatio-temporal (space–time) separable models are always space–time fully symmetric. Note that any spatio-temporal stationary cross-covariance function is space–time symmetric, i.e., $C_{ij}(\mathbf{h}, u) = C_{ij}(-\mathbf{h}, -u)$ for any spatio-temporal lag combinations $(\mathbf{h}, u)$, $i, j = 1, \ldots, p$. A stricter form of space–time symmetry, termed full symmetry, occurs when $C_{ij}(\mathbf{h}, u) = C_{ij}(-\mathbf{h}, u) = C_{ij}(\mathbf{h}, -u) = C_{ij}(-\mathbf{h}, -u)$. Full symmetry means that the cross-covariance between variable i at site $\mathbf{s}_1$, at time t_1, and variable j at site $\mathbf{s}_2$, at time t_2, where $\mathbf{s}_2 = \mathbf{s}_1 + \mathbf{h}$ and $t_2 = t_1 + u$, is identical to that of variable i at site $\mathbf{s}_1$, at time t_2, and variable j at site $\mathbf{s}_2$, at time t_1. This assumption of full symmetry is known to not hold in reality, especially when modeling environmental and earth sciences data that are influenced by natural occurring forces, for example, atmospheric flows. These types of phenomena are so prevalent that their appropriate models constitute a special subclass and are discussed in a separate succeeding section.

The different spatio-temporal extensions of the stationary LMC offer spatio-temporal separable stationary cross-covariance models with different types of separability. When the univariate covariances in the LMC are written as a product of two univariate covariance functions, one purely spatial and the other purely temporal, the LMC model is fully separable. Otherwise, when the univariate covariances are space–time nonseparable, the LMC model is a variable-separable model. A collection of different spatio-temporal stationary LMC is provided in De Iaco et al. (2019), and a list of different types of separability (full, space, time, and variable, to name a few) is discussed in Apanasovich and Genton (2010).

Other works on the several adaptations of the multivariate purely spatial Matérn to space–time were recently contributed in the literature such as Bourotte et al. (2016) and Ip and Li (2016, 2017). In a Bayesian formalism, Zammit-Mangion et al. (2015) proposed a multivariate spatio-temporal model by merging stochastic partial differential equations with the spatio-temporal random field theory. Rodrigues and Diggle (2010) proposed a spatio-temporal extension of the stationary convolution models.

Book-length discussions recently appeared in the literature on the topic of spatio-temporal models such as Montero et al. (2015), Christakos (2017), Wikle et al. (2019), and Corzo and Varouchakis (2019). Christakos (2017) offered an extensive coverage of spatio-temporal geostatistics from theory to applications in the univariate and multivariate settings. Corzo and Varouchakis (2019) provided a review of the recently available spatio-temporal univariate models.

3.2. Spatio-temporal nonstationary cross-covariance functions

Again, assuming second-order stationarity in space and in time is a convenient starting point, however, models that accommodate more realistic assumptions such as second-order nonstationarity in space and/or time are needed. These models offer more sophistication than those in the previous sections. Hence, a very sparse literature on spatio-temporal nonstationary cross-covariance functions is expected.

Ip and Li (2015) examined the possibility of changing the spatio-temporal covariance structure depending on the temporal location. The idea is simple and it addresses the problem of second-order nonstationarity in space and/or time. In their work, Ip and Li (2015) outlined several theorems that allow the $Np \times Np$ matrix, $\mathbf{C}(\cdot; t_1, t_2)$, to assume different parametric forms for any t_1 and t_2, where N is the number of spatial locations.

A number of newly developed spatio-temporal nonstationary models were proposed only in the univariate setting such as the improved latent space approach (ILSA) of Xu and Gardoni (2018). A spatio-temporal Karhunen–Loève expansion developed in Choi (2014) may be used to construct covariance functions that are nonstationary in space and/or time. Following the work of Bornn et al. (2012), Shand and Li (2017) formulated a spatio-temporal dimension expansion approach by performing a straightforward expansion of the temporal dimension. Again, a multivariate version of the approach has not yet been proposed and can be done as follows: Consider a multivariate spatio-temporal nonstationary random field $\mathbf{Z}(\mathbf{s}, t) = \left\{Z_1(\mathbf{s}, t), \ldots, Z_p(\mathbf{s}, t)\right\}^\top$ such that $\mathbf{Z}(\mathbf{s}, \tilde{\boldsymbol{\eta}}; t, \tilde{\boldsymbol{\xi}}) = \left\{Z_1(\mathbf{s}, \boldsymbol{\eta}_1; t, \boldsymbol{\xi}_1), \ldots, Z_p(\mathbf{s}, \boldsymbol{\eta}_p; t, \boldsymbol{\xi}_p)\right\}^\top$, where $\mathbf{s} \in \mathbb{R}^d$, $\tilde{\boldsymbol{\eta}} = \left\{\boldsymbol{\eta}_1^\top, \ldots, \boldsymbol{\eta}_p^\top\right\}^\top \in \mathbb{R}^{pd'}$, $\boldsymbol{\eta}_i \in \mathbb{R}^{d'}$, $\tilde{\boldsymbol{\xi}} = \left\{\boldsymbol{\xi}_1^\top, \ldots, \boldsymbol{\xi}_p^\top\right\}^\top \in \mathbb{R}^{pd''}$, and $\boldsymbol{\xi}_i \in \mathbb{R}^{d''}$, $d' + d'' > 0$, is a multivariate spatio-temporal stationary random field, for $i = 1, \ldots, p$. The components i, j of $\mathbf{Z}(\mathbf{s}, \tilde{\boldsymbol{\eta}}; t, \tilde{\boldsymbol{\xi}})$, taken at spatio-temporal locations $(\mathbf{s}_1, \boldsymbol{\eta}_{1i}; t_1, \boldsymbol{\xi}_{1i})$ and $(\mathbf{s}_2, \boldsymbol{\eta}_{2j}; t_2, \boldsymbol{\xi}_{2j})$, after accounting for the extra spatial and temporal dimensions, have a spatio-temporal stationary cross-covariance $C_{ij}\left\{(\mathbf{s}_1, \boldsymbol{\eta}_{1i}) - (\mathbf{s}_2, \boldsymbol{\eta}_{2j}); (t_1, \boldsymbol{\xi}_{1i}) - (t_2, \boldsymbol{\xi}_{2j})\right\}$.

4. Physically-motivated space–time models

A recurring theme in the elaborate discussions in Christakos (2017) is the need for physically-motivated models. This is especially true for environmental datasets for which variables should not be analyzed in a vacuum. Certain physical features that are inherent to the physical nature of these variables are appropriately required to be incorporated in the cross-covariance function. Otherwise, modeling efforts are deemed useless. Multivariate spatio-temporal geostatistics offers alternatives to computationally expensive physics models by developing complex covariance function models that can satisfactorily represent complicated physical phenomena. Waymire et al. (1984) advocated the use of statistical models to describe ground-level observations and derived a physically realistic stochastic behavior based on empirical observations. Examples of physically realistic statistical models are the spatio-temporal covariance models satisfying the Taylor's hypothesis (Taylor, 1938).

A univariate spatio-temporal stationary covariance function $C(\mathbf{h}, u)$ on $\mathbb{R}^d \times \mathbb{R}$ satisfies Taylor's hypothesis if there exists $\mathbf{v} \in \mathbb{R}^d$ such that $C(\mathbf{0}, u) = C(\mathbf{v}u, 0)$, $u \in \mathbb{R}$. The Taylor's hypothesis tells us that the marginal spatial and temporal covariances can proxy each other and this is good because one form might be more accessible than the other. Statistical theories of turbulence have sprung from this simple relationship and several laboratory experiments have confirmed this equivalence, at least at certain scales; see, e.g., Li et al. (2009) and references therein.

Bras and Rodríguez-Iturbe (1976) and Lovejoy and Mandelbrot (1985) studied the modeling implications of the Taylor's hypothesis by first modeling rainfall as a purely spatial phenomenon and then considering it as a spatio-temporal phenomenon using Taylor's hypothesis. Gupta and Waymire (1987) were one of the first to rigorously develop the implications of Taylor's hypothesis in spatio-temporal statistics. They defined a process

$$Z(\mathbf{s}, t) = \tilde{Z}(\mathbf{s} - \mathbf{v}t) \tag{1}$$

with spatio-temporal stationary covariance function $C(\mathbf{h}, u) = C^S(\mathbf{h} - \mathbf{v}u)$ and called it the frozen field. Here, $\tilde{Z}$ is a purely spatial stationary random field and C^S is its purely spatial stationary covariance function. Cox and Isham (1988) offered more flexibility to the frozen field model by replacing the constant velocity with a random velocity, $\mathbf{V} \in \mathbb{R}^d$, resulting in a spatio-temporal covariance function model of the form $C(\mathbf{h}, u) = \mathrm{E}_{\mathbf{V}}\left\{C^S(\mathbf{h} - \mathbf{V}u)\right\}$. We call this model the non-frozen random field model. The constant $\mathbf{v}$ in the frozen field model is usually treated as the mean of the random variable $\mathbf{V}$. The vectors $\mathbf{V}$ and $\mathbf{v}$ are commonly referred to as the random and constant transport or advection velocity vectors, respectively. A random field simulated from the frozen field model can exhibit a dimple effect. A dimple effect points to a phenomenon where the covariance between $Z(\mathbf{s}_1, t_1)$ and $Z(\mathbf{s}_2, t_2)$ is stronger than that of $Z(\mathbf{s}_1, t_1)$ and $Z(\mathbf{s}_2, t_1)$, where $t_2 = t_1 + 1$; see Kent et al. (2011). The frozen and non-frozen models, as well as the dimple effect, are physically justifiable by observations influenced by transport phenomena that are usually caused by predominant winds, waves, and flows, to name but a few. The stochastic representation (1) makes physical sense in modeling observations influenced by transport phenomena under the Lagrangian reference frame. This reference frame has its roots in physics and is a way of describing the development of a phenomenon in space and in time while moving or traveling with it.

Covariance functions centered around modeling a process (1) are collectively termed "spatio-temporal covariance functions under the Lagrangian framework" (Gneiting, 2002; Gneiting et al., 2007). These covariance functions use the Lagrangian reference frame to build spatio-temporal covariance functions from purely spatial covariance functions. A survey of existing literature suggests that there is no detailed Lagrangian formulation in the multivariate nonstationary arena. Hence, we took significant strides towards developing and unifying the modeling of multivariate spatio-temporal transport datasets using specialized covariance functions under the Lagrangian framework.

5. The Lagrangian framework

The Lagrangian framework transforms what had been primarily a purely spatial covariance function into a spatio-temporal covariance function. Under this framework, the transport property is exploited, and one readily obtains substantial performance benefits of a spatio-temporal model using a primarily purely spatial one. The working premise is that the purely spatial random field retains its spatial properties while being transported and the models depend for effectiveness on the advection velocity vector. It effectively establishes that the derived spatio-temporal covariance function inherits all the properties of the underlying purely spatial covariance function.

This technique seems to be most widely used in engineering, and only few developments in theory and applications in geostatistics have been proposed. In this section, we review the recent works done on this special construction approach and examine how more sophisticated spatio-temporal models may be conceived.

5.1. Stationary Lagrangian covariance functions

A brief review of the genesis of spatio-temporal covariance functions under the Lagrangian framework has already been presented in Section 4. Ma (2003) established an umbrella theorem that formalizes the validity of purely spatial covariance functions turning into spatio-temporal covariance functions. The frozen and non-frozen field models are stationary in their inception. From time to time, the frozen field model is used to model waves (Ailliot et al., 2011), wind (Gneiting et al., 2007; Ezzat et al., 2018), solar irradiance (Lonij et al., 2013; Inoue et al., 2012), and cloud cover data (Shinozaki et al., 2016). Gneiting et al. (2007) proposed to model an Irish wind dataset using a convex combination of a classical and a Lagrangian spatio-temporal covariance function. Because of prior knowledge of a prevailing westerly wind pattern, the Lagrangian spatio-temporal covariance function assumed the form $C(\mathbf{h}, u) = (1 - \frac{1}{2v_1}|h_1 - uv_1|)_+$, where $\mathbf{h} = (h_1, h_2)^\top$, $\mathbf{v} = (v_1, 0)^\top$, and $(\cdot)_+ = \max(\cdot, 0)$. Christakos et al. (2017) used a different term for this random field and called it a traveling random field that was then used to model the spread of diseases. In his book, Christakos (2017) provided a more in-depth discussion of the traveling random field.

