Comparing two spatial variables with the probability of agreement

Jonathan Acosta[1], Ronny Vallejos[2], Aaron M. Ellison[3,4], Felipe Osorio[2] and Mário de Castro[5]

[1]Departamento de Estadística, Pontificia Universidad Católica de Chile, Chile
[2]Departamento de Matemática, Universidad Técnica Federico Santa María, Chile
[3]Harvard University Herbaria, Harvard University, Cambridge, MA, USA
[4]Sound Solutions for Sustainable Science, Boston, MA, USA
[5]Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo, São Carlos, SP, Brazil

SUMMARY: Computing the agreement between two continuous sequences is of great interest in statistics when comparing two instruments or one instrument with a gold standard. The probability of agreement (PA) quantifies the similarity between two variables of interest, and it is useful for determining what constitutes a practically important difference. In this article we introduce a generalization of the PA for the treatment of spatial variables. Our proposal makes the PA dependent on the spatial lag. We establish the conditions for which the PA decays as a function of the distance lag for isotropic stationary and nonstationary spatial processes. Estimation is addressed through a first-order approximation that guarantees the asymptotic normality of the sample version of the PA. The sensitivity of the PA with respect to the covariance parameters is studied for finite sample size. The new method is described and illustrated with real data involving autumnal changes in the green chromatic coordinate (G_{cc}), an index of "greenness" that captures the phenological stage of tree leaves, is associated with carbon flux from ecosystems, and is estimated from repeated images of forest canopies.

KEY WORDS: bivariate Gaussian spatial process; spatiotemporal process; covariance functions; G_{cc} index.

1. Introduction

The comparison of two sequences is a fundamental problem in several scientific disciplines, and it can be addressed in many different ways. For example, the Student's t-test, the correlation coefficient, and the Wilcoxon two-sample rank test are three techniques that, under different assumptions, provide information about some aspects of comparisons between two independent populations (Lin et al., 2012). When the goal is to measure the agreement between two variables to validate an assay, a process, or a newly developed instrument, it may be relevant to evaluate whether its performance is concordant with other existing ones or a "gold standard" (Lin et al., 2002).

The probability of agreement (PA) measures the level of agreement between two continuous sequences. It was first introduced in a series of papers by Stevens et al. (Stevens and Anderson-Cook, 2017; Stevens et al., 2017, 2018, 2020; Stevens and Lu, 2020) as an alternative approach to some existing agreement measures, including the concordance correlation coefficient between measurements generated by two different methods (Lin, 1989) and its extensions. Subsequently, Leal et al. (2019) studied the PA in a context of local influence and de Castro and Galea (2021) developed Bayesian PA methods to compare measurement systems with

either homoscedastic or heteroscedastic measurement errors. We note that the literature on PA is extensive and many relevant uses for it in different scenarios have been published. However, in all cases, the general definition is always the probability that the difference between two statistical quantities is negligible and not statistically distinguishable. A salient example is by Ponnet et al. (2021), who adapted the probability of concordance or C-index to the specific needs of a discrete frequency and severity model that typically is used during the technical pricing of a non-life insurance product.

In this paper, we generalize the PA (sensu Lin et al., 2002; Stevens and Anderson-Cook, 2017; Stevens et al., 2017) to the case of two georeferenced sequences in the plane, sequences that also may vary in time. We were motivated to extend and generalize the PA into the spatiotemporal domain because of the widespread need to analyze near-Earth remotely sensed data that are used to make inferences about ongoing climatic change and its effects on environmental dynamics at a range of spatial extents (e.g., Richardson et al., 2007, 2018; Yang et al, 2013). Our exemplar is the problem of comparing and assessing the PA of repeated images taken from fixed platforms to determine changes in the times of year when trees in the northern hemisphere emerge from dormancy and produce new leaves ("spring green-up") and when the leaves of these same trees senesce in the fall before the trees go dormant for the winter. There is substantial evidence that because of ongoing, anthropogenically-driven climate change, spring green-up is, on average, occurring earlier in the year

---

[1]This preprint has not undergone improvements or corrections. The Version of record of this article is published in *Biometrics*.

(e.g., Richardson et al., 2007; Keenan et al, 2014) and leaf senescence is occurring later in the fall (e.g., Moon, 2022). During the "growing season" between spring green-up and fall leaf senescence, forests in the northern hemisphere remove a substantial amount of carbon dioxide from the atmosphere (Barichivich et al., 2012). Changes in the carbon cycle have important implications not only for ecological dynamics and forest processes, but also for economic activities including tourism and markets for carbon offsets (e.g., Le Quéré et al., 2009; Piao et al., 2019). Our spatiotemporal generalization of the PA provides a method to quantify whether these changes are non-negligible and meaningful, and test their statistical significance.

The generalization we describe here makes the PA dependent on a spatial lag, similar to how the variogram and the covariance functions are used in spatial statistics. Specific conditions for the variance of the difference between the two sequences are imposed so that the PA is a decreasing function when the norm of the spatial lag increases. This monotonic feature of the PA is established for the bivariate Matérn and Wendland covariance functions. We then extend the PA for the case of spatiotemporal processes so that two images of the same scene taken at different times –such as annual images of trees leafing out or senescing– can be compared. In our spatiotemporal extension of the PA, we consider that the process includes a temporal and, perhaps, a spatial trend, and random noise. The resulting process is nonstationary in the mean and is flexible enough to account for a number of different trends in time.

Estimation of the PA is addressed via plug-ins and the delta method, assuming that the estimates of the parameters of the covariance function exist and are asymptotically normal (Mardia and Marshall, 1984). A simple expression for the asymptotic variance of the sample PA is also derived. To assess the properties of the PA for finite sample sizes, we carried out two numerical experiments: a sensitivity study that illustrates that PA decreases as a function of the norm of the spatial lag, and a Monte Carlo simulation study to estimate the PA and the relevant parameters of two spatial covariance models. We then applied our spatial PA to a temporal sequence of images of a forest canopy and estimated the PA of a sequence of images taken of the same scene at different times (described in Section 2). Images like these are used routinely to identify seasonal changes in the unfolding, maturation, and senescence (with accompanying fall colors) of individual trees and entire forest canopies, and to estimate fluxes of carbon, water vapor, and other gases between the forest and the atmosphere (e.g., Richardson et al., 2018).

In Section 3 we give some additional, albeit brief, background on PA. We then introduce the idea of PA for spatial processes and establish the main theoretical results for stationary processes (Section 4) and spatiotemporal processes (Section 5). Section 6 discusses estimation of PA, which is then illustrated with numerical simulation experiments (Section 7) and an analysis of the forest-canopy images (Section 8). We conclude with an outline for future research and developments in this area (Section 9). Proofs of the main results are given in the Appendix, the rest of the proofs are relegated to the Supplementary Material of this paper.

2. Imagery

We analyzed a series of 15 annual images of the same scene taken in mid-October from 2008 to 2022 (Figure 1A). This is the time of year when leaves of deciduous trees are senescing, which leads to the spectacular display of fall colors in New England (USA), Japan, and other parts of the northern hemisphere. We chose images taken in a 3-day window each year (12–15 October), as this represents the approximate historical "peak" of fall colors in New England. As the regional climate has warmed, however, this peak has begun to shift towards later dates. Identifying the rate and spatial patterning of this shift is of interest to ecologists, foresters, tourism boards, and economists (e.g., Moon, 2022).

The 15 images we used were taken from the database of the PhenoCam Network,[1] a network of more than 700 fixed observation sites across North America and elsewhere in the world that since 2008 has collected high-frequency and high-resolution imagery with networked digital cameras to track the timing of vegetation change ("phenology") in a range of ecosystems (Seyednasrollah, 2019). We used images from the Harvard Forest, where they have been captured at 30–60-minute intervals since 2008 with a 2048×1636-pixel CMOS sensor in an outdoor StarDot NetCam XL 3MP camera (Richardson, 2021).