One nice property of Lagrangian spatio-temporal covariance functions is that, in general, along the main direction of transport, the spatio-temporal relationship is asymmetric. Consequently, other observations exhibiting space–time asymmetry in their covariance structure, although not obviously influenced by the transport effect, may utilize models under this framework.

Although the Lagrangian framework easily extends purely spatial covariance functions to space–time, a model under this framework, the stationary frozen field model, has the disadvantage that the spatio-temporal covariance functions it produces are not anisotropic, for any $u \neq 0$. One should not confuse anisotropy and asymmetry. Anisotropy is a property involving only the spatial arguments of the covariance function, whereas asymmetry involves both the spatial and temporal arguments. Porcu et al. (2006) proposed an anisotropic version of this model by partitioning the spatial lag and the advection velocity vector into smaller components: $C(\mathbf{h}, u) = \mathrm{E}_{\mathbf{V}_1,\mathbf{V}_2}\left[\mathcal{L}\left\{\gamma_1\left(\mathbf{h}_1 - \mathbf{V}_1 u\right), \gamma_2\left(\mathbf{h}_2 - \mathbf{V}_2 u\right)\right\}\right]$, $\mathbf{h}_1, \mathbf{V}_1 \in \mathbb{R}^{d_1}$, $\mathbf{h}_2, \mathbf{V}_2 \in \mathbb{R}^{d_2}$, where γ_1, γ_2 are purely spatial stationary variograms with $\gamma_1(\mathbf{0}) = \gamma_2(\mathbf{0}) = 0$ and $d_1 + d_2 = d$. This is a valid spatio-temporal asymmetric component-wise anisotropic stationary covariance function in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}$. Here $\mathcal{L}$ denotes a bivariate Laplace transform with representation $\mathcal{L}(\theta_1, \theta_2) = \int_{[0,\infty)^2} \exp\left(-r_1\theta_1 - r_2\theta_2\right) \mathrm{d}F(r_1, r_2)$, where F is a bivariate probability measure and $\mathcal{L}(0, 0) = 1$.

Another major drawback of the stationary frozen field model is that the model itself does not permit the dampening of the spatio-temporal covariance, meaning that the maximum covariances at different temporal lags are always equal. However, real data do not exhibit this property. Objects that are transported can be subjected to diffusion, i.e., they get transported to different directions at any time point (Hwang et al., 2018). This drawback was addressed by the stationary non-frozen model of Cox and Isham (1988) that allowed a dampened maximum covariance at nonzero temporal lags. The stationary non-frozen model does not have an explicit form except for some special distributions of the random advection velocity vector $\mathbf{V}$ and purely spatial covariance C^S. Schlather (2010) derived the explicit form when $\mathbf{V} \sim \mathcal{N}_d(\boldsymbol{\mu}_\mathbf{V}, \boldsymbol{\Sigma}_\mathbf{V})$ and C^S is the stationary squared exponential covariance function:

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

It is apparent from the model above that the maximum possible covariance decreases as u increases. Furthermore, this model can introduce anisotropy via $\boldsymbol{\Sigma}_\mathbf{V}$ at nonzero u; see Salvaña et al. (2020) for illustrations of the stationary non-frozen field models. Fitting a stationary frozen model to data generated from a stationary non-frozen Lagrangian covariance function can lead to poor parameter estimates and kriging errors, as shown by Salvaña et al. (2020).

5.2. Nonstationary Lagrangian covariance functions

Nonstationarity can also be incorporated into any Lagrangian covariance function. Following Ma (2003), Salvaña and Genton (2020) proposed the inception of new spatio-temporal covariance functions from purely spatial nonstationary ones. Indeed, if $C^S(\mathbf{s}_1, \mathbf{s}_2)$ is a valid purely spatial nonstationary covariance function on $\mathbb{R}^d$, then, $C(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = \mathrm{E}_{\mathbf{V}}\left\{C^S(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2)\right\}$ for $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$ and $t_1, t_2 \in \mathbb{R}$, is a valid spatio-temporal nonstationary covariance function on $\mathbb{R}^d \times \mathbb{R}$ provided that the expectation exists. This model leads to spatio-temporal covariance functions that are nonstationary in both space and time.

The model above implies that the second-order nonstationary spatial profile of the random field is constant at any time point, as the underlying purely spatial nonstationary covariance function, C^S, is independent of any temporal arguments. Notwithstanding, this model can be made more general for some classes of purely spatial nonstationary covariance functions, e.g., deformation models (Sampson and Guttorp, 1992), normal scale-mixture models (Paciorek and Schervish, 2006), and basis function models, such that the second-order nonstationary spatial profile of the random field may be time varying and that one can have second-order nonstationarity in space and/or time.

5.3. Stationary Lagrangian cross-covariance functions

The Lagrangian paradigm, originally formulated in the univariate setting, has been successfully extended to the multivariate realm in a recent unpublished manuscript by Salvaña et al. (2020), where the validity of the Theorem 1 in Ma (2003) was established in the multivariate setting. Particularly, if $\mathbf{C}^S(\mathbf{h})$ is a purely spatial matrix-valued stationary covariance function on $\mathbb{R}^d$, then $\mathbf{C}(\mathbf{h}, u) = \mathrm{E}_{\mathbf{V}}\left\{\mathbf{C}^S(\mathbf{h} - \mathbf{V}u)\right\}, (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$, is a valid spatio-temporal matrix-valued stationary covariance function on $\mathbb{R}^d \times \mathbb{R}$ provided that the expectation exists.

For any u, $\mathbf{C}(\mathbf{h}, u)$ inherits the isotropy of its univariate counterpart, in all its marginals and cross-covariances. Furthermore, when $\mathbf{V} = \mathbf{v}$, $\mathbf{C}(\mathbf{h}, u)$ becomes the multivariate stationary frozen field and it carries the same drawbacks present in the univariate stationary frozen field model. The multivariate stationary non-frozen field provides a natural alternative that is anisotropic for $u \neq 0$ and that allows for dampened maximum cross-covariance at $u \neq 0$. Another option for a Lagrangian stationary cross-covariance function that is anisotropic for $u \neq 0$ is the Lagrangian latent dimension model of Apanasovich and Genton (2010). The fundamental difference between the model of Apanasovich and Genton (2010) and the formulation of Salvaña et al. (2020) is that the former requires a spatio-temporal stationary covariance function and then uses latent dimensions to transform the spatio-temporal stationary covariance function into a spatio-temporal stationary cross-covariance function. In the latter model, one starts with a purely spatial stationary cross-covariance function and turns it into a spatio-temporal stationary cross-covariance function by virtue of the advection velocity vector.

The multivariate setup also opens the question of whether it is permissible that each variable can have a different $\mathbf{V}$ that affects them. One approach to directly solve this problem is to construct a spatio-temporal random field $\mathbf{Z}(\mathbf{s}, t) = \left\{\tilde{Z}_1(\mathbf{s} - \mathbf{V}_1 t), \ldots, \tilde{Z}_p\left(\mathbf{s} - \mathbf{V}_p t\right)\right\}^\top$, where $\tilde{\mathbf{Z}}$ is a multivariate purely spatial stationary random field and $\mathbf{V}_1, \ldots, \mathbf{V}_p$ are random vectors in $\mathbb{R}^p$, which may or may not be correlated. However, in this construction, one cannot keep the stationarity of the cross-covariance function. Another option is to use the LMC with different uncorrelated random vectors $\mathbf{V}_r,\ r = 1, \ldots, R,\ 1 \le R \le p$, for each of the R uncorrelated latent univariate random fields. With this approach, the resulting spatio-temporal cross-covariance function remains stationary.

The last of this progression of models is the nonstationary version of the preceding model, which again, has not yet been proposed in the literature. We develop this new class in Section 6 and we show how we can model more interesting multivariate spatio-temporal random fields.

6. Nonstationary cross-covariance functions under the Lagrangian framework

The theorem below ensures the validity of the Lagrangian spatio-temporal nonstationary cross-covariance functions.

Theorem 1. *Let* $\mathbf{V}$ *be a random vector on* $\mathbb{R}^d$. *If* $\mathbf{C}^S(\mathbf{s}_1, \mathbf{s}_2)$ *is a valid purely spatial matrix-valued nonstationary covariance function on* $\mathbb{R}^d$, *i.e.,* $\mathbf{C}^S(\mathbf{s}_1, \mathbf{s}_2) = \left\{C^S_{ij}(\mathbf{s}_1, \mathbf{s}_2)\right\}_{i,j=1}^p$, *then* $\mathbf{C}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) =$

$E_{\mathbf{V}}\left[\mathbf{C}^S(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2)\right]$, for $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$, and $t_1, t_2 \in \mathbb{R}$, is a valid spatio-temporal matrix-valued nonstationary covariance function on $\mathbb{R}^d \times \mathbb{R}$ provided that the expectation exists.

The proof and all subsequent proofs are relegated to the Appendix. Theorem 1 suggests that one way to build spatio-temporal nonstationary cross-covariance functions is to take any purely spatial nonstationary cross-covariance function and apply a Lagrangian transformation to the coordinates. The resulting covariance function is multivariate, spatio-temporal, nonstationary, and Lagrangian. Further scrutiny is necessary for the models arising from the construction approach in Theorem 1 and is given in Sections 6.1–6.3. A list of stylized examples applying Theorem 1 is also included. Section 6.4 discusses the estimation procedure for Lagrangian models. Section 6.5 illustrates the new cross-covariance function models and their corresponding simulated realizations.

6.1. Lagrangian nonstationary linear model of coregionalization

There are several nonstationary versions of the classical LMC but we select the nonstationary LMC proposed by Fouedjio (2018), as an example, and extend it to space–time using Theorem 1. Let $\mathbf{V}_r, \ \ r = 1, \ldots, R, \ \ 1 \leq R \leq p$, be random vectors on $\mathbb{R}^d$ that characterize the different uncorrelated random advection velocities. If $\mathbf{C}^S(\mathbf{s}_1, \mathbf{s}_2)$ is a valid purely spatial nonstationary LMC on $\mathbb{R}^d$, i.e., $\mathbf{C}^S(\mathbf{s}_1, \mathbf{s}_2) = \sum_{r=1}^R \rho_r \left[\left\{(\mathbf{s}_1 - \mathbf{s}_2)^\top \mathbf{D}_r(\mathbf{s}_1, \mathbf{s}_2)^{-1}(\mathbf{s}_1 - \mathbf{s}_2)\right\}^{1/2}\right] \mathbf{A}_r(\mathbf{s}_1)\mathbf{A}_r(\mathbf{s}_2)^\top$, then

$$\mathbf{C}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = \sum_{r=1}^R \mathrm{E}_{\mathbf{V}_r} \Big\{\rho_r \Big(\Big[\{\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{V}(t_1 - t_2)\}^\top \mathbf{D}_r(\mathbf{s}_1 - \mathbf{V}_r t_1, \mathbf{s}_2 - \mathbf{V}_r t_2)^{-1} \times \{\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{V}(t_1 - t_2)\} \Big]^{1/2} \Big) \mathbf{A}_r(\mathbf{s}_1 - \mathbf{V}_r t_1)\mathbf{A}_r(\mathbf{s}_2 - \mathbf{V}_r t_2)^\top \Big\},$$

where $\rho_r(\cdot)$ is a valid univariate stationary correlation function of a normal scale-mixture type on $\mathbb{R}^d$ and $\mathbf{A}_r$ is a $p \times R$ matrix, is a valid spatio-temporal matrix-valued nonstationary covariance function on $\mathbb{R}^d \times \mathbb{R}$, for any $1 \leq R \leq p$. The condition that $\mathbf{V}_r, \ r = 1, \ldots, R$, are uncorrelated random vectors is set because the underlying univariate random fields are assumed to be uncorrelated. The case in which they may be dependent is on the works.

6.2. Lagrangian spatially varying parameters cross-covariance functions

Introducing spatially varying parameters in a cross-covariance function is a common approach of converting a stationary cross-covariance function to a nonstationary one. The purely spatial nonstationary LMC in Section 6.1 is in fact an example of the spatially varying parameters approach of Paciorek and Schervish (2006) for the normal scale-mixture type of covariance functions. Based on the univariate formulation of Paciorek and Schervish (2006), Kleiber and Nychka (2012) introduced the purely spatial Matérn nonstationary cross-covariance function. By applying Theorem 1, we obtain the Lagrangian spatio-temporal Matérn nonstationary cross-covariance function:

$$\begin{aligned} &C_{ij}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) \\ &= \rho_{ij} \mathrm{E}_{\mathbf{V}} \big\{\sigma_{ij}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2) \\ &\quad \times \left[\{\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{V}(t_1 - t_2)\}^\top \mathbf{D}_{ij}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2)^{-1}\{\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{V}(t_1 - t_2)\}\right]^{\nu_{ij}} \\ &\quad \times \mathcal{K}_{\nu_{ij}} \left[\{\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{V}(t_1 - t_2)\}^\top \mathbf{D}_{ij}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2)^{-1}\{\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{V}(t_1 - t_2)\}\right]\big\}, \end{aligned} \tag{2}$$

for $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$, $t_1, t_2 \in \mathbb{R}$, and $i, j = 1, \ldots, p$. The purely spatial parameters are as follows: $\nu_{ij} > 0$ is the smoothness parameter, $\rho_{ij} \in [-1, 1]$ is the colocated correlation parameter, and $\sigma_{ij}(\cdot)$ is the spatially varying variance parameter, for $i, j = 1, \ldots, p$. Note that ρ_{ij} may also be allowed to vary as a function of its spatial location; see Kleiber and Nychka (2012). Here, $\mathcal{K}_\nu$ is the modified Bessel function of the second kind of order ν, $\mathbf{V}$ is a random vector on $\mathbb{R}^d$, $\sigma_{ij}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2) = |\mathbf{D}_i(\mathbf{s}_1 - \mathbf{V}t_1)|^{1/4}|\mathbf{D}_j(\mathbf{s}_2 - \mathbf{V}t_2)|^{1/4}|\mathbf{D}_{ij}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2)|^{-1/2}$, $\mathbf{D}_{ij}(\mathbf{s}_1, \mathbf{s}_2) = \frac{1}{2}\left\{\mathbf{D}_i(\mathbf{s}_1) + \mathbf{D}_j(\mathbf{s}_2)\right\}$, and $\mathbf{D}_i(\mathbf{s})$ is a $d \times d$ positive definite kernel matrix for variable i, $i = 1, \ldots, p$, at $\mathbf{s}$ that controls the spatially varying local anisotropy and which can be defined via its spectral decomposition, i.e., for $d = 2$:

$$\mathbf{D}_i(\mathbf{s}) = \begin{bmatrix} \cos\{\phi_i(\mathbf{s})\} & -\sin\{\phi_i(\mathbf{s})\} \\ \sin\{\phi_i(\mathbf{s})\} & \cos\{\phi_i(\mathbf{s})\} \end{bmatrix} \begin{bmatrix} \lambda_{1_i}(\mathbf{s}) & 0 \\ 0 & \lambda_{2_i}(\mathbf{s}) \end{bmatrix} \begin{bmatrix} \cos\{\phi_i(\mathbf{s})\} & \sin\{\phi_i(\mathbf{s})\} \\ -\sin\{\phi_i(\mathbf{s})\} & \cos\{\phi_i(\mathbf{s})\} \end{bmatrix},$$

where $\lambda_{1_i}(\mathbf{s}), \lambda_{2_i}(\mathbf{s}) > 0$ are the eigenvalues representing the spatial ranges and $\phi_i(\mathbf{s}) \in (0, \pi/2)$ represents the angle of rotation.

The model in (2) can be generalized to accommodate a time varying $\mathbf{D}_i$ using the following proposition.

Proposition 1. *Let* $\mathbf{V}$ *be a random vector on* $\mathbb{R}^d$ *and let* $\mathbf{D}_i^t(\mathbf{s})$ *be a time varying* $d \times d$ *positive definite kernel matrix at spatial location* $\mathbf{s}$ *and temporal location* t, $i = 1, \ldots, p$. *If* C_{ij}^S *is a valid purely spatial nonstationary cross-covariance function of a normal scale-mixture type on* $\mathbb{R}^d$, *then*

$$\begin{aligned} C_{ij}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = E_{\mathbf{V}} \Big\{ & \sigma_{ij}^{t_1,t_2}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2) \\ & \times C_{ij}^S\Big(\Big[\{\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{V}(t_1 - t_2)\}^\top \mathbf{D}_{ij}^{t_1,t_2}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2)^{-1} \{\mathbf{s}_1 - \mathbf{s}_2 - \mathbf{V}(t_1 - t_2)\}\Big]^{1/2}\Big) \Big\}, \end{aligned} \tag{3}$$

is a valid spatio-temporal nonstationary cross-covariance function on $\mathbb{R}^d \times \mathbb{R}$ *provided that the expectation exists. Here* $\sigma_{ij}^{t_1,t_2}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2) = \left|\mathbf{D}_i^{t_1}(\mathbf{s}_1 - \mathbf{V}t_1)\right|^{1/4} \left|\mathbf{D}_j^{t_2}(\mathbf{s}_2 - \mathbf{V}t_2)\right|^{1/4} \left|\mathbf{D}_{ij}^{t_1,t_2}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2)\right|^{-1/2}$, $\mathbf{D}_{ij}^{t_1,t_2}(\mathbf{s}_1, \mathbf{s}_2) = \frac{1}{2}\left\{\mathbf{D}_i^{t_1}(\mathbf{s}_1) + \mathbf{D}_j^{t_2}(\mathbf{s}_2)\right\}$, *and* $\mathbf{D}_i^t$ *can be defined via its spectral decomposition, i.e., for* $d = 2$ *and* $i = 1, \ldots, p$:

$$\mathbf{D}_i^t(\mathbf{s}) = \begin{bmatrix} \cos\{\phi_i(\mathbf{s}, t)\} & -\sin\{\phi_i(\mathbf{s}, t)\} \\ \sin\{\phi_i(\mathbf{s}, t)\} & \cos\{\phi_i(\mathbf{s}, t)\} \end{bmatrix} \begin{bmatrix} \lambda_{1_i}(\mathbf{s}, t) & 0 \\ 0 & \lambda_{2_i}(\mathbf{s}, t) \end{bmatrix} \begin{bmatrix} \cos\{\phi_i(\mathbf{s}, t)\} & \sin\{\phi_i(\mathbf{s}, t)\} \\ -\sin\{\phi_i(\mathbf{s}, t)\} & \cos\{\phi_i(\mathbf{s}, t)\} \end{bmatrix},$$

where $\lambda_{1_i}(\mathbf{s}, t), \lambda_{2_i}(\mathbf{s}, t) > 0$ *are the eigenvalues representing the spatial ranges at spatio-temporal location* $(\mathbf{s}, t)$ *and* $\phi_i(\mathbf{s}, t) \in (0, \pi/2)$ *represents the angle of rotation at location* $(\mathbf{s}, t)$.

6.3. Lagrangian multivariate deformation model

The univariate deformation model was first proposed by Sampson and Guttorp (1992). No multivariate extensions were proposed since. Here, we present a multivariate extension.

Theorem 2. *If* $\tilde{C}_{ij}^S(\mathbf{s}_1 - \mathbf{s}_2)$ *is a valid purely spatial stationary cross-covariance function on* $\mathbb{R}^d$, *then*

$$C_{ij}^S(\mathbf{s}_1, \mathbf{s}_2) = \tilde{C}_{ij}^S\left\{\|f_i(\mathbf{s}_1) - f_j(\mathbf{s}_2)\|\right\}, \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d,$$

where f_i, $i = 1, \ldots, p$, *represent deterministic non-linear smooth bijective functions of the geographical space onto the deformed space, is a valid purely spatial nonstationary cross-covariance function on* $\mathbb{R}^d$.

The nonstationary cross-covariance functions derived using Theorem 2 naturally yield models with variable asymmetry features. The simplest case is when $f_i = f$, for all $i = 1, \ldots, p$. When applying Theorem 1 to the cross-covariance functions in Theorem 2, we obtain the multivariate Lagrangian spatio-temporal deformation models as follows. Let $\mathbf{V}$ be a random vector on $\mathbb{R}^d$. If $\tilde{C}_{ij}^S(\mathbf{s}_1 - \mathbf{s}_2)$ is a valid purely spatial stationary cross-covariance function on $\mathbb{R}^d$, then

$$C_{ij}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = \mathrm{E}_{\mathbf{V}}\left[\tilde{C}_{ij}^S\left\{\|f_i(\mathbf{s}_1 - \mathbf{V}t_1) - f_j(\mathbf{s}_2 - \mathbf{V}t_2)\|\right\}\right], \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d, \tag{4}$$

where f_i, $i = 1, \ldots, p$, represent deterministic non-linear smooth bijective functions of the geographical space onto the deformed space, is a valid spatio-temporal nonstationary cross-covariance function on $\mathbb{R}^d \times \mathbb{R}$ provided that the expectation exists.

The deformation function can also vary in time, i.e., f can also depend on the temporal location, f^t. As such, the model in (4) can be generalized in the following proposition.

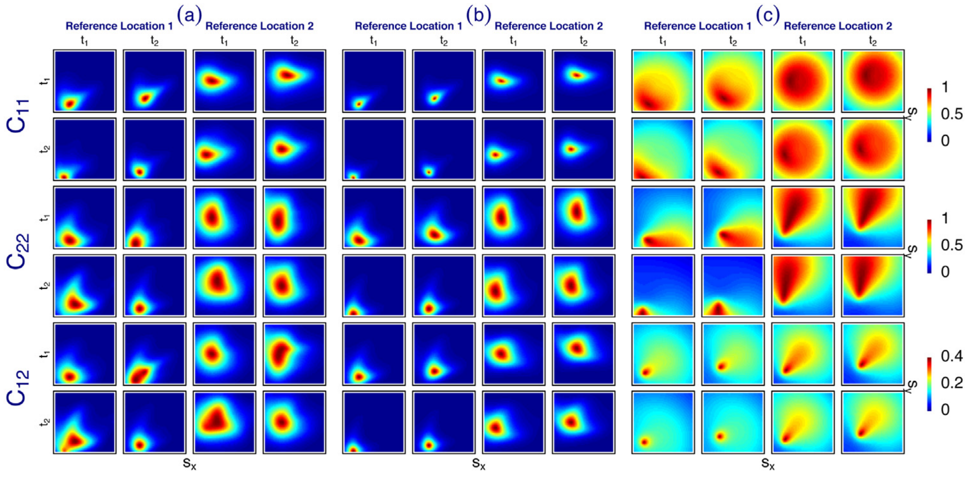

Fig. 1. Heatmaps of the spatio-temporal marginals and cross-covariance functions for the proposed models: (a) frozen Lagrangian nonstationary LMC, (b) frozen Lagrangian spatially varying parameters model, and (c) frozen Lagrangian deformation model. Reference locations 1 and 2 are marked in Fig. 2.

Proposition 2. *Let* $\mathbf{V}$ *be a random vector on* $\mathbb{R}^d$ *and let* f_i^t *be a time-varying deformation function,* $i = 1, \ldots, p$. *If* $\tilde{C}_{ij}^S(\mathbf{s}_1 - \mathbf{s}_2)$ *is a valid purely spatial stationary cross-covariance function on* $\mathbb{R}^d$, *then*

$$C_{ij}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = E_{\mathbf{V}}\left[\tilde{C}_{ij}^S\left\{\|f_i^{t_1}(\mathbf{s}_1 - t_1\mathbf{V}) - f_j^{t_2}(\mathbf{s}_2 - t_2\mathbf{V})\|\right\}\right], \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d, \quad t_1, t_2 \in \mathbb{R}, \qquad (5)$$

is a valid spatio-temporal nonstationary cross-covariance function on $\mathbb{R}^d \times \mathbb{R}$ *provided that the expectation exists.*

6.4. Estimation

For stationary Lagrangian models, estimation can be done via least squares or maximum likelihood and was suggested by Salvaña et al. (2020) to proceed in a multi-step fashion in both approaches, starting with retrieving the purely spatial parameters (marginals and crosses), followed by the advection vector parameters and other temporal parameters. The fact that, for stationary Lagrangian models, we can write out the spatial margins, free of any temporal parameters, is especially convenient for estimation. However, for nonstationary Lagrangian models, a joint estimation of purely spatial and advection vector parameter is necessary, as the spatial and temporal margins can no longer be split up; see Salvaña et al. (2020) and Salvaña and Genton (2020) for a more in-depth discussion of the estimation of Lagrangian spatio-temporal stationary and nonstationary models.