Our interest is in changes in the forest canopy, so we clipped a 570×660-pixel rectangular section from each image (Figure 1B). The edges of the clipped image (black rectangle in Figure 1A) were chosen to maximize the size of the clipped image while avoiding sky (top), wires (left) and other instrumentation on the extended boom (lower right). The high-resolution clipped image (Figure 1B) was then downscaled ≈ 200-fold (to 38×44 pixels; Figure 1C) for further analysis and estimation of spatiotemporal PA. Downscaling was done by rasterizing the image using adjoining 15×15-pixel windows; we used the mean RGB value from these windows in the rasterized image (Figure 1C). This downscaling was done for two reasons. First, "reasonable" values of the spatial lag $\|\boldsymbol{h}\|$ (i.e., < 15) would have been within a single leaf of the high-resolution image, and it is of more interest to look at changes among leaves and among entire trees. Second, we determined that estimating the PA of the two high-resolution clipped images would require ≈ 6 PB of RAM, whereas the estimation of the downscaled images, which were close to the same size as those used for our simulation studies (Figures S4 and S5), preserved sufficient visual differences among trees while being computationally more manageable.

We note that our clipped rectangle is different in shape from, but approximately the same size as, the "region of interest" (ROI) defined and analyzed by researchers who use these images for phenological studies (the unmasked area in Figure 1D). The ROI for each PhenoCam site is identified as the area of the image that maximizes the amount of vegetation of interest (i.e., deciduous forest at Harvard Forest) while avoiding sky, topographic features, instrumentation, other human artefacts (e.g., buildings, wires), and other areas of the image that could give seasonally biased results (e.g., soil covered by snow) (Richardson et al., 2007). The
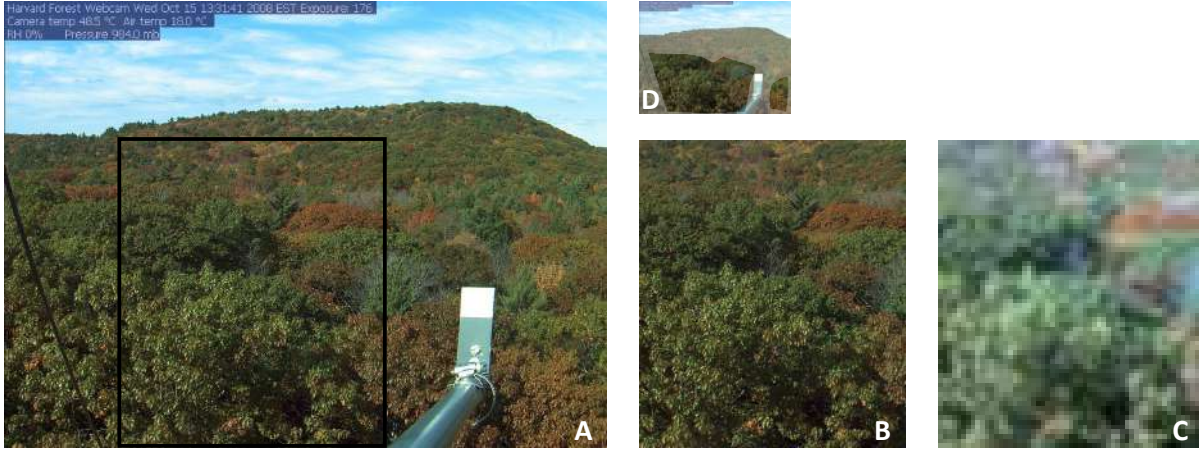
---

[1]https://phenocam.nau.edu/

Figure 1. **A.** The PhenoCam image taken by a stationary camera from the EMS tower at Harvard Forest, Massachusetts, USA on 15 October 2008 at 13:31 (UTC −4). A rectangular section (570 × 660 pixels) of the image was clipped (**B**; black outline in **A**) and then down-scaled (to 38 × 44) pixels (**C**) for estimation of the spatial PA. Note that our rectangular image is different from the clipped "region of interest" (ROI; unmasked area in **D**) analyzed by the PhenoCam network.

original, clipped, and rasterized images, and code used for rasterizing and analyzing these images are all available on GitHub at `https://github.com/JAcosta-Hub/Comparing-two-spatial-variables-with-the-probability-of-agreement`.

3. Mathematical Background and Preliminaries

Following Leal et al. (2019), we assume that $(X_1, Y_1)^\top, \ldots, (X_n, Y_n)^\top$ is a random sample from a bivariate normal distribution with mean vector $\boldsymbol{\mu} = (\mu_X, \mu_Y)^\top$, and covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix},$$

where $(\cdot)^\top$ means transposition. Then, a method to quantify the degree of agreement between the variables X and Y relies on the differences (D_i) between their corresponding values: $D_i = X_i - Y_i$, $i = 1, \ldots, n$. The probability of agreement is defined as

$$\psi_c = \mathrm{P}(|D_i| \leqslant c), \ c > 0, \tag{1}$$

where c denotes the maximum acceptable difference from a practical perspective, also called the equivalence margin. In such a case the interval $(-c, c)$ is often called the region of practical equivalence (ROPE). In clinical studies this interval is called the clinically acceptable difference (CAD).

The selection of c is an important decision that requires careful consideration and domain expertise. In general, the value of c is chosen by the practitioner and is dependent on the specific application and the relative risks that are being considered. In any specific study, the investigator should evaluate the sensitivity of the PA to the chosen value of c. Stevens et al. (2017) created an R Shiny app that enables a user to evaluate the sensitivity of the PA to the chosen value of c. Plotting the PA as a function of the parameters of the model, including the value of c, also can provide valuable information. Likewise, knowledge acquired from Monte Carlo

simulation studies can be highly valuable in understanding the dependency of the PA on c in various scenarios. Further analysis and criteria regarding the use of clinical tolerance limits to evaluate agreement have recently been explored by Taffé (2022).

Because of the normality assumption, the PA as given in (1) takes the form

$$\psi_c = \Phi\left(\frac{c - \mu_D}{\sigma_D}\right) - \Phi\left(-\frac{c + \mu_D}{\sigma_D}\right), \tag{2}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution, $\mu_D = \mu_X - \mu_Y$ and $\sigma_D^2 = \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}$. How large the PA should be to consider the variables interchangeable is up to the practitioner; Stevens et al. (2017) suggest using $\psi_c \geqslant 0.95$ as a guideline.

Under the assumption of normality, inference for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ can be addressed via maximum likelihood (ML) (Anderson, 2003, Section 3.2). By substituting such estimates into (2), an ML estimate for ψ_c, denoted by $\widehat{\psi_c}$, can be obtained. Under mild assumptions, Leal et al. (2019) established the asymptotic normality of $\widehat{\psi_c}$, which relies on the asymptotic distribution of the ML under normality and the delta method. It is then straightforward to obtain approximate confidence intervals and test hypotheses about ψ_c.