6.5. Illustrations

In Fig. 1, for $p = 2$, we illustrate the different proposed models, and their corresponding bivariate realizations are shown in Fig. 2. In the Lagrangian nonstationary LMC example in Fig. 1(a), we set $\mathbf{V}_1 = \mathbf{v}_1 = (0.1, 0.1)^\top$ and $\mathbf{V}_2 = \mathbf{v}_2 = (-0.1, -0.1)^\top$. The advection velocities in the frozen models in Fig. 1(b) and Fig. 1(c) are set to $\mathbf{V} = \mathbf{v} = (0.1, 0.1)^\top$. In Fig. 1(c), the first deformation is a point-source $f_1(\mathbf{s}) = \mathbf{b} + (\mathbf{s} - \mathbf{b})\|\mathbf{s} - \mathbf{b}\|$, $\mathbf{b} = (0.5, 0.5)^\top$, and the second deformation is of the form $f_2(\mathbf{s}) = \mathbf{b} + (\mathbf{s} - \mathbf{b})\left\{1 + c_1 \exp(-c_2\|\mathbf{s} - \mathbf{b}\|^2)\right\}$, $\mathbf{b} = (0.15, 0.15)^\top$, $c_1 = 6$, and $c_2 = 5$. Here f_2 is the same model used by Iovleff and Perrin (2004) in their simulation study.

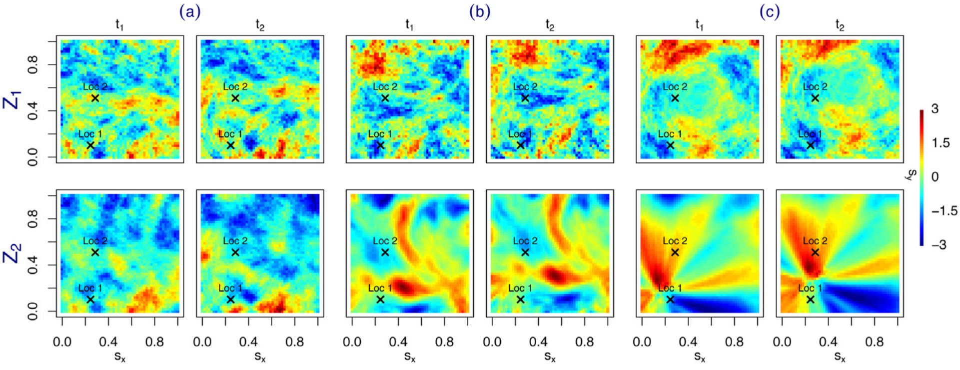

Fig. 2. Simulated realizations for the proposed models: (a) frozen Lagrangian nonstationary LMC, (b) frozen Lagrangian spatially varying parameters model, and (c) frozen Lagrangian deformation model. Two reference locations are represented by crosses, and are used in Fig. 1.

7. Application to regional climate model output

The aforementioned multivariate purely spatial and spatio-temporal stationary and nonstationary models are tested on a bivariate regional climate model output that includes both temperature and precipitation, the two being possibly influenced by transport, on a portion of the Midwest of the United States. The dataset is gridded and covers an area of approximately 1000 km $\times$ 1600 km. The dataset is exactly the same as the one analyzed in Genton and Kleiber (2015). Contrary to the spatio-temporal vantage point we are proposing, after annual trend removal, they treated the yearly temporal replicates of temperature and log precipitation measurements over the summer months (June, July, and August) for the years 1981–2004 as temporally independent and fitted eight bivariate purely spatial stationary and nonstationary covariance function models to the dataset. This approach is limited because it misses the temporal structure due to the fact that the analysis is constrained to be purely spatial. In this work, using the same dataset used by Genton and Kleiber (2015), we view the repeated measurements in time as spatio-temporally dependent.

7.1. Spatio-temporal data analysis

Fig. 3 shows eight consecutive snapshots of the temperature and log precipitation residual fields. From the figure, one can see that the year-on-year spatial profile of the average temperature and log precipitation changes. Specifically, the lowest temperature occurs in different regions every year. Similarly, the region with the highest mean log precipitation varies. Presence of atmospheric flows can cause this phenomenon. The prevailing advection direction can be detected visually by following the blue and red blobs for the temperature and log precipitation residual fields, respectively. Furthermore, the frozen field assumption can be outright dismissed as the transport direction and magnitude seem different for every two consecutive frames.

Fig. 3 also shows that indeed, at any time point, temperature and log precipitation may have equal correlation scales and that the two are negatively correlated. Moreover, the temperature residual field is smoother than the log precipitation residual field and that their smoothness are consistent all throughout the temporal domain under study. Following Genton and Kleiber (2015), let $T(\mathbf{s}, t)$ and $P(\mathbf{s}, t)$ be the temperature and log precipitation residual measurements at spatial location $\mathbf{s}$ and temporal location t. Augmenting their purely spatial bivariate analysis, we seek to find the best bivariate spatio-temporal model for the phenomenon at hand. We fit one bivariate purely spatial and eight bivariate spatio-temporal models as follows:

- M1: Parsimonious bivariate purely spatial Matérn with spatially varying variances and colocated correlation coefficients modeled using thin plate splines.

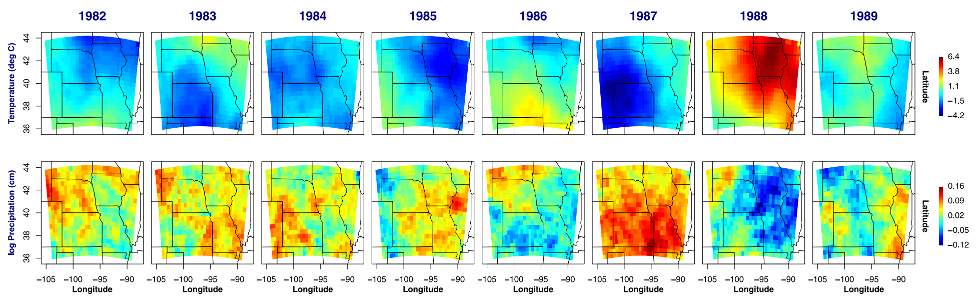

Fig. 3. Bivariate dataset of Genton and Kleiber (2015) with temporal resolution of 92 days (June to August), for the years 1982–1989. The plots in the last column are exactly the plots found in Fig. 1 of Genton and Kleiber (2015). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

- M2: Non-frozen Lagrangian stationary LMC with single advection velocity vector, i.e., $T(\mathbf{s}, t) = A_{11}Z_1(\mathbf{s} - \mathbf{V}_1 t)$ and $P(\mathbf{s}, t) = A_{21}Z_1(\mathbf{s} - \mathbf{V}_1 t) + A_{22}Z_2(\mathbf{s} - \mathbf{V}_1 t)$, where Z_1 and Z_2 are independent mean zero purely spatial processes generated from Matérn correlations, $\mathcal{M}(\mathbf{h}; a_r, \nu_r)$, and $\mathbf{V}_r \sim \mathcal{N}_2(\boldsymbol{\mu}_{\mathbf{V}_r}, \boldsymbol{\Sigma}_{\mathbf{V}_r})$, $r = 1, 2$, are random advection velocity vectors. Here $\mathcal{M}(\mathbf{h}; a, \nu)$ is the univariate Matérn correlation with scale and smoothness parameters a and ν, respectively.
- M3: Non-frozen Lagrangian nonstationary LMC with single advection velocity vector and spatially varying coefficients modeled using thin plate splines, i.e., $T(\mathbf{s}, t) = A_{11}(\mathbf{s} - \mathbf{V}_1 t)Z_1(\mathbf{s} - \mathbf{V}_1 t)$ and $P(\mathbf{s}, t) = A_{21}(\mathbf{s} - \mathbf{V}_1 t)Z_1(\mathbf{s} - \mathbf{V}_1 t) + A_{22}(\mathbf{s} - \mathbf{V}_1 t)Z_2(\mathbf{s} - \mathbf{V}_1 t)$. Z_1 and Z_2 are the same as those in M2.
- M4: Non-frozen Lagrangian stationary LMC with multiple advection velocity vectors, i.e., $T(\mathbf{s}, t) = A_{11}Z_1(\mathbf{s} - \mathbf{V}_1 t)$ and $P(\mathbf{s}, t) = A_{21}Z_1(\mathbf{s} - \mathbf{V}_1 t) + A_{22}Z_2(\mathbf{s} - \mathbf{V}_2 t)$. Z_1 and Z_2 are the same as those in M2.
- M5: Non-frozen Lagrangian nonstationary LMC with multiple advection velocity vectors and spatially varying coefficients modeled using thin plate splines, i.e., $T(\mathbf{s}, t) = A_{11}(\mathbf{s} - \mathbf{V}_1 t)Z_1(\mathbf{s} - \mathbf{V}_1 t)$ and $P(\mathbf{s}, t) = A_{21}(\mathbf{s} - \mathbf{V}_1 t)Z_1(\mathbf{s} - \mathbf{V}_1 t) + A_{22}(\mathbf{s} - \mathbf{V}_2 t)Z_2(\mathbf{s} - \mathbf{V}_2 t)$. Z_1 and Z_2 are the same as those in M2.
- M6: Non-frozen Lagrangian parsimonious bivariate stationary Matérn: for $i, j = 1, 2$,

$$C_{ij}(\mathbf{h}, u) = \rho_{ij}\sigma_i\sigma_j \mathrm{E}_{\mathbf{V}}\left\{\mathcal{M}\left(\mathbf{h} - \mathbf{V}u; a, \nu_{ij}\right)\right\}.$$

- M7: Non-frozen Lagrangian parsimonious bivariate Matérn with spatially varying variances and colocated correlation coefficients modeled using thin plate splines: for $i, j = 1, 2$,

$$C_{ij}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = \mathrm{E}_{\mathbf{V}}\left\{\rho_{ij}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2)\sigma_{ij}(\mathbf{s}_1 - \mathbf{V}t_1, \mathbf{s}_2 - \mathbf{V}t_2)\mathcal{M}\left(\mathbf{h} - \mathbf{V}u; a, \nu_{ij}\right)\right\}.$$

- M8: Bivariate spatio-temporal Gneiting–Matérn of Bourotte et al. (2016) with a frozen Lagrangian parsimonious bivariate stationary Matérn. This model is a linear combination of a bivariate spatio-temporal fully symmetric stationary covariance function and a bivariate spatio-temporal asymmetric stationary covariance function of the form: for $i, j = 1, 2$,

$$C_{ij}(\mathbf{h}, u) = \rho_{ij}\sigma_i\sigma_j \left\{(1 - \Lambda)\frac{1}{\alpha|u|^{2\xi} + 1}\mathcal{M}(\mathbf{h}; a, \nu_{ij}) + \Lambda\mathcal{M}\left(\mathbf{h} - \mathbf{v}u; a, \nu_{ij}\right)\right\},$$

  where $\alpha > 0$ and $\xi \in (0, 1]$ describe the temporal range and smoothness, respectively. Here $\Lambda \in [0, 1]$ is the temporal asymmetry parameter which represents the degree of lack of symmetry in time. This temporal asymmetry parameter is key to detect possible transport effect. When $\Lambda \neq 0$ and $\mathbf{v} \neq \mathbf{0}$, the variables are most likely influenced by an advection velocity and are being transported. The model above is very flexible since a wide range of multivariate spatio-temporal random fields can be modeled, from static to moving.

Table 1
Maximum likelihood parameters estimates of the best performing frozen and non-frozen models in terms of the BIC. The advection velocity parameters are in degrees while the scale parameters a and a_i, $i = 1, 2$, are in kilometers. The spatially varying variance and colocated correlation coefficients are no longer shown.

Model	$\hat{\nu}_1$	$\hat{\nu}_2$	a	a_1	a_2	ξ	α	Δ	$\mathbf{v}$	$\mu_{\mathbf{V}_1}$	$\Sigma_{\mathbf{V}_1}$	$\mu_{\mathbf{V}_2}$	$\Sigma_{\mathbf{V}_2}$
M5	0.261	0.307	–	355	1026	–	–	–	–	$(-1.096, 1.441)^\top$	$\begin{pmatrix} 0.011 & 0 \\ 0 & 0.037 \end{pmatrix}$	$(-1.139, 2.643)^\top$	$\begin{pmatrix} 0.001 & 0 \\ 0 & 0.002 \end{pmatrix}$
M9	1.275	0.602	323	–	–	0.91	360	0.442	$(-0.313, 0.205)^\top$	–	–	–	–

- M9: Bivariate spatio-temporal Gneiting–Matérn of Bourotte et al. (2016) with a frozen Lagrangian parsimonious bivariate Matérn (similar to M8) with spatially varying variances and colocated correlation coefficients modeled using thin plate splines: for $i, j = 1, 2$,

$$C_{ij}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = \rho_{ij}(\mathbf{s}_1 - \mathbf{v}t_1, \mathbf{s}_2 - \mathbf{v}t_2)\sigma_{ij}(\mathbf{s}_1 - \mathbf{v}t_1, \mathbf{s}_2 - \mathbf{v}t_2) \times \left\{(1-\Delta)\frac{1}{\alpha|u|^{2\xi}+1}\mathcal{M}(\mathbf{h}; a, \nu_{ij}) + \Delta\mathcal{M}\left(\mathbf{h} - \mathbf{v}u; a, \nu_{ij}\right)\right\}.$$

A few remarks regarding the chosen models above are in order. Because frozen models generally do not perform well when fitted to random fields that are not frozen, as they do not allow diffusion or dissipation of covariances and cross-covariances at nonzero temporal lags, we do not fit frozen versions of models M2 to M7. We still included, however, a variant of the frozen field models such as models M8 and M9 since the non-Lagrangian portion of the models takes care of the dissipation of covariances and cross-covariances at nonzero temporal lags. For brevity, we limit the nonstationary models to only capture spatially varying variances and cross-correlation coefficients since that was the approach undertaken by Genton and Kleiber (2015), to which we aim to make a comparison regarding purely spatial vs. spatio-temporal fits. Moreover, prior knowledge of the topography of the region under study signifies that a spatially varying variance and colocated correlation coefficients model is sufficient as every site in the region is subjected to almost similar, mainly agricultural, topographical features. The formulation of the LMC models is tailored after the technique of Genton and Kleiber (2015) to bestow on the temperature variable a smoother spatial random field. Finally, unlike Genton and Kleiber (2015), we do not have the luxury of independent spatio-temporal replicates to produce empirical estimates of the spatially varying variance and colocated correlation coefficients. Hence, the spatially varying parameters are assumed to vary smoothly over space and are modeled via thin plate splines; refer to Salvaña and Genton (2020) for a justification of this estimation approach.

7.2. Model performance

The model parameters are estimated via maximum likelihood. The negative log-likelihoods are minimized using the *optim* function with quasi-Newton method "BFGS" in R (R Core Team, 2019). It took approximately 3 h to fit the purely spatial models while fitting spatio-temporal models took 20 h using a 32-core Intel Xeon Gold 6148 with 2.6 GHz clock speed. The interpolation performances of the models are evaluated by the Akaike and Bayesian information criteria. Table 1 collects the maximum likelihood parameter estimates of the best performing frozen and non-frozen models in terms of the BIC (see Table 2). Indeed, the two fields are negatively correlated with an average correlation coefficient of -0.573 under stationary models M6 and M8. The estimated value for the spatio-temporal asymmetry parameter Δ is nonzero, i.e., $\hat{\Delta} = 0.442$. This means that there is a transport behavior that will ultimately be missed if one results to using only spatio-temporal non-Lagrangian models. The advection velocity vector estimates from all the Lagrangian models imply approximately the same Northwest mean direction of transport.

We want to find out if spatio-temporal models have additional benefits over purely spatial models in interpolation and prediction. Comparisons of their performance can be done directly simply by using the likelihood. Also a viable approach in performance comparison between the two modeling paradigms, purely spatial and spatio-temporal, is introducing a magnitude adjustment to

Table 2
A summary of the models and their in-sample (log likelihood, AIC, and BIC) and out-of-sample prediction scores (average RMSE). The in-sample scores were computed using the full data. The lower the AIC, BIC, and RMSE values, the better. The reverse is true for the log likelihood. The best scores are in bold. For concise comparison, we include the fit of three models in Genton and Kleiber (2015) and their corresponding out-of-sample prediction scores.

					Portion of data screened			
	Model	Log likelihood	AIC	BIC	5%	10%	15%	20%
Spatial	M1 (Nonstationary)	67,173	−134, 218	−133, 935	0.079	0.070	0.083	0.080
	Nonstationary Parsimonious Matérn in Genton and Kleiber (2015)	67,242	−134, 476	−134, 446	0.077	0.073	0.078	0.072
	Stationary Parsimonious Matérn in Genton and Kleiber (2015)	66,234	−132, 456	−132, 410	0.078	0.077	0.077	0.080
	Stationary LMC in Genton and Kleiber (2015)	65,611	−131, 208	−131, 155	0.074	0.078	0.079	0.078
Spatio-Temporal	M2 (Stationary)	67,564	−135, 110	−135, 042	0.034	0.047	0.048	0.059
	M3 (Nonstationary)	67,722	−135, 508	−135, 371	0.032	0.047	0.048	0.060
	M4 (Stationary)	68,771	−137, 520	−137, 436	0.028	0.032	0.042	0.049
	M5 (Nonstationary)	**68, 952**	**−137, 864**	**−137, 712**	**0.027**	**0.032**	**0.041**	**0.049**
	M6 (Stationary)	67,435	−134, 854	−134, 793	0.034	0.047	0.048	0.059
	M7 (Nonstationary)	67,499	−134, 674	−133, 928	0.029	0.039	0.043	0.052
	M8 (Stationary)	67,563	−135, 098	−134, 992	0.030	0.042	0.055	0.068
	M9 (Nonstationary)	68,514	−136, 826	−136, 057	0.029	0.036	0.046	0.051

the likelihood function; see Ribatet et al. (2012). This was the approach taken by Sharkey and Winter (2019) to quantify loss of information when fitting purely spatial models given spatio-temporally dependent data. Another approach, which we follow in this paper, is to conduct a pseudo cross-validation study and measure co-kriging performance. In particular, we introduce different degrees of data screening. In the first round, we screen 5% of the 620 available spatial locations, at each t, $t = 1, \ldots, 24$. Then, we increase the number of values screened at an increment of 5%. When a spatial location is chosen to be screened, all the variables observed on that location are screened. At each round, we compute the root mean square error (RMSE). The RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{|\mathcal{S}|T}\sum_{t=1}^{T}\sum_{r\in\mathcal{S}}\|\mathbf{Z}(\mathbf{s}_r, t) - \widehat{\mathbf{Z}}(\mathbf{s}_r, t)\|^2}, \tag{6}$$

where $T = 24$ and $\mathcal{S}$, with cardinality $|\mathcal{S}|$, is the set of screened spatial locations indices and $\mathcal{S}$ does not change across t. Here $\widehat{\mathbf{Z}}(\mathbf{s}_r, t)$ is the vector of predicted values for the two variables at time t at the unobserved location $\mathbf{s}_r$, $r \in \mathcal{S}$, and $\widehat{\mathbf{Z}}(\mathbf{s}_r, t)$ is computed using the co-kriging formula in Section 1. Each round is repeated 1000 times with different sets of randomly chosen screened spatial locations. We expect that spatio-temporal models will have lower average RMSE than the purely spatial ones since they can borrow more information from neighboring temporal sites to more accurately predict screened data. Table 2 summarizes the spatio-temporal co-kriging performances of the different models. The log likelihood values of all spatial and spatio-temporal models are at par with each other. The nonstationary models generally perform better than their stationary counterparts. Furthermore, the spatio-temporal non-frozen Lagrangian models and the frozen Lagrangian model M9 have some of the best interpolation performance, with the non-frozen Lagrangian nonstationary LMC with multiple advection velocities as the preferred model in all metrics. This was expected since the model offers more flexibility by allowing different magnitudes and directions of advection. While the in-sample metrics (log likelihood, AIC, and BIC) provide limited evidence that spatio-temporal modeling should be pursued on this dataset, the out-of-sample metrics in Table 2 say otherwise. The average co-kriging RMSE is less when using spatio-temporal models on this bivariate dataset. Moreover, the discrepancies between the prediction performance of purely spatial and spatio-temporal models are more pronounced as more spatial locations are screened. Hence, we conclude that spatio-temporal models provide a large improvement over the purely spatial models.

8. Purely spatial and spatio-temporal nonstationary cross-covariance functions on the sphere

All the methods previously mentioned in this paper produce valid cross-covariance functions on the sphere when evaluated using the chordal distance. The chordal distance is the length of the shortest straight line between two locations on the sphere. However, the concept of a straight line does not make sense on a sphere. On curved surfaces such as the sphere, the amount of departure that the surfaces make from being a plane should be accounted for (Jacobson and Jacobson, 2005). As a consequence, the shortest path between two locations on the sphere is rightfully represented by a curve or a geodesic. The length of the curve separating two locations on the sphere is called the great circle distance. The great circle distance, however, renders positive definiteness of a covariance function model a serious concern. Furthermore, replacing the great circle distance with chordal distance, to which the latter has an extensive array of positive definite functions in Euclidean space to choose from, may lead to certain problems in interpolation and prediction, as the chordal distance underestimates the great circle distance (Porcu et al., 2016). These complications provided the impetus for extending existing models on the Euclidean space to work with the great circle distance and for developing new methods that work on data obtained on the sphere. Much recent progress has been made including the results of Jeong and Jun (2015), Guinness and Fuentes (2016), Jeong et al. (2017), Porcu et al. (2018), and White and Porcu (2019b). The papers from Arafat Hassan Mohammed (2017) and Guella et al. (2018) provided rigorous characterization of strictly positive definite covariance and cross-covariance functions on the sphere, respectively. An analogue of the univariate purely spatial stationary Matérn that works on the sphere was introduced by Alegría et al. (2018). Variable asymmetry on spherical processes was also studied in that paper. Lastly, White and Porcu (2019a) modeled air pollution using valid models on the sphere.

Other studies detailing the construction and characterization of nonstationary covariance functions on the sphere were published by Jun and Stein (2007, 2008), Hitczenko and Stein (2012), and Jun (2014). A study by Jun (2011) involved deriving models from scalar potentials using differential operators. A physically-motivated construction was used in the models of Fan et al. (2018), specifically for divergence-free and curl-free random vector fields. A paper by Li and Zhu (2016) extended the kernel convolution approach of Paciorek and Schervish (2006) to introduce nonstationary models on the sphere.

Alegría and Porcu (2017) and Porcu et al. (2018) were the first to discuss the validity of the covariance functions under the Lagrangian framework on the sphere, i.e., the covariance functions were evaluated without the use of the Euclidean distance but the great circle distance instead. More specifically, the transport was modeled through a random rotation matrix $\boldsymbol{\mathcal{R}} \in \mathbb{R}^{(d+1)\times(d+1)}$, and not through the random advection velocity vector $\mathbf{V} \in \mathbb{R}^{d+1}$. Moreover, $\boldsymbol{\mathcal{R}}$ was chosen such that it is an orthogonal matrix with a determinant equal to 1. Consider the sphere $\mathbb{S}^2$ with unit radius, i.e., $\mathbb{S}^2 = \left\{\mathbf{s} \in \mathbb{R}^3, \|\mathbf{s}\| = 1\right\}$, where $\|\cdot\|$ is the usual Euclidean distance, as our spatial domain. The spatial location $\mathbf{s} \in \mathbb{R}^3$ has a spherical coordinate representation $\mathbf{s} = (\phi, \theta)^\top$, where $\phi = L\pi/180$ and $\theta = l\pi/180$ are the polar and azimuthal angles, and $(L, l) \in [-90°, 90°] \times [-180°, 180°]$ is the spatial location given in latitude, L, and longitude, l. The Lagrangian spatio-temporal covariance function $C(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = \mathrm{E}_{\boldsymbol{\mathcal{R}}}\left[C^S\left\{d_{GC}(\boldsymbol{\mathcal{R}}^{t_1}\mathbf{s}_1, \boldsymbol{\mathcal{R}}^{t_2}\mathbf{s}_2)\right\}\right], \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{S}^2$, $\boldsymbol{\mathcal{R}} \in \mathbb{R}^{3\times 3}$, where C^S is a purely spatial geodesically isotropic covariance function evaluating its arguments using the great circle distance, $d_{GC}(\mathbf{s}_1, \mathbf{s}_2) = \arccos(\langle \mathbf{s}_1, \mathbf{s}_2\rangle) = \arccos\{\sin\phi_1 \sin\phi_2 + \cos\phi_1 \cos\phi_2 \cos(\theta_1 - \theta_2)\}$, and $\boldsymbol{\mathcal{R}}^t = Q\left[\mathrm{diag}\left\{\exp\left(i\kappa_k t\right)\right\}_{k=1}^3\right]Q^{-1}$, such that the eigenvalues λ_k of $\boldsymbol{\mathcal{R}}$ can be uniquely written as $\lambda_k = \exp(i\kappa_k)$, for $k = 1, 2, 3$, is a valid spatio-temporal covariance function on the sphere provided that the expectation exists. Extending this model to accommodate $p > 1$ variables is straightforward. Let $\boldsymbol{\mathcal{R}}_i$ be a random rotation matrix on $\mathbb{R}^{3\times 3}$, for $i = 1, \ldots, p$. Suppose C_{ij}^S is a valid purely spatial geodesically isotropic cross-covariance function on $\mathbb{S}^2$, then $C_{ij}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = \mathrm{E}_{\boldsymbol{\mathcal{R}}_i, \boldsymbol{\mathcal{R}}_j}\left[C_{ij}^S\left\{d_{GC}(\boldsymbol{\mathcal{R}}_i^{t_1}\mathbf{s}_1, \boldsymbol{\mathcal{R}}_j^{t_2}\mathbf{s}_2)\right\}\right]$, is a valid spatio-temporal cross-covariance function on $\mathbb{S}^2 \times \mathbb{R}$, provided that the expectation exists. The proof can be found in the Appendix.