A random field, denoted by $\{Y(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{D} \subset \mathbb{R}^d\}$ or $Y(\boldsymbol{s})$, is a collection of random variables indexed by $\mathcal{D}$, where each $\boldsymbol{s} \in \mathcal{D}$ is a location (Cressie, 1993). The mean of $Y(\boldsymbol{s})$ is the function $\mu : \mathbb{R}^d \to \mathbb{R}$ defined through $\mu(\boldsymbol{s}) = \mathrm{E}(Y(\boldsymbol{s}))$ and the covariance function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is defined as $C(\boldsymbol{s}_1, \boldsymbol{s}_2) = \mathrm{cov}(Y(\boldsymbol{s}_1), Y(\boldsymbol{s}_2))$. $Y(\boldsymbol{s})$ is called a weakly stationary random field if the mean function is constant, i.e., $\mu(\boldsymbol{s}) = \mu \in \mathbb{R}$, and the covariance function satisfies $C(\boldsymbol{s}_1, \boldsymbol{s}_2) = C(\boldsymbol{s}_1 - \boldsymbol{s}_2, \boldsymbol{0}) = C_0(\boldsymbol{s}_1 - \boldsymbol{s}_2)$. The covariance function of a stationary process will be denoted as $C(\boldsymbol{h})$, where $\boldsymbol{h} = \boldsymbol{s}_1 - \boldsymbol{s}_2$. Thus, $\mathrm{var}(Y(\boldsymbol{s})) = C(\boldsymbol{0}) = \sigma^2 > 0$, and $\rho(\boldsymbol{h}) = C(\boldsymbol{h})/C(\boldsymbol{0})$ is the correlation function. $Y(\boldsymbol{s})$ is said to be an intrinsically stationary random field if $Y(\boldsymbol{s} + \boldsymbol{h}) - Y(\boldsymbol{s})$ is stationary. For intrinsically stationary random fields, the

mapping $2\gamma(\boldsymbol{h}) = \text{var}\,(Y(\boldsymbol{s}+\boldsymbol{h}) - Y(\boldsymbol{s}))$ is the variogram. For a stationary random field, the relationship $\gamma(\boldsymbol{h}) = C(\boldsymbol{0}) - C(\boldsymbol{h})$ holds. The random field is called isotropic if the covariance function (or the variogram) depends exclusively on the distance between the spatial locations. In such a case, we write $C(\|\boldsymbol{h}\|)$ or $\gamma(\|\boldsymbol{h}\|)$, where $\|\cdot\|$ denotes the Euclidean norm. The *sill* represents the variance of the random field, and it is obtained through $\lim_{\|\boldsymbol{h}\|\to\infty} \gamma(\|\boldsymbol{h}\|) = C(0) = \sigma^2$. The *range* is the distance for which any pair of observations are not correlated, and is obtained as the smallest u such that $\gamma(u) = sill$. In situations when $u \to \infty$, it is useful to define the *practical range* as the smallest u for which the semivariogram reaches the 95% of the *sill*; that is, the value of u that satisfies $\gamma(u) = 0.95\sigma^2$.

4. Probability of Agreement for Stationary Processes

In this section we introduce the PA in the context of geo-referenced variables, where $X(\boldsymbol{s})$ and $Y(\boldsymbol{s})$ are the variables of interest and $\boldsymbol{s}$ represents the spatial coordinates in the plane. Formally, let $\boldsymbol{Z}(\boldsymbol{s}) = (X(\boldsymbol{s}), Y(\boldsymbol{s}))^\top$, $\boldsymbol{s} \in \mathcal{D} \subset \mathbb{R}^2$, be a bivariate second-order stationary random field with $\boldsymbol{s}, \boldsymbol{h} \in \mathbb{R}^2$, mean $(\mu_X, \mu_Y)^\top$, and covariance function

$$\boldsymbol{C}(\boldsymbol{h}) = \begin{pmatrix} C_X(\boldsymbol{h}) & C_{XY}(\boldsymbol{h}) \\ C_{YX}(\boldsymbol{h}) & C_Y(\boldsymbol{h}) \end{pmatrix},$$

where $C_X(\boldsymbol{h}) = \text{cov}(X(\boldsymbol{s}), X(\boldsymbol{s} + \boldsymbol{h}))$, $C_Y(\boldsymbol{h}) = \text{cov}(Y(\boldsymbol{s}), Y(\boldsymbol{s} + \boldsymbol{h}))$ and $C_{XY}(\boldsymbol{h}) = C_{YX}(\boldsymbol{h}) = \text{cov}(X(\boldsymbol{s}), Y(\boldsymbol{s} + \boldsymbol{h}))$. Define the difference

$$D(\boldsymbol{s}, \boldsymbol{h}) = X(\boldsymbol{s}) - Y(\boldsymbol{s} + \boldsymbol{h}). \tag{3}$$

This difference measures the discrepancy between the processes when there is a separation vector equal to $\boldsymbol{h}$ between them. In particular, if $\boldsymbol{h} = \boldsymbol{0}$ the usual difference is recovered and the PA that uses the difference $X(\boldsymbol{s}) - Y(\boldsymbol{s})$ cannot characterize the spatial dependence effectively.

Assume that $\boldsymbol{Z}(\boldsymbol{s})$ is a Gaussian process with mean $\boldsymbol{\mu} = (\mu_X, \mu_Y)^\top$ and covariance function $\boldsymbol{C}(\boldsymbol{h})$, $\boldsymbol{h} \in \mathcal{D}$. Then, $D(\boldsymbol{s}, \boldsymbol{h}) \sim \mathcal{N}(\mu_D, \sigma_D^2(\boldsymbol{h}))$, where $\mu_D = \mu_X - \mu_Y$ and $\sigma_D^2(\boldsymbol{h}) = C_X(\boldsymbol{0}) + C_Y(\boldsymbol{0}) - 2C_{XY}(\boldsymbol{h})$. Then, the PA between processes $X(\boldsymbol{s})$ and $Y(\boldsymbol{s} + \boldsymbol{h})$ is

$$\psi_c(\boldsymbol{h}) = \text{P}(|D(\boldsymbol{s}, \boldsymbol{h})| \leqslant c), \tag{4}$$

which takes the form

$$\psi_c(\boldsymbol{h}) = \Phi\left(\frac{c - \mu_D}{\sigma_D(\boldsymbol{h})}\right) - \Phi\left(-\frac{c + \mu_D}{\sigma_D(\boldsymbol{h})}\right), \tag{5}$$

The probabilities in (4) and (5) can be written, without loss of generality, as $\psi_c(\boldsymbol{h}) = \psi_c(\|\boldsymbol{h}\|)$, indicating the dependency of the PA on the norm of $\boldsymbol{h}$ when processes $X(\boldsymbol{s})$ and $Y(\boldsymbol{s})$, and also $\sigma_D(\boldsymbol{h})$, are all isotropic. In that case, the PA can be plotted as a function of $\|\boldsymbol{h}\|$ in a similar way as the covariance function is plotted for several parametric processes. Then, the PA in (1) can be obtained as a particular case of (4). In fact, $\psi_c = \psi_c(\boldsymbol{0})$.

In the remainder of this section, we emphasize the monotonicity of the PA as a function of the spatial lag $\boldsymbol{h}$. If the PA exhibits a decreasing trend, we can determine the minimum value of $\|\boldsymbol{h}\|$ at which PA stabilizes or is negligible, akin

to how the spatial range is defined for the semivariogram. Moreover, the monotonicity of the PA allows us to construct hypothesis tests about the true PA for a fixed $\boldsymbol{h}$. As we will see in Section 6, the monotonicity of the PA allows us to simplify the test.

We first consider the Matérn covariance function (Matérn, 1986) to illustrate (5). This function is widely used in spatial statistics because of its theoretical properties and its flexibility for modeling local behavior of spatial correlations (Stein, 1999). The Matérn covariance function takes the form

$$M(\boldsymbol{h}, \nu, a) = \frac{2^{1-\nu}}{\Gamma(\nu)}(a\|\boldsymbol{h}\|)^\nu K_\nu(a\|\boldsymbol{h}\|), \tag{6}$$

where $\Gamma(\cdot)$ denotes the gamma function, $K_\nu(\cdot)$ is the modified Bessel function of the second kind, $a > 0$ is a parameter that controls the rate of decay of the correlation, and $\nu > 0$ is the smoothing parameter that is related to the behavior of the correlation near the origin. A special case of the Matérn function is when $\nu = m + 1/2$, $m \in \mathbb{N} \cup \{0\}$. Then,

$$M(\boldsymbol{h}, m + 1/2, a)$$
$$= \exp(-a\|\boldsymbol{h}\|) \sum_{k=0}^{m} \frac{(m+k)!}{(2m)!}\binom{m}{k}(2a\|\boldsymbol{h}\|)^{m-k}. \tag{7}$$

By choosing $m = 0$, we have the simplest form $M(\boldsymbol{h}, 1/2, a) = \exp(-a\|\boldsymbol{h}\|)$.