9. Discussion

Important research progress has been made since the review paper of Genton and Kleiber (2015). Many modeling techniques and approaches were developed, and many research avenues were explored. In this paper, we have reviewed recent advances in the field of multivariate spatio-temporal geostatistics, and presented a variety of models that can adequately describe different behaviors of multivariate spatio-temporal datasets. We devoted a significant part of the paper to introduce and formulate new spatio-temporal covariance models under the Lagrangian framework. The Lagrangian framework provides a recipe for extending purely spatial models to space–time and the models derived from this formulation are generally space–time asymmetric. Although here we attribute the space–time asymmetry to transport caused by an advection velocity, the modeling approach can still be used as long as the space–time asymmetry behavior is observed. The only limitation of the models under the Lagrangian framework is that they are more appropriately applied when the random field is transported.

Ongoing research that is being done in parallel to the writing of this review paper includes the Lagrangian multivariate stationary and nonstationary models on the sphere and the Lagrangian spatio-temporal dimension expansion. Extending the Lagrangian framework to modeling extreme events and non-Gaussian random fields is also in the works. Lagrangian Markov random field models for multivariate lattice data is also an interesting direction to explore.

The Lagrangian formulation of known spatial copula models is an interesting research problem. Significant work has been done in the copula space for multivariate nonstationary random fields. Krupskii and Genton (2019) proposed a new copula model that can capture more complex dependence behavior such as strong joint tail dependence and variable asymmetry. The concept of spatial asymmetry is a feature that is also recently studied in the copula space. Bárdossy and Hörning (2017) offered a procedure in detecting spatial asymmetry by using the concept of reversibility in time series to purely spatial random fields. The novelty of their work lies in the ability to detect directional dependence from a single purely spatial snapshot of a spatio-temporal random field. These two papers may provide a key starting point for constructing methodologies to detect and model space–time asymmetries from a purely spatial dataset.

An additional aim of research should be to identify distributions that lead to explicit forms of the covariance function $C(\mathbf{h}, u) = \mathrm{E}_{\mathbf{V}}\left\{C^S(\mathbf{h} - \mathbf{V}u)\right\}$. A major challenge in this area is to hasten the evaluation of the non-frozen Lagrangian covariance model. The usage of the non-frozen model is inherently difficult because of the presence of the expectation. Since only a few known specialized cases result in explicit forms, sophisticated models have to be evaluated numerically. Development of techniques in performing this numerical evaluation rapidly is also an open problem. Approximations to the non-frozen model may be attempted.

The Lagrangian framework can be used to extend multivariate purely spatial variograms to space–time. Given the relationship between cross-covariance functions and cross-variograms under joint second-order stationarity, i.e., $\boldsymbol{\gamma}(\mathbf{h}) = \mathbf{C}(\mathbf{0}) - \frac{1}{2}\{\mathbf{C}(\mathbf{h}) + \mathbf{C}(-\mathbf{h})\}$, one obtains a spatio-temporal cross-variogram: $\boldsymbol{\gamma}(\mathbf{h}, u) = \mathbf{C}(-\mathbf{v}u) - \frac{1}{2}\{\mathbf{C}(\mathbf{h} - \mathbf{v}u) + \mathbf{C}(\mathbf{v}u - \mathbf{h})\}$. Hence, multivariate models such as those proposed in Chen and Genton (2019) can readily be extended to space–time.

For the sake of conciseness, we only provided sufficient discussions on concepts and models, highlighted only their distinctive features, and restricted the discussions on models with clear avenues for future research. However, three essential topics that were omitted require mention. First, computational issues when fitting massive multivariate spatial and spatio-temporal datasets are a practical consideration that should be addressed. Furthermore, fitting complex models consumes a lot of computing power. This is largely due to heavy parameterization of more complex models. Parameter estimation and prediction become excruciatingly slow as n and p increase. Cost in computation should not far exceed the gain in prediction. Otherwise, there is substantial disincentive in fitting more advanced models. Nevertheless, this challenge presents an opportunity to support usage of sophisticated models on large datasets. Ton et al. (2018) highlighted three viable strategies to overcome scalability issues, including low rank approximations, sparse approximation methods, and spectral methods. Low rank approximations involve approximating the full covariance matrix with a matrix of smaller rank. Often, basis functions at pre-specified knots are utilized for this purpose. A recent work of Kleiber et al. (2019) utilized basis function representations, with coefficients taken from a multivariate lattice process, and gave alternatives to commonly used multivariate purely spatial models. Dimension reduction may also be achieved by clustering via Dirichlet processes. A complete treatment of this model is found in Shirota et al. (2019). Baugh and Stein (2018) proposed an approximation to the full likelihood for purely spatial nonstationary Gaussian processes using recursive skeletonization factorizations. The full recursive skeletonization factorization procedure is laid out in Minden et al. (2017). Litvinenko et al. (2019) introduced the hierarchical matrix or $\mathcal{H}$-matrix approximation of a dense log-likelihood. A known technique in linear algebra, the $\mathcal{H}$-matrix approximation involves partitioning the full covariance matrix into sub-blocks, followed by low-rank approximation of the majority of the sub-blocks.

The second approach, the sparse approximation methods, introduces sparsity in the dense full covariance matrix via compactly supported covariance functions. Hence, for this purpose, a great deal of attention is being given to flexible compactly supported covariance function models and covariance tapering; see Genton and Kleiber (2015) and references therein for a full discussion on this second approach. Porcu et al. (2020) provided spatio-temporal compactly supported models. The compact supports in their models are dynamical in the sense that the compact supports depend on the spatial and temporal lags. Bevilacqua et al. (2016) studied the implications of fitting multivariate covariance tapered models on two fronts: statistical efficiency and computational complexity. They concluded that their proposed models lead to some loss in computational efficiency but kept the estimation equations unbiased. Another alternative to compactly supported covariance functions are the nearest neighbor Gaussian process (NNGP) models (Datta et al., 2016). These models induce sparsity on the full precision matrix and they work under the graphical models framework. Recently, Taylor-Rodriguez et al. (2019) combined this approach with spatial factor models (SFM) to come up with the SF-NNGP model for LIDAR and ground measurements of forest variables, with large p and large n. A specialized treatment is demanded for SFM with large p and large n, but not all variables are observed on the spatial locations under study. This problem was tackled by Ren and Banerjee (2013) using an adaptive Bayesian factor model. Hybrid approaches involving low rank approximations and sparse approximation methods are also done in practice and were thoroughly reviewed in Zhang et al. (2019).

Lastly, spectral approaches exploit the spectral representation of the full covariance matrix. Mosammam (2016) proposed the half spectral composite likelihood targeted for large n problems. His approach involves rewriting the full likelihood as a function of the periodogram and the spectral density function evaluated at $(\mathbf{h}, \tau)$, where $\mathbf{h}$ is the spatial lag and τ is the temporal frequency. This avoids the expensive inversion and determinant computation of the large full covariance matrix. Other spectral approaches are listed in Ton et al. (2018).

When one includes spatial (and temporal) nonstationarity into the mix of complex features present in the data, the models above cannot be appropriately applied as they are defined only in the stationary case. New models addressing large scale multivariate spatio-temporal nonstationary phenomena, similar to the work of Kleiber and Porcu (2015) in the purely spatial stationary case, are demanded.

The scalability issues mentioned in the previous paragraphs may be overcome using high performance computations such as the ExaGeoStat software developed mainly for large n problems with dense full covariance matrices (Abdulah et al., 2018a). ExaGeoStat employs the most advanced parallel architectures, combined with cutting edge dense linear algebra libraries. ExaGeoStat was also fine-tuned to work on the Tile Low-Rank representation of the dense full covariance matrix (Abdulah et al., 2018b).

The second important topic regarding multivariate spatio-temporal modeling which was not yet mentioned in this work is the need for efficient estimation techniques for large n and p problems. A good estimation technique is necessary to provide good prediction performance. Already numerous estimation techniques have been developed: least squares, maximum likelihood, restricted maximum likelihood, composite likelihood, and other nonparametric approaches. However, theoretical developments in estimation techniques in the multivariate nonstationary context lag behind and should be attempted. Tajbakhsh et al. (2019) formulated the generalized sparse precision matrix selection (GSPS) algorithm for fitting variable separable purely spatial cross-covariance models and guaranteed theoretical convergence of the estimators. The GSPS method is predicated on a linear algebra result which states that "if the elements of a matrix show a decay property, then the elements of its inverse also show a similar behavior" (Jaffard, 1990; Benzi, 2016). The GSPS is a two-stage approach. The first stage involves approximating the precision matrix of the full data by an unparameterized sparse matrix using Gaussian Markov random field (GMRF) approximation via maximum likelihood. The second stage entails inversion of the fitted precision matrix and fitting a parametrized cross-covariance matrix via least squares. The convexity of the precision matrix in the first stage makes computation less demanding. Castrillon-Candás et al. (2016) formulated a new set of contrasts for their proposed multi-level restricted maximum likelihood. Horrell and Stein (2015) highlighted the complications brought by the composite likelihood to datasets with enormous spatial and temporal separation lags. According to them, the composite likelihood has no clear criteria in choosing the subsets of the data and their corresponding conditioning sets. In practice, observations with small spatial and temporal lags are grouped together. However, this is not the case with their polar-orbiting satellite dataset. Hence, they developed the Interpolation likelihood or I-likelihood which eradicates all these issues.

Lastly, new constructing principles that are capable of modeling environmental phenomena more realistically, without sacrificing critical features to much simpler assumptions, should be explored. The pervasiveness of large spatio-temporal data has given us the ability to extract even the most hidden features of a dataset. These features should be represented in the spatio-temporal cross-covariance functions. Active areas of work such as Bayesian models and stochastic partial differential equations (SPDE) were not discussed here explicitly, but these offer different perspectives and strategies in modeling.

Appendix. Proofs

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

$$\sum_{l=1}^{n}\sum_{r=1}^{n} \boldsymbol{\lambda}_l^\top \mathbf{C}(\mathbf{s}_l, \mathbf{s}_r; t_l, t_r)\boldsymbol{\lambda}_r = \sum_{l=1}^{n}\sum_{r=1}^{n} \boldsymbol{\lambda}_l^\top \mathrm{E}_{\mathbf{V}}\left\{\mathbf{C}^S\left(\mathbf{s}_l - \mathbf{V}t_l, \mathbf{s}_r - \mathbf{V}t_r\right)\right\}\boldsymbol{\lambda}_r$$
$$= \mathrm{E}_{\mathbf{V}}\left\{\sum_{l=1}^{n}\sum_{r=1}^{n} \boldsymbol{\lambda}_l^\top \mathbf{C}^S\left(\tilde{\mathbf{s}}_l, \tilde{\mathbf{s}}_r\right)\boldsymbol{\lambda}_r\right\} \geq 0$$

for all $n \in \mathbb{Z}^+$ and $\{(\mathbf{s}_1, t_1), \dots, (\mathbf{s}_n, t_n)\} \in \mathbb{R}^d \times \mathbb{R}$, where the last inequality follows from the assumption that $\mathbf{C}^S$ is a valid purely spatial matrix-valued nonstationary covariance function on $\mathbb{R}^d$. □

Proof of Theorem 2. The validity is established by considering a purely spatial random field $\mathbf{Z}(\mathbf{s}) = \left[Z_1\left\{f_1(\mathbf{s})\right\}, \dots, Z_p\left\{f_p(\mathbf{s})\right\}\right]^\top$. □