For a bivariate Gaussian random field, the Matérn covariance function has been extended (Gneiting et al., 2010) as

$$C_X(\|\boldsymbol{h}\|) = \sigma_X^2 M(\boldsymbol{h}, \nu_X, a_X),$$
$$C_Y(\|\boldsymbol{h}\|) = \sigma_Y^2 M(\boldsymbol{h}, \nu_Y, a_Y) \tag{8}$$

and

$$C_{XY}(\|\boldsymbol{h}\|) = \rho_{XY}\sigma_X\sigma_Y M(\boldsymbol{h}, \nu_{XY}, a_{XY}), \tag{9}$$

where $\sigma_X^2 > 0, \sigma_Y^2 > 0$, and ρ_{XY} is the co-located correlation coefficient between $X(\boldsymbol{s})$ and $Y(\boldsymbol{s})$. Specific conditions for the parameters are required so that the covariance model described in (8) and (9) is positive definite. In this model, we have that $\sigma_D^2(\|\boldsymbol{h}\|) = \sigma_X^2 + \sigma_Y^2 - 2\rho_{XY}\sigma_X\sigma_Y M(\boldsymbol{h}, \nu_{XY}, a_{XY})$.

For illustrative purposes, consider $\sigma_X = 1$, $\sigma_Y = 2$, $\sigma_{XY} = 1.8$, $a_{XY} = 2$, $\rho_{XY} = 0.9$ and $\nu = \nu_{XY} = \{0.5, 1.5, 2.5\}$. We plot $\psi_c(\|\boldsymbol{h}\|)$ versus $\|\boldsymbol{h}\|$ for $\|\boldsymbol{h}\| \in \{0, 1, \ldots, 15\}$, $c = \{1.5, 2, 2.5\}$, and using the Matérn covariance function (Figure S1). In all cases, $\psi_c(\|\boldsymbol{h}\|)$ decreases as a function of $\|\boldsymbol{h}\|$ and the curves decay more rapidly to 0 as ν decreases. This is a consequence of the monotonic property of the Matérn covariance as shown in the following example.

EXAMPLE 1: Let $\boldsymbol{Z}(\boldsymbol{s})$ be a bivariate second-order stationary Gaussian process with the Matérn covariance function given in (8) and (9). Assume that $\nu_{XY} = m + 1/2$ and, without loss of generality, assume that the Gaussian process has mean 0 and $a = 1$ in (7). Then,

$$\psi_c(\|\boldsymbol{h}\|) = \Phi\left(\frac{c}{\sigma_D(\|\boldsymbol{h}\|)}\right) - \Phi\left(-\frac{c}{\sigma_D(\|\boldsymbol{h}\|)}\right)$$
$$= 2\Phi\left(\frac{c}{\sigma_D(\|\boldsymbol{h}\|)}\right) - 1.$$

This implies that

$$\psi_c'(\|\boldsymbol{h}\|) = 2\varphi\left(\frac{c}{\sigma_D(\|\boldsymbol{h}\|)}\right)\left(-\frac{c}{\sigma_D(\|\boldsymbol{h}\|)}\right)$$
$$\times \frac{1}{2\sqrt{\sigma_X^2 + \sigma_Y^2 - \rho_{XY}\sigma_X\sigma_Y M(\boldsymbol{h}, \nu_{XY}, a_{XY})}}$$
$$\times \{-\rho_{XY}\sigma_X\sigma_Y M'(\boldsymbol{h}, \nu_{XY}, a_{XY})\}, \tag{10}$$

where $\varphi(\cdot)$ is the probability density function of a standard normal random variable. Because the right hand side of (10) is positive, to prove that $\psi_c(\|\boldsymbol{h}\|)$ is decreasing as a function of $\|\boldsymbol{h}\|$, it is enough to get the sign of the derivative of $M(\boldsymbol{h}, \nu_{XY}, a_{XY})$ with respect to $\|\boldsymbol{h}\|$.

It should be noted that $\psi_c(\|\boldsymbol{h}\|)$ will not necessarily be a decreasing function of $\|\boldsymbol{h}\|$. As an example, consider a bivariate process with mean $(\mu, \mu)^\top$ and an isotropic separable covariance function $\boldsymbol{C}(\|\boldsymbol{h}\|)$, where $C_X(\|\boldsymbol{h}\|) = C_Y(\|\boldsymbol{h}\|) = \sigma^2(\phi/\|\boldsymbol{h}\|)\sin(\|\boldsymbol{h}\|/\phi)$, and $C_{XY}(\|\boldsymbol{h}\|) = \rho_{XY}\sigma^2(\phi/\|\boldsymbol{h}\|)\sin(\|\boldsymbol{h}\|/\phi)$. For simplicity, set $\sigma^2 = 1$ and $\phi = 1$. Then, $\sigma_D^2(h) = 2\left(1 - \rho_{XY}\sin(\|\boldsymbol{h}\|)/\|\boldsymbol{h}\|\right)$. For $c = 1$ and $\rho_{XY} = 1/2$, $\psi_c(\pi/2) = 0.6082$, $\psi_c(3\pi/2) = 0.4986$ and $\psi_c(5\pi/2) = 0.5351$, hence $\psi_c(\|\boldsymbol{h}\|)$ is not decreasing in $\|\boldsymbol{h}\|$.

However, for certain parametric models, as in Example 1, $\psi_c(\|\boldsymbol{h}\|)$ is a monotonic function. A sufficient condition for a parametric covariance model that ensures that $\psi_c(\|\boldsymbol{h}\|)$ is a decreasing monotonic function is given in the following result.

THEOREM 1: *Suppose that $\psi_c(\|\boldsymbol{h}\|)$ is as in (5). If $\sigma_D(\|\boldsymbol{h}\|)$ is an increasing function of $\|\boldsymbol{h}\|$, then $\psi_c(\|\boldsymbol{h}\|)$ is a decreasing function of $\|\boldsymbol{h}\|$.*

Theorem 2 shows that for the bivariate Matérn covariance function in (8) and (9), $\sigma_D^2(\|\boldsymbol{h}\|)$ is an increasing function of $\|\boldsymbol{h}\|$.

THEOREM 2: *Suppose that $\sigma_D^2(\|\boldsymbol{h}\|)$ is obtained using the Matérn covariance model and assume that $\rho_{XY} \geqslant 0$. Then $\sigma_D(\|\boldsymbol{h}\|)$ is an increasing function of $\|\boldsymbol{h}\|$. In consequence, the conditions of Theorem 1 are satisfied and $\psi_c(\|\boldsymbol{h}\|)$ is a decreasing function of $\|\boldsymbol{h}\|$.*

Example 1, therefore, is a direct consequence of Theorems 1 and 2. Moreover, in Theorem 2 it is not necessary to restrict the smoothness parameter of bivariate Matérn covariance model, $\nu_{XY} = m + 1/2$, as in Example 1.

The generalized Wendland family of covariance functions (Gneiting, 2002) is defined, for an integer $\kappa > 0$, as

$$\mathcal{GW}(h; \kappa, \mu) = \frac{1}{B(2\kappa, \mu + 1)}\int_h^1 u(u^2 - h^2)^{\kappa - 1}(1 - u)^\mu du, \tag{11}$$

if $0 \leqslant h < 1$ and is defined as 0, if $h \geqslant 1$, where $B(\cdot, \cdot)$ denotes the beta function and μ must be positive with a lower bound as given in Gneiting (2002). By continuity, for $\kappa = 0$, we have

$$\mathcal{GW}(h; 0, \mu) = \begin{cases} (1 - h^2)^\mu, & \text{if } 0 \leqslant h < 1, \\ 0, & \text{if } h \geqslant 1. \end{cases}$$

Specific conditions for μ can also be found in Bevilacqua et al. (2019).

For a bivariate Gaussian random field, the Wendland-Gneiting covariance function has been extended (Daley et al., 2015) and is defined as

$$C_X(\|\boldsymbol{h}\|) = \sigma_X^2 c_{11}b_{11}^{\nu+2\kappa+1}B(\nu + 2\kappa - 1, \gamma_{11} + 1)$$
$$\mathcal{GW}(\|\boldsymbol{h}\|/b_{11}; \kappa, \nu + \gamma_{11} + 1), \tag{12}$$
$$C_Y(\|\boldsymbol{h}\|) = \sigma_Y^2 c_{22}b_{22}^{\nu+2\kappa+1}B(\nu + 2\kappa - 1, \gamma_{22} + 1)$$
$$\mathcal{GW}(\|\boldsymbol{h}\|/b_{22}; \kappa, \nu + \gamma_{22} + 1), \tag{13}$$
$$C_{XY}(\|\boldsymbol{h}\|) = \rho_{XY}\sigma_X\sigma_Y c_{12}b_{12}^{\nu+2\kappa+1}B(\nu + 2\kappa - 1, \gamma_{12} + 1)$$
$$\mathcal{GW}(\|\boldsymbol{h}\|/b_{12}; \kappa, \nu + \gamma_{12} + 1), \tag{14}$$

where $\sigma_X^2 > 0, \sigma_Y^2 > 0$ and ρ_{XY} is the co-located correlation coefficient between $X(\boldsymbol{s})$ and $Y(\boldsymbol{s})$. Specific conditions for the parameters c_{ii}, b_{ii}, γ_{ii}, $i = 1, 2$, b_{12}, c_{12}, γ_{12}, ν and κ are needed so that the covariance model described in (12)–(14) is positive definite (Daley et al., 2015). In this model, we have that

$$\sigma_D^2(\|\boldsymbol{h}\|) = \sigma_X^2 + \sigma_Y^2 - 2\rho_{XY}\sigma_X\sigma_Y c_{12}b_{12}^{\nu+2\kappa+1}$$
$$\times B(\nu + 2\kappa - 1, \gamma_{12} + 1)\mathcal{GW}\left(\frac{\|\boldsymbol{h}\|}{b_{12}}; \kappa, \nu + \gamma_{12} + 1\right). \tag{15}$$

THEOREM 3: *Suppose that $\sigma_D^2(\|\boldsymbol{h}\|)$ is as in (15), and assume that $\rho_{XY} \geqslant 0$ and $c_{12} \geqslant 0$. Then, $\sigma_D(\|\boldsymbol{h}\|)$ is an increasing function of $\|\boldsymbol{h}\|$. In consequence, the conditions of Theorem 1 are satisfied and $\psi_c(\|\boldsymbol{h}\|)$ is a decreasing function of $\|\boldsymbol{h}\|$.*

5. Probability of Agreement for Spatiotemporal Processes

Assume that $Z(\boldsymbol{s}, t)$, $\boldsymbol{s} \in \mathcal{D} \subset \mathbb{R}^2$, $t \in \mathbb{Z}_0^+$ is a stationary Gaussian spatiotemporal process with mean 0 and covariance function $C(\boldsymbol{h}, u) = \text{cov}(Z(\boldsymbol{s}, t), Z(\boldsymbol{s} + \boldsymbol{h}, t + u))$, where $\boldsymbol{s} = (s_1, s_2)^\top$ represents the spatial coordinates, and t denotes the temporal coordinate.

The correlation function and the semivariogram of the spatiotemporal process are defined analogously to the spatial case, that is, $\rho(\boldsymbol{h}, u) = C(\boldsymbol{h}, u)/C(\boldsymbol{0}, 0)$, and $\gamma(\boldsymbol{h}, u) = C(\boldsymbol{0}, 0) - C(\boldsymbol{h}, u)$ (Sherman, 2011). The practical spatial range, in this case, is the value of $\|\boldsymbol{h}\|$ at which the semivariogram reaches the 95% of the *sill* (i.e. $C(\boldsymbol{0}, 0) = \sigma^2$); that is, the value of $\|\boldsymbol{h}\|$ that satisfies $\sigma^2\{1 - \rho(\|\boldsymbol{h}\|, u)\} = 0.95\sigma^2$, where $\rho(\|\boldsymbol{h}\|, u)$ denotes the correlation function of the spatiotemporal process.

Let $Y(\boldsymbol{s}, t) = \mu(\boldsymbol{s}, t) + Z(\boldsymbol{s}, t)$, where the mean function $\mu(\boldsymbol{s}, t) = \boldsymbol{F}(\boldsymbol{s}, t)^\top\boldsymbol{\beta}$, with $\boldsymbol{F}(\boldsymbol{s}, t)$ known and $\boldsymbol{\beta}$ is a vector of unknown parameters. For example, if $\boldsymbol{F}(\boldsymbol{s}, t) = (1, s_1, s_2, t)^\top$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)^\top$, then $\mu(\boldsymbol{s}, t) = \beta_0 + \beta_1 s_1 + \beta_2 s_2 + \beta_3 t$ represents a linear trend in the spatial and temporal coordinates.

Similarly to the spatial case, let us look at the increments in time and space, defining the difference

$$D(\boldsymbol{s}, t, \boldsymbol{h}, u) = Y(\boldsymbol{s}, t) - Y(\boldsymbol{s} + \boldsymbol{h}, t + u). \tag{16}$$

The quantity defined in (16) measures the discrepancy between the process and itself for a spatial separation $\boldsymbol{h}$ and

temporal separation u. Then, under the Gaussian assumption, $D(\boldsymbol{s}, t, \boldsymbol{h}, u) \sim \mathcal{N}(\mu_D(\boldsymbol{h}, u), \sigma_D^2(\boldsymbol{h}, u))$, where $\mu_D(\boldsymbol{h}, u) = \{\boldsymbol{F}(\boldsymbol{s} + \boldsymbol{h}, t + u) - \boldsymbol{F}(\boldsymbol{s}, t)\}^\top \boldsymbol{\beta}$ and $\sigma_D^2(\boldsymbol{h}, u) = 2C(\boldsymbol{0}, 0) - 2C(\boldsymbol{h}, u) := 2\gamma(\boldsymbol{h}, u)$, and $\gamma(\boldsymbol{h}, u)$ is the spatiotemporal semivariogram (Sherman, 2011). Consequently, a natural extension of the probability of agreement between $Y(\boldsymbol{s}, t)$ and $Y(\boldsymbol{s} + \boldsymbol{h}, t + u)$ is

$$\psi_c(\boldsymbol{h}, u) = \mathrm{P}(|D(\boldsymbol{s}, t, \boldsymbol{h}, u)| \leqslant c). \qquad (17)$$

Therefore, the PA takes the form

$$\psi_c(\boldsymbol{h}, u) = \Phi\left(\frac{c - \mu_D(\boldsymbol{h}, u)}{\sigma_D(\boldsymbol{h}, u)}\right) - \Phi\left(-\frac{c + \mu_D(\boldsymbol{h}, u)}{\sigma_D(\boldsymbol{h}, u)}\right), \quad (18)$$

where $\Phi(\cdot)$ is as in (2). As an illustration, if $\mu(\boldsymbol{s}, t) = \beta_0 + \beta_1 t$ (linear trend in time) and $C(\boldsymbol{h}, u) = \sigma^2 \exp(-\|\boldsymbol{h}\|/\phi_s) \exp(-|u|/\phi_t)$ (separable exponential covariance model) the mean and the variance are

$$\mu_D(\boldsymbol{h}, u) = \beta_1 u \quad \text{and}$$

$$\sigma_D^2(\boldsymbol{h}, u) = 2\sigma^2 \left[1 - \exp\left(-\frac{\|\boldsymbol{h}\|}{\phi_s}\right) \exp\left(-\frac{|u|}{\phi_t}\right)\right].$$

Thus, (18) can be written as

$$\psi_c(\boldsymbol{h}, u) = \Phi\left(\frac{c - u\beta_1}{\sigma_D(\boldsymbol{h}, u)}\right) - \Phi\left(-\frac{c + u\beta_1}{\sigma_D(\boldsymbol{h}, u)}\right).$$

No matter the sign of β_1, $\psi_c(\boldsymbol{h}, u)$ decreases as u increases, which is in agreement with the fact that the PA become smaller when separation over time is enlarged.