Proof of Proposition 1. Let $Q_{\mathbf{s}_l,\mathbf{s}_r}^{t_l,t_r} = \left[\{\mathbf{s}_l - \mathbf{s}_r - \mathbf{V}(t_l - t_r)\}^\top \mathbf{D}_{ij}^{t_l,t_r}(\mathbf{s}_l - \mathbf{V}t_l, \mathbf{s}_r - \mathbf{V}t_r)^{-1}\{\mathbf{s}_l - \mathbf{s}_r - \mathbf{V}(t_l - t_r)\}\right]^{1/2}$ and $\boldsymbol{\lambda}_l \in \mathbb{R}^p$. Then:

$$\sum_{l=1}^{n}\sum_{r=1}^{n} \boldsymbol{\lambda}_l^\top \mathbf{C}(\mathbf{s}_l, \mathbf{s}_r; t_l, t_r)\boldsymbol{\lambda}_r$$
$$= \sum_{l=1}^{n}\sum_{r=1}^{n} \boldsymbol{\lambda}_l^\top \mathrm{E}_{\mathbf{V}}\left[\left\{\sigma_{ij}^{t_l,t_r}(\mathbf{s}_l - \mathbf{V}t_l, \mathbf{s}_r - \mathbf{V}t_r)\tilde{C}_{ij}^S\left(Q_{\mathbf{s}_l,\mathbf{s}_r}^{t_l,t_r}\right)\right\}_{i,j=1}^p\right]\boldsymbol{\lambda}_r$$
$$= \mathrm{E}_{\mathbf{V}}\left[\sum_{l=1}^{n}\sum_{r=1}^{n} \boldsymbol{\lambda}_l^\top \left\{\frac{\left|\mathbf{D}_i^{t_l}(\mathbf{s}_l - \mathbf{V}t_l)\right|^{1/4}\left|\mathbf{D}_j^{t_r}(\mathbf{s}_r - \mathbf{V}t_r)\right|^{1/4}}{\left|\frac{\mathbf{D}_i^{t_l}(\mathbf{s}_l - \mathbf{V}t_l) + \mathbf{D}_j^{t_r}(\mathbf{s}_r - \mathbf{V}t_r)}{2}\right|^{1/2}}\int_0^\infty \exp\left(-\omega Q_{\mathbf{s}_l,\mathbf{s}_r}^{t_l,t_r}\right)\right.\right.$$
$$\left.\left.\times\, g_i^{t_l}(\omega, \mathbf{s}_l - \mathbf{V}t_l)g_j^{t_r}(\omega, \mathbf{s}_r - \mathbf{V}t_r)\mathrm{d}\mu(\omega)\right\}_{i,j=1}^p \boldsymbol{\lambda}_r\right] \geq 0,$$

where $g_i^t(\cdot,\cdot)$ is a density on $[0,\infty)$, for $i = 1,\ldots,p$, and we used the representations of the normal scale-mixture. The inequality in the last row follows from the fact that the term inside the expectation is positive definite as explicitly shown in Paciorek and Schervish (2006) and Kleiber and Nychka (2012, Theorem 1 proofs). □

Proof of Proposition 2. Given a multivariate purely spatial random field $\tilde{\mathbf{Z}}(\mathbf{s})$, with second-order nonstationarity, define a multivariate deformed spatio-temporal random field $\mathbf{Z}(\mathbf{s},t) = \left[\tilde{Z}_1\left\{f_1^t\left(\mathbf{s}-\mathbf{V}t\right)\right\},\ldots,\tilde{Z}_p\left\{f_p^t\left(\mathbf{s}-\mathbf{V}t\right)\right\}\right]^\top$, where f_i^t is a temporally varying spatial deformation, $i = 1,\ldots,p$. The covariance between variable i taken at spatio-temporal location $(\mathbf{s}_1,t_1)$ and variable j taken at spatio-temporal location $(\mathbf{s}_2,t_2)$ is

$$\begin{aligned}\operatorname{cov}\left\{Z_i(\mathbf{s}_1,t_1),Z_j(\mathbf{s}_2,t_2)\right\} &= \mathrm{E}_{\mathbf{V}}\left(\operatorname{cov}\left[\tilde{Z}_i\left\{f_i^{t_1}\left(\mathbf{s}_1-\mathbf{V}t_1\right)\right\},\tilde{Z}_j\left\{f_j^{t_2}\left(\mathbf{s}_2-\mathbf{V}t_2\right)\right\}\right]\right)\\ &= \mathrm{E}_{\mathbf{V}}\left[\tilde{C}_{ij}\left\{\|f_i^{t_1}\left(\mathbf{s}_1-\mathbf{V}t_1\right)-f_j^{t_2}\left(\mathbf{s}_2-\mathbf{V}t_2\right)\|\right\}\right]. \qquad \square\end{aligned}$$

Proof of the Lagrangian spatio-temporal cross-covariance in Section 8. In the same line of reasoning as Alegría and Porcu (2017) and Porcu et al. (2018), consider a multivariate purely spatial stationary random field $\tilde{\mathbf{Z}}(\mathbf{s})$ on $\mathbb{S}^2$. Define a multivariate spatio-temporal random field $\mathbf{Z}(\mathbf{s},t) = \{\tilde{Z}_1\left(\boldsymbol{\mathcal{R}}_1^t\mathbf{s}\right),\ldots,\tilde{Z}_p\left(\boldsymbol{\mathcal{R}}_p^t\mathbf{s}\right)\}^\top$. The spatio-temporal cross-covariance between variables i and j taken at spatio-temporal locations $(\mathbf{s}_1,t_1),(\mathbf{s}_2,t_2) \in \mathbb{S}^2\times\mathbb{R}$ is

$$\begin{aligned}\operatorname{cov}\left\{Z_i(\mathbf{s}_1,t_1),Z_j(\mathbf{s}_2,t_2)\right\} &= \mathrm{E}_{\boldsymbol{\mathcal{R}}_i,\boldsymbol{\mathcal{R}}_j}\left[\operatorname{cov}\left\{\tilde{Z}_i\left(\boldsymbol{\mathcal{R}}_i^{t_1}\mathbf{s}_1\right),\tilde{Z}_j\left(\boldsymbol{\mathcal{R}}_j^{t_2}\mathbf{s}_2\right)\right\}\right]\\ &= \mathrm{E}_{\boldsymbol{\mathcal{R}}_i,\boldsymbol{\mathcal{R}}_j}\left[C_{ij}^S\left\{d_{GC}(\boldsymbol{\mathcal{R}}_i^{t_1}\mathbf{s}_1,\boldsymbol{\mathcal{R}}_j^{t_2}\mathbf{s}_2)\right\}\right],\end{aligned}$$

provided that the expectation exists. □

References

Abdulah, S., Ltaief, H., Sun, Y., Genton, M.G., Keyes, D.E., 2018a. ExaGeoStat: A high performance unified software for geostatistics on manycore systems. IEEE Trans. Parallel Distrib. Syst. 29 (12), 2771–2784.

Abdulah, S., Ltaief, H., Sun, Y., Genton, M.G., Keyes, D.E., 2018b. Parallel approximation of the maximum likelihood estimation for the prediction of large-scale geostatistics simulations. In: 2018 IEEE International Conference on Cluster Computing, CLUSTER. IEEE, pp. 98–108.

Ailliot, P., Baxevani, A., Cuzol, A., Monbet, V., Raillard, N., 2011. Space–time models for moving fields with an application to significant wave height fields. Environmetrics 22 (3), 354–369.

Alegría, A., Cuevas, F., Diggle, P., Porcu, E., 2018. A family of covariance functions for random fields on spheres. CSGB Research Report. Department of Mathematics, Aarhus University.

Alegría, A., Porcu, E., 2017. The dimple problem related to space–time modeling under the Lagrangian framework. J. Multivariate Anal. 162, 110–121.

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

Apanasovich, T.V., Genton, M.G., Sun, Y., 2012. A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. J. Amer. Statist. Assoc. 107 (497), 180–193.

Arafat Hassan Mohammed, A., 2017. Mathematical Developments on Isotropic Positive Definite Functions on Spheres (Ph.D. thesis). Universitat Jaume I, Spain.

Bakka, H., Vanhatalo, J., Illian, J.B., Simpson, D., Rue, H., 2019. Non-stationary Gaussian models with physical barriers. Spatial Stat. 29, 268–288.

Bárdossy, A., Hörning, S., 2017. Process-driven direction-dependent asymmetry: identification and quantification of directional dependence in spatial fields. Math. Geosci. 49 (7), 871–891.

Baugh, S., Stein, M.L., 2018. Computationally efficient spatial modeling using recursive skeletonization factorizations. Spatial Stat. 27, 18–30.

Benzi, M., 2016. Localization in matrix computations: Theory and applications. In: Exploiting Hidden Structure in Matrix Computations: Algorithms and Applications. Springer, pp. 211–317.

Bevilacqua, M., Fassò, A., Gaetan, C., Porcu, E., Velandia, D., 2016. Covariance tapering for multivariate Gaussian random fields estimation. Stat. Methods Appl. 25 (1), 21–37.

Bevilacqua, M., Hering, A.S., Porcu, E., 2015. On the flexibility of multivariate covariance models: comment on the paper by Genton and Kleiber. Statist. Sci. 30 (2), 167–169.

Bornn, L., Shaddick, G., Zidek, J.V., 2012. Modeling nonstationary processes through dimension expansion. J. Amer. Statist. Assoc. 107 (497), 281–289.

Bourotte, M., Allard, D., Porcu, E., 2016. A flexible class of non-separable cross-covariance functions for multivariate space–time data. Spatial Stat. 18, 125–146.
Bras, R.L., Rodríguez-Iturbe, I., 1976. Rainfall generation: A nonstationary time-varying multidimensional model. Water Resour. Res. 12 (3), 450–456.
Castrillon-Candás, J.E., Genton, M.G., Yokota, R., 2016. Multi-level restricted maximum likelihood covariance estimation and kriging for large non-gridded spatial datasets. Spatial Stat. 18, 105–124.
Chen, W., Genton, M.G., 2019. Parametric variogram matrices incorporating both bounded and unbounded functions. Stoch. Environ. Res. Risk Assess. 33, 1669–1679.
Choi, M., 2014. Time-Dependent Karhunen-LoèVe Type Decomposition Methods for SPDEs (Ph.D. thesis). Brown University Providence, Rhode Island.
Christakos, G., 2017. Spatiotemporal Random Fields: Theory and Applications. Elsevier.
Christakos, G., Zhang, C., He, J., 2017. A traveling epidemic model of space–time disease spread. Stoch. Environ. Res. Risk Assess. 31 (2), 305–314.
Corzo, G., Varouchakis, E.A., 2019. Spatiotemporal Analysis of Extreme Hydrological Events. Elsevier.
Cox, D.R., Isham, V., 1988. A simple spatial-temporal model of rainfall. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 415, 317–328.
Cressie, N., Burden, S., Davis, W., Krivitsky, P.N., Mokhtarian, P., Suesse, T., Zammit-Mangion, A., 2015. Capturing multivariate spatial dependence: Model, estimate and then predict. Statist. Sci. 30 (2), 170–175.
Cressie, N., Zammit-Mangion, A., 2016. Multivariate spatial covariance models: a conditional approach. Biometrika 103 (4), 915–935.
Datta, A., Banerjee, S., Finley, A.O., Gelfand, A.E., 2016. Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. J. Amer. Statist. Assoc. 111 (514), 800–812.
De Iaco, S., Palma, M., Posa, D., 2019. Choosing suitable linear coregionalization models for spatio-temporal data. Stoch. Environ. Res. Risk Assess. 33 (7), 1419–1434.
Ezzat, A.A., Jun, M., Ding, Y., 2018. Spatio-temporal asymmetry of local wind fields and its impact on short-term wind forecasting. IEEE Trans. Sustain. Energy 9 (3), 1437–1447.
Fan, M., Paul, D., Lee, T.C., Matsuo, T., 2018. Modeling tangential vector fields on a sphere. J. Amer. Statist. Assoc. 113 (524), 1625–1636.
Fouedjio, F., 2017. Second-order non-stationary modeling approaches for univariate geostatistical data. Stoch. Environ. Res. Risk Assess. 31 (8), 1887–1906.
Fouedjio, F., 2018. A fully non-stationary linear coregionalization model for multivariate random fields. Stoch. Environ. Res. Risk Assess. 32 (6), 1699–1721.
Fouedjio, F., Desassis, N., Rivoirard, J., 2016. A generalized convolution model and estimation for non-stationary random functions. Spatial Stat. 16, 35–52.
Fouedjio, F., Desassis, N., Romary, T., 2015. Estimation of space deformation model for non-stationary random functions. Spatial Stat. 13, 45–61.
Fuglstad, G.-A., Simpson, D., Lindgren, F., Rue, H., 2015. Does non-stationary spatial data always require non-stationary random fields? Spatial Stat. 14, 505–531.
Gelfand, A.E., Schmidt, A.M., Banerjee, S., Sirmans, C., 2004. Nonstationary multivariate process modeling through spatially varying coregionalization. TEST 13 (2), 263–312.
Genton, M.G., Kleiber, W., 2015. Cross-covariance functions for multivariate geostatistics (with discussion). Statist. Sci. 30 (2), 147–163.
Gnann, S.J., Allmendinger, M.C., Haslauer, C.P., Bárdossy, A., 2018. Improving copula-based spatial interpolation with secondary data. Spatial Stat. 28, 105–127.
Gneiting, T., 2002. Nonseparable, stationary covariance functions for space–time data. J. Amer. Statist. Assoc. 97 (458), 590–600.
Gneiting, T., Genton, M.G., Guttorp, P., 2007. Geostatistical space–time models, stationarity, separability, and full symmetry. In: Finkenstaedt, B., Held, L., Isham, V. (Eds.), Statistics of Spatio-Temporal Systems. Monographs in Statistics and Applied Probability. Chapman & Hall/CRC Press, Boca Raton, Florida, pp. 151–175.
Gneiting, T., Kleiber, W., Schlather, M., 2010. Matérn cross-covariance functions for multivariate random fields. J. Amer. Statist. Assoc. 105 (491), 1167–1177.
Guella, J.C., Menegatto, V.A., Porcu, E., 2018. Strictly positive definite multivariate covariance functions on spheres. J. Multivariate Anal. 166, 150–159.
Guinness, J., Fuentes, M., 2016. Isotropic covariance functions on spheres: Some properties and modeling considerations. J. Multivariate Anal. 143, 143–152.
Gupta, V.K., Waymire, E., 1987. On Taylor's hypothesis and dissipation in rainfall. J. Geophys. Res.: Atmos. 92 (D8), 9657–9660.
Hitczenko, M., Stein, M.L., 2012. Some theory for anisotropic processes on the sphere. Stat. Methodol. 9 (1–2), 211–227.
Horrell, M.T., Stein, M.L., 2015. A covariance parameter estimation method for polar-orbiting satellite data. Statist. Sinica 25, 41–59.
Hwang, Y., Barut, E., Yeo, K., 2018. Statistical-physical estimation of pollution emission. Statist. Sinica 28, 921–940.
Inoue, T., Sasaki, T., 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.
Iovleff, S., Perrin, O., 2004. Estimating a nonstationary spatial structure using simulated annealing. J. Comput. Graph. Statist. 13 (1), 90–105.
Ip, R.H., Li, W.K., 2015. Time varying spatio-temporal covariance models. Spatial Stat. 14, 269–285.