Note that we can also use Theorems 1-3 and (18) to consider two different spatial processes in time. In this case, the covariance function has the same structure: $C_X(\boldsymbol{h}) = C_Y(\boldsymbol{h})$. Indeed, by replacing $\sigma_D(\|\boldsymbol{h}\|)$ by $\sigma_D(\|\boldsymbol{h}\|, u)$ in Theorem 1, we obtain the same result if $\sigma_D(\|\boldsymbol{h}\|, u)$ is a increasing function of $\|\boldsymbol{h}\|$ for fixed u. Moreover, for fixed $\|\boldsymbol{h}\|$, if $\sigma_D(\|\boldsymbol{h}\|, u)$ is an increasing function of u then $\psi_c(\|\boldsymbol{h}\|, u)$ is a decreasing function of u. The proof is virtually identical to that of Theorem 1.

In particular, for the separable exponential covariance model, the negative exponential function is always strictly decreasing. Therefore in this case, $\sigma_D^2(\boldsymbol{h}, u)$ is a strictly increasing function of $\|\boldsymbol{h}\|$ for any fixed u, and u similarly increases for any fixed $\boldsymbol{h}$.

6. Estimation

The purpose of this section is to describe the estimation of the PA defined in (5). Here we emphasize that the variance of the difference depends on the parameters of the correlation structure, which we denote as $\sigma_D(\boldsymbol{h}) = \sigma_D(\boldsymbol{h}, \boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \mathbb{R}^q$, and $q \in \mathbb{N}$ is a parameter vector associated with the covariance function. The next definition stresses the dependence of the PA on $\boldsymbol{\theta}$.

DEFINITION 1: Suppose that $(X(\boldsymbol{s}), Y(\boldsymbol{s}))^\top$ is a bivariate second-order stationary random field with $\boldsymbol{s}, \boldsymbol{h} \in \mathbb{R}^2$, mean $(\mu_X, \mu_Y)^\top$, and parametric covariance function $\boldsymbol{C}(\boldsymbol{h}; \theta)$. The PA between processes $X(\boldsymbol{s})$ and $Y(\boldsymbol{s} + \boldsymbol{h})$ is defined through

$$\psi_c(\boldsymbol{h}; \mu_D, \boldsymbol{\theta}) = \Phi\left(\frac{c - \mu_D}{\sigma_D(\boldsymbol{h}; \boldsymbol{\theta})}\right) - \Phi\left(-\frac{c + \mu_D}{\sigma_D(\boldsymbol{h}; \boldsymbol{\theta})}\right),$$

where $\mu_D = \mu_X - \mu_Y$ and $\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta}) = C_X(\boldsymbol{0}; \boldsymbol{\theta}) + C_Y(\boldsymbol{0}; \boldsymbol{\theta}) - 2C_{XY}(\boldsymbol{h}; \boldsymbol{\theta})$.

If $\widehat{\mu}_D$ and $\widehat{\boldsymbol{\theta}}$ are estimators of μ_D and $\boldsymbol{\theta}$, respectively, obtained from the sample $(X(\boldsymbol{s}_1), Y(\boldsymbol{s}_1))^\top, \ldots, (X(\boldsymbol{s}_n), Y(\boldsymbol{s}_n))^\top$, then the plug-in estimator of $\psi_c(\boldsymbol{h}; \mu_D, \boldsymbol{\theta})$ is denoted as $\widehat{\psi}_c(\boldsymbol{h}) := \psi_c(\boldsymbol{h}; \widehat{\mu}_D, \widehat{\boldsymbol{\theta}})$.

LEMMA 1: If $\boldsymbol{V}_{\boldsymbol{\theta}}^{-1/2}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is consistent and $\boldsymbol{V}_{\boldsymbol{\theta}}^{-1/2}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}_q(\boldsymbol{0}, \boldsymbol{I}_q)$ as $n \to \infty$, then $\widehat{\sigma}_D(\boldsymbol{h}; \boldsymbol{\theta}) = \sqrt{C_X(\boldsymbol{0}; \widehat{\boldsymbol{\theta}}) + C_Y(\boldsymbol{0}; \widehat{\boldsymbol{\theta}}) - 2C_{XY}(\boldsymbol{h}; \widehat{\boldsymbol{\theta}})}$ is consistent and asymptotically Gaussian with $\mathrm{E}(\widehat{\sigma}_D(\boldsymbol{h}; \boldsymbol{\theta})) \approx \sqrt{C_X(\boldsymbol{0}; \boldsymbol{\theta}) + C_Y(\boldsymbol{0}; \boldsymbol{\theta}) - 2C_{XY}(\boldsymbol{h}; \boldsymbol{\theta})}$ and

$$\mathrm{var}(\widehat{\sigma}_D(\boldsymbol{h}; \boldsymbol{\theta})) \approx \frac{1}{4\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})} \nabla \sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})^\top \boldsymbol{V}_{\boldsymbol{\theta}} \nabla \sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta}), \quad (19)$$

where $V_{\boldsymbol{\theta}}$ is a $q \times q$ definite positive matrix, $\boldsymbol{I}_q$ is the identity matrix of order q, and ∇ is the gradient operator.

THEOREM 4: Suppose that $(X(\boldsymbol{s}), Y(\boldsymbol{s}))^\top$ is a bivariate stationary Gaussian random field. Assume that $\widehat{\mu}_D$ is consistent, $(\widehat{\mu}_D - \mu_D)/\sqrt{V_{\mu_D}(\boldsymbol{\theta})} \xrightarrow{d} \mathcal{N}(0, 1)$, $\widehat{\boldsymbol{\theta}}$ is consistent, $\boldsymbol{V}_{\boldsymbol{\theta}}^{-1/2}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}_q(\boldsymbol{0}, \boldsymbol{I}_q)$ and $\widehat{\mu}_D$ is independent of $\widehat{\boldsymbol{\theta}}$. Then, $\widehat{\psi}_c(\boldsymbol{h})$ is consistent and asymptotically Gaussian with $\mathrm{E}(\widehat{\psi}_c(\boldsymbol{h})) \approx \psi_c(\boldsymbol{h}; \mu_D, \boldsymbol{\theta})$ and

$$\mathrm{var}(\widehat{\psi}_c(\boldsymbol{h})) \approx \frac{1}{2\pi\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})} \exp\left\{-\frac{(c - \mu_D)^2}{\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})}\right\} \left\{V_{\mu_D} + \frac{(c - \mu_D)^2}{\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})} V_{\sigma_D}\right\}$$

$$+ \frac{1}{2\pi\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})} \exp\left\{-\frac{(c + \mu_D)^2}{\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})}\right\} \left\{V_{\mu_D} + \frac{(c + \mu_D)^2}{\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})} V_{\sigma_D}\right\}$$

$$- \frac{1}{\pi\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})} \exp\left\{-\frac{c^2 + \mu_D^2}{\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})}\right\} \left\{V_{\mu_D} - \frac{c^2 - \mu_D^2}{\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})} V_{\sigma_D}\right\}.$$

As a consequence of the limiting distribution established in Theorem 4, an approximate hypothesis test for the PA can be constructed. Consider the null hypothesis $\mathrm{H}_0 : \psi_c(\|\boldsymbol{h}\|, \mu_D, \boldsymbol{\theta}) = \psi_c^{(0)}$, $0 \leqslant \psi_c^{(0)} \leqslant 1$, versus one of the following three alternative hypotheses $\mathrm{H}_1 : \psi_c(\|\boldsymbol{h}\|, \mu_D, \boldsymbol{\theta}) \neq (> \text{ or } <) \psi_c^{(0)}$, this hypothesis test is relevant because it compares the PA with the nominal value suggested by Stevens et al. (2017) for a fixed $\boldsymbol{h}$. In practice, under the conditions of Theorem 1, the test of $\mathrm{H}_0 : \psi_c(0, \mu_D, \boldsymbol{\theta}) = 0.95$ versus $\mathrm{H}_1 : \psi_c(0, \mu_D, \boldsymbol{\theta}) < 0.95$ can be considered; if H_0 is rejected, then $\mathrm{H}_0 : \psi_c(\|\boldsymbol{h}\|, \mu_D, \boldsymbol{\theta}) = 0.95$ is rejected for all $\|\boldsymbol{h}\|$, because of the monotone property of the PA.

In a spatiotemporal context, denote the covariance function parameterized by $\boldsymbol{\theta}$ as $C(\boldsymbol{h}, u; \boldsymbol{\theta})$. Then, the PA in this case is $\psi_c(\boldsymbol{h}, u; \boldsymbol{\beta}, \boldsymbol{\theta})$, similar to that defined in (18), with $\sigma_D(\boldsymbol{h}, u) = \sigma_D(\boldsymbol{h}, u; \boldsymbol{\theta})$. The plug-in estimator of $\psi_c(\boldsymbol{h}, u; \boldsymbol{\beta}, \boldsymbol{\theta})$ is $\widehat{\psi}_c(\boldsymbol{h}, u) = \psi_c(\boldsymbol{h}, u; \widetilde{\boldsymbol{\beta}}, \widehat{\boldsymbol{\theta}})$. Thus, if $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\theta}}$ are consistent estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$, respectively, Theorem 4 applies to $\widehat{\psi}_c(\boldsymbol{h}, u)$ considering $X(\boldsymbol{s}) = Y(\boldsymbol{s}, t)$ and $Y(\boldsymbol{s}) = Y(\boldsymbol{s}, t + u)$ for a fixed u. The analogous hypothesis testing problem in this context is $\mathrm{H}_0 : \psi_c(\|\boldsymbol{h}\|, u, \boldsymbol{\beta}, \boldsymbol{\theta}) = \psi_c^{(0)}$, $0 \leqslant \psi_c^{(0)} \leqslant 1$, versus one of the following three alternative hypotheses $\mathrm{H}_1 : \psi_c(\|\boldsymbol{h}\|, u, \boldsymbol{\beta}, \boldsymbol{\theta}) \neq (> \text{ or } <) \psi_c^{(0)}$.

7. Numerical Experiments

We carried out two numerical experiments to gain more insights into the properties of the PA for finite sample sizes. The first was a sensitivity analysis that examined how variation in key parameters affected the estimation of the PA for Gaussian random fields with a specific covariance function. The second was a Monte Carlo simulation study, based on 500 runs, that considered spatiotemporal processes with a linear trend and either separable or non-separable covariance structures. In all cases, the estimates were obtained using a pairwise maximum likelihood method implemented in the R programming language version 4.0.5 (R Core Team, 2022). Code is available at https://github.com/JAcosta-Hub/Comparing-two-spatial-variables-with-the-probability-of-agreement.

7.1 *Bivariate Gaussian Random Field*

Let $\boldsymbol{Z}(\boldsymbol{s}) = (X(\boldsymbol{s}), Y(\boldsymbol{s}))^\top$ be a bivariate stationary Gaussian random field with a Matérn covariance function as in (8) and (9), where $\sigma_X^2 = 1$, $\nu_X = \nu_Y = \nu_{XY} = 0.5$, $a_X = a_Y = 1$, $a_{XY} \in \{0.1, 0.15, 0.2, 0.25, 0.3\}$, $\mu_X = 1$, $\mu_Y \in \{0, 0.25, 0.5, 0.75, 1\}$, $\rho_{XY} \in \{0, 0.25, 0.5, 0.75, 1\}$, and $\sigma_Y^2 \in \{0.8, 0.9, 1, 1.1, 1.2\}$. Assuming that $c = 1$ and a fixed covariance parameter, we examined the behavior of the PA, $\psi_c(\|\boldsymbol{h}\|)$ (Figure 2). In all cases, we observed that $\psi_c(\|\boldsymbol{h}\|)$ is a decreasing function of $\|\boldsymbol{h}\|$ and for large $\|\boldsymbol{h}\|$, reaches a fixed value corresponding to the uncorrelated case. As expected, for a fixed $\|\boldsymbol{h}\|$, PA increases with the correlation in the data (ρ_{XY}). Finally, as either μ_D or σ_D increases, PA decreases.

Based on 1000 runs of a bivariate Gaussian random field with separable and nonseparable Matérn covariance functions, the size and the power of the test $H_0 : \psi_c(\boldsymbol{h}, \theta) \geqslant 0.95$ versus $H_1 : \psi_c(\boldsymbol{h}, \theta) < 0.95$ were evaluated for several different scenarios, considering a well/misspecification of covariance functions. Tables S1 and S2 show the estimations, and Figures S1 and S2 show the rejection rates, in all cases the misspecification has a higher rejection rate. More details of the experiment can be found in Appendix C.1.

7.2 *Gaussian Spatiotemporal Process*

The Monte Carlo simulation study considered a spatiotemporal process as described in Section 5. For the purposes of this study, we assumed that $\mu(\boldsymbol{s}, t) = a_0 + a_1 t$ and for the correlation structure, we considered both separable and non-separable cases:

- Exponential (separable):

$$\rho(\boldsymbol{h}, u) = \exp\left(-\frac{\|\boldsymbol{h}\|}{\phi_s}\right) \exp\left(-\frac{|u|}{\phi_t}\right),$$

- Iacocesare (non-separable):

$$\rho(\boldsymbol{h}, u) = \left\{1 + \left(\frac{\|\boldsymbol{h}\|}{\phi_s}\right)^{\alpha_s} + \left(\frac{|u|}{\phi_t}\right)^{\alpha_t}\right\}^{-\beta}$$

where α_s, α_t, and β are positive parameters for Iacocesare model (De Iaco et al, 2002). Because $\rho(\boldsymbol{h}, u)$ is a strictly decreasing function in $\boldsymbol{h}$ and u, Iacocesare's nonseparable covariance function implies that $\sigma_D^2(\boldsymbol{h}, u)$ is increasing in $\|\boldsymbol{h}\|$ for any fixed u, and increasing in u for any fixed $\boldsymbol{h}$. In the absence of a nugget, the covariance function is $C(\boldsymbol{h}, u) = \sigma^2 \rho(\boldsymbol{h}, u)$.

The estimates of the spatial scale (extent) parameter ϕ_s in the covariance terms had very large ranges, extending from < 0 to greater than the maximum size of the grid. Thus, we only estimated PA for values of ϕ_s such that $0 < \phi_s < $ the maximum size of the grid (in this case, 50). More details of the Monte Carlo simulations are given Appendix C.2.

The parameter estimates for the spatiotemporal process with a separable covariance structure and two different values for the trend parameter a_1 (-0.1 and 0.1) are given in Table 1; the estimators were practically unbiased and consistent. The corresponding estimates of PA for different values of h, u and c, and fixed parameters given in Table 1 are shown in Figures S4–S7. The corresponding results from the simulations of a spatiotemporal process with a non-separable covariance structure and identical values for the trend parameter a_1 (-0.1 and 0.1) are given, respectively, in Table 1 and Figures S8–S12, and in Table 1 and Figure 3. As with the separable case, the estimators were practically unbiased and consistent. The model with the Iacocesare covariance had a slightly better performance than the model with the exponential-separable covariance and a higher percent of valid cases used for estimating the parameters. In Tables S6–S12 we show that the PA is decreasing as a function of $\|\boldsymbol{h}\|$ for the spatiotemporal Gaussian processes.

As a consequence of the numerical experiments carried out in Sections 7.1 and 7.2 and Appendices C.1 and C.2, the value of c is strictly related to the scale of the data, thus we suggest that c should be proportional to the scale. That is, $c = \tilde{c} * \sigma$, where σ is the scale of the random field, and $\tilde{c} \in [l, 2.0]$, with $l = 1.6$ (nonseparable case) or $l = 1.7$ (separable case). In practice, this simple rule might be emulated for other covariance structures.