Ip, R.H., Li, W., 2016. Matérn cross-covariance functions for bivariate spatio-temporal random fields. Spatial Stat. 17, 22–37.
Ip, R.H., Li, W., 2017. A class of valid Matérn cross-covariance functions for multivariate spatio-temporal random fields. Statist. Probab. Lett. 130, 115–119.
Jacobson, M.Z., Jacobson, M.Z., 2005. Fundamentals of Atmospheric Modeling. Cambridge university press.
Jaffard, S., 1990. Propriétés des matrices bien localisées près de leur diagonale et quelques applications. In: Annales de l'Institut Henri Poincaré (C) Non Linear Analysis, vol. 7, Elsevier, pp. 461–476.
Jeong, J., Jun, M., 2015. A class of Matérn-like covariance functions for smooth processes on a sphere. Spatial Stat. 11, 1–18.
Jeong, J., Jun, M., Genton, M.G., 2017. Spherical process models for global spatial statistics. Statist. Sci. 32 (4), 501–513.
Jun, M., 2011. Non-stationary cross-covariance models for multivariate processes on a globe. Scand. J. Stat. 38 (4), 726–747.
Jun, M., 2014. Matérn-based nonstationary cross-covariance models for global processes. J. Multivariate Anal. 128, 134–146.
Jun, M., Stein, M.L., 2007. An approach to producing space–time covariance functions on spheres. Technometrics 49 (4), 468–479.
Jun, M., Stein, M.L., 2008. Nonstationary covariance models for global data. Ann. Appl. Stat. 2 (4), 1271–1289.
Kent, J.T., Mohammadzadeh, M., Mosammam, A.M., 2011. The dimple in Gneiting's spatial-temporal covariance model. Biometrika 98 (2), 489–494.
Kleiber, W., 2016. High resolution simulation of nonstationary Gaussian random fields. Comput. Statist. Data Anal. 101, 277–288.
Kleiber, W., Nychka, D., 2012. Nonstationary modeling for multivariate spatial processes. J. Multivariate Anal. 112, 76–91.
Kleiber, W., Nychka, D., Bandyopadhyay, S., 2019. A model for large multivariate spatial data sets. Statist. Sinica 29, 1085–1104.
Kleiber, W., Porcu, E., 2015. Nonstationary matrix covariances: compact support, long range dependence and quasi-arithmetic constructions. Stoch. Environ. Res. Risk Assess. 29 (1), 193–204.
Krupskii, P., Genton, M.G., 2019. A copula model for non-Gaussian multivariate spatial data. J. Multivariate Anal. 169, 264–277.
Laga, I., Kleiber, W., 2017. The modified Matérn process. Stat 6 (1), 241–247.
Lajaunie, C., Renard, D., Quentin, A., Le Guen, V., Caffari, Y., 2019. A non-Homogeneous model for Kriging dosimetric data. Math. Geosci. 1–17.
Li, B., Murthi, A., Bowman, K.P., North, G.R., Genton, M.G., Sherman, M., 2009. Statistical tests of Taylor's hypothesis: An application to precipitation fields. J. Hydrometeorol. 10 (1), 254–265.
Li, Y., Zhu, Z., 2016. Modeling nonstationary covariance function with convolution on sphere. Comput. Statist. Data Anal. 104, 233–246.
Litvinenko, A., Sun, Y., Genton, M.G., Keyes, D.E., 2019. Likelihood approximation with hierarchical matrices for large spatial datasets. Comput. Statist. Data Anal. 137, 115–132.
Lonij, V.P., Brooks, A.E., Cronin, A.D., Leuthold, M., Koch, K., 2013. Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors. Sol. Energy 97, 58–66.
Lovejoy, S., Mandelbrot, B.B., 1985. Fractal properties of rain, and a fractal model. Tellus A 37 (3), 209–232.
Ma, C., 2003. Families of spatio-temporal stationary covariance models. J. Statist. Plann. Inference 116 (2), 489–501.
Marcotte, D., 2015. TASC3D: a program to test the admissibility in 3D of non-linear models of coregionalization. Comput. Geosci. 83, 168–175.
Marcotte, D., 2019. Some observations on a recently proposed cross-correlation model. Spatial Stat. 30, 65–70.
Minden, V., Damle, A., Ho, K.L., Ying, L., 2017. Fast spatial gaussian process maximum likelihood estimation via skeletonization factorizations. Multiscale Model. Simul. 15 (4), 1584–1611.
Montero, J.-M., Fernández-Avilés, G., Mateu, J., 2015. Spatial and Spatio-Temporal Geostatistical Modeling and Kriging, Vol. 998. John Wiley & Sons.
Mosammam, A.M., 2016. Half spectral composite likelihood approach for estimating spatial–temporal covariance functions. Spatial Stat. 16, 19–34.
Paciorek, C.J., Schervish, M.J., 2006. Spatial modelling using a new class of nonstationary covariance functions. Environmetrics 17 (5), 483–506.
Porcu, E., Alegria, A., Furrer, R., 2018. Modeling temporally evolving and spatially globally dependent data. Internat. Statist. Rev. 86 (2), 344–377.
Porcu, E., Bevilacqua, M., Genton, M.G., 2016. Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. J. Amer. Statist. Assoc. 111 (514), 888–898.
Porcu, E., Bevilacqua, M., Genton, M.G., 2020. Nonseparable space–time covariance functions with dynamical compact supports. Statist. Sinica 30, (in press).
Porcu, E., Gregori, P., Mateu, J., 2006. Nonseparable stationary anisotropic space–time covariance functions. Stoch. Environ. Res. Risk Assess. 21 (2), 113–122.
R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/.
Ren, Q., Banerjee, S., 2013. Hierarchical factor models for large spatially misaligned data: a low-rank predictive process approach. Biometrics 69 (1), 19–30.
Ribatet, M., Cooley, D., Davison, A.C., 2012. Bayesian inference from composite likelihoods, with an application to spatial extremes. Statist. Sinica 22, 813–845.
Risser, M.D., 2015. Spatially-Varying Covariance Functions for Nonstationary Spatial Process Modeling (Ph.D. thesis). The Ohio State University.

Risser, M.D., Calder, C.A., 2015. Regression-based covariance functions for nonstationary spatial modeling. Environmetrics 26 (4), 284–297.
Rodrigues, A., Diggle, P.J., 2010. A class of convolution-based models for spatio-temporal processes with non-separable covariance structure. Scand. J. Stat. 37 (4), 553–567.
Salvaña, M.L., Genton, M.G., 2020. Lagrangian spatio-temporal nonstationary covariance functions. Unpublished Manuscript.
Salvaña, M.L., Lenzi, A., Genton, M.G., 2020. Spatio-temporal cross-covariance functions under the Lagrangian framework. Unpublished Manuscript.
Sampson, P.D., Guttorp, P., 1992. Nonparametric estimation of nonstationary spatial covariance structure. J. Amer. Statist. Assoc. 87 (417), 108–119.
Schlather, M., 2010. Some covariance models based on normal scale mixtures. Bernoulli 16 (3), 780–797.
Shand, L., Li, B., 2017. Modeling nonstationarity in space and time. Biometrics 73 (3), 759–768.
Sharkey, P., Winter, H.C., 2019. A Bayesian spatial hierarchical model for extreme precipitation in Great Britain. Environmetrics 30 (1), e2529.
Shinozaki, K., Yamakawa, N., Sasaki, T., Inoue, T., 2016. Areal solar irradiance estimated by sparsely distributed observations of solar radiation. IEEE Trans. Power Syst. 31 (1), 35–42.
Shirota, S., Gelfand, A.E., Banerjee, S., 2019. Spatial joint species distribution modeling using Dirichlet processes. Statist. Sinica 29 (3), 1127–1154.
Tajbakhsh, S.D., Aybat, N.S., Del Castillo, E., 2019. Generalized sparse precision matrix selection for fitting multivariate Gaussian random fields to large data sets. Statist. Sinica 28, 941–962.
Taylor-Rodriguez, D., Finley, A.O., Datta, A., Babcock, C., Andersen, H.-E., Cook, B.D., Morton, D.C., Banerjee, S., 2019. Spatial factor models for high-dimensional and large spatial data: An application in forest variable mapping. Statist. Sinica 29, 1155–1180.
Tobler, W., 1970. A computer movie simulating urban growth in the detroit region. Econ. Geogr. 46, 234–240.
Ton, J.-F., Flaxman, S., Sejdinovic, D., Bhatt, S., 2018. Spatial mapping with Gaussian processes and nonstationary Fourier features. Spatial Stat. 28, 59–78.
Wang, L., 2008. Karhunen-Loève Expansions and their Applications (Ph.D. thesis). London School of Economics and Political Science, United Kingdom.
Waymire, E., Gupta, V.K., Rodriguez-Iturbe, I., 1984. A spectral theory of rainfall intensity at the meso-β scale. Water Resour. Res. 20 (10), 1453–1465.
White, P., Porcu, E., 2019a. Nonseparable covariance models on circles cross time: A study of Mexico City ozone. Environmetrics 30 (5), e2558.
White, P., Porcu, E., 2019b. Towards a complete picture of stationary covariance functions on spheres cross time. Electron. J. Stat. 13 (2), 2566–2594.
Wikle, C.K., Zammit-Mangion, A., Cressie, N., 2019. Spatio-Temporal Statistics with R. CRC Press.
Xu, H., Gardoni, P., 2018. Improved latent space approach for modelling non-stationary spatial–temporal random fields. Spatial Stat. 23, 160–181.
Zammit-Mangion, A., Rougier, J., Schön, N., Lindgren, F., Bamber, J., 2015. Multivariate spatio-temporal modelling for assessing Antarctica's present-day contribution to sea-level rise. Environmetrics 26 (3), 159–177.
Zhang, B., Sang, H., Huang, J.Z., 2019. Smoothed full-scale approximation of Gaussian process models for computation of large spatial datasets. Statist. Sinica 29, 1711–1737.