8. Probability of Agreement of the Forest Images

8.1 *Values Used for the Comparisons*

We calculated the green chromatic coordinate (G_{cc}) of each pixel in the ROI of the images (Figure 1C). G_{cc} is an index of "greenness" that captures the phenological stage of tree leaves, is associated with carbon flux from ecosystems, and is estimated from repeated images of forest canopies (Richardson et al., 2018). For a pixel in an RGB image, G_{cc} is calculated as $G_{cc} = G_{DN}/(G_{DN} + R_{DN} + B_{DN})$, where $(\cdot)_{DN}$ is the digital number of the green (G), red (R) and blue (B) channels, respectively, assigned to each pixel in a digital image. The PhenoCam network calculates G_{cc} for each pixel and reports the mean, and the 50th, 75th, and 90th percentiles of the G_{cc} for the ROI of each image, and their 1-day and 3-day running means and percentiles (Richardson et al., 2018). Our estimates of the mean G_{cc} calculated by the PhenoCam network for the corresponding ROI fell within the range of G_{cc} values for each pixel of our clipped and rasterized images (Figure 4).

8.2 *Estimating the PA of the Forest Images*

The empirical variogram of the original data showed strong temporal dependence, so we used linear regression independent of spatial covariance to remove the effect of a deterministic trend (slope = 0.00235, intercept = 0.3854; *p*-value
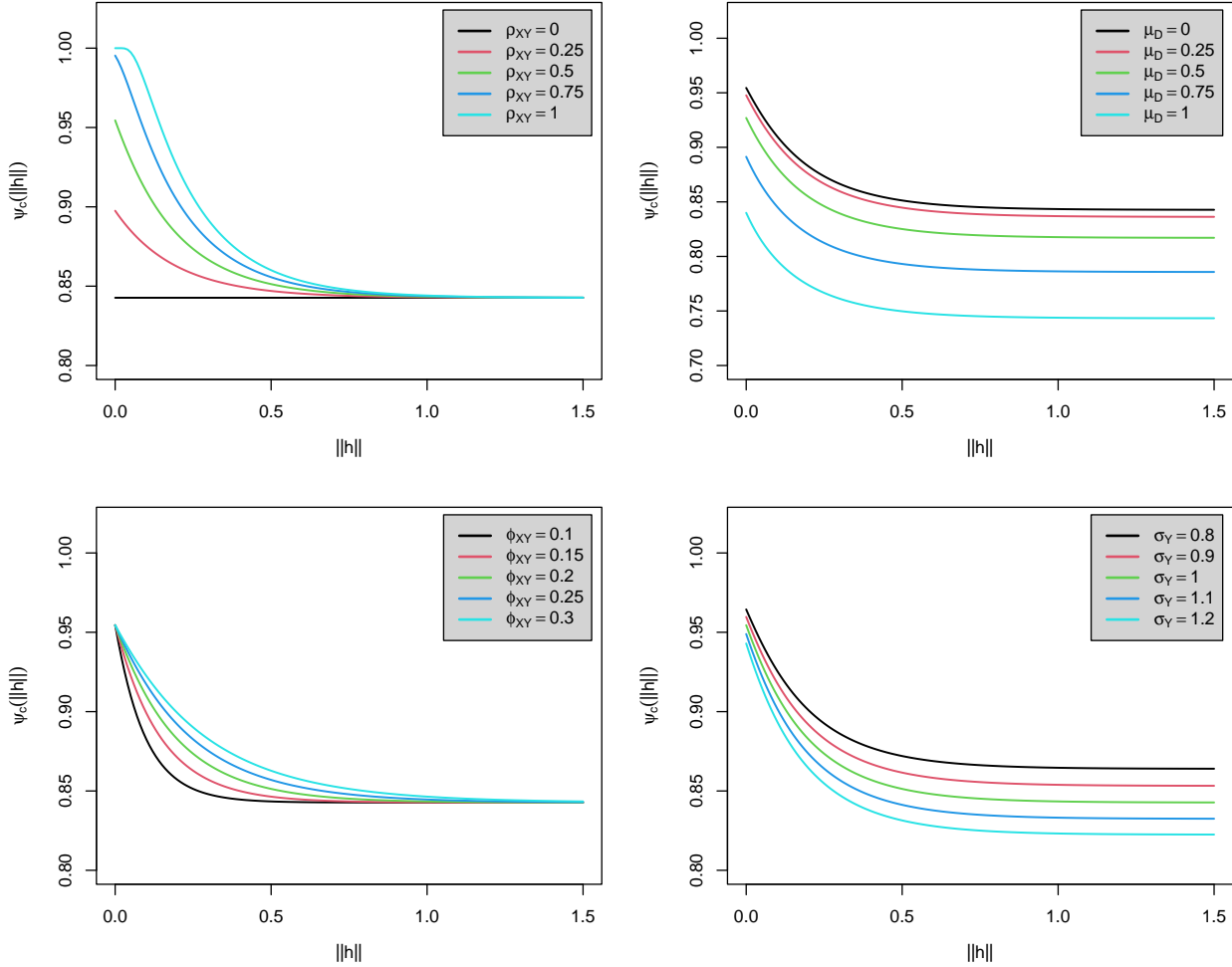
Figure 2. The effect of variation in ρ_{XY} (top left), μ_D (top right), ϕ_{XY} (bottom left), and σ_Y (bottom right), on the behavior of PA as a function of lag $\|h\|$ in a bivariate Gaussian random field with a Matérn covariance structure. Note differences in range limits of the y-axis among the four panels.

< 0.0001 for both). This change in G_{cc} corresponds to a delay in leaf senescence on the order of 0.5–0.7 days/yr, marginally larger than the delay estimated through 2010 by Gill et al. (2015) when climatic warming was somewhat slower than it is now. Figure S13 shows the marginal empirical variograms after this trend had been removed (i.e., the empirical variogram of the residuals of the simple linear regression).

To model G_{cc}, we assume that G_{cc} is a realization of a spatiotemporal process where the spatial coordinates, $\boldsymbol{s} = (s_1, s_2)^\top$ are the pixel positions, and t is the temporal instant when the photo was taken. We also assume that the mean is linear and the covariance function is isotropic and stationary. That is, $G_{cc} := G_{cc}(\boldsymbol{s}, t) = \beta_0 + \beta_1 t + Z(\boldsymbol{s}, t)$, where $Z(\boldsymbol{s}, t)$ is a stationary Gaussian spatiotemporal process with mean 0 and covariance function $C(\|\boldsymbol{h}\|, u)$ of the spatial lag $\|\boldsymbol{h}\|$ and the temporal lag u. In this case, $F(\boldsymbol{s}, t) = (1, t)^\top$ and $\boldsymbol{\beta} = (\beta_0, \beta_1)^\top$, so $\mu(\boldsymbol{s}, t) = \beta_0 + \beta_1 t$. We considered different covariance models (separable and non-separable), each with fixed nugget effect equal to 0.

The values that maximize the objective function (log com-posite likelihood) for the exponential and Iacocesare models, respectively, were $19\,668\,351.95$ and $19\,682\,852.64$, and the values of the associated pseudo-Akaike information criterion were, respectively, $-39\,336\,693.90$ and $-39\,365\,689.29$. These results suggested a better fit to the data when using the Iacocesare non-separable covariance model. The parameter estimates (and the values we used to initialize the estimation routine) for the temporal trend and the covariance model are given in Table 2.

The practical spatial range, defined as the distance at which 95% of the sill (i.e., σ^2 in this case) is reached, will depend on the temporal separation. For the Iacocesare covariance model, this distance can be obtained by using the estimates of the parameters in this equation:

Practical range $=$

$$\phi_s \left\{ 20^{1/\beta} - 1 - \left(\frac{|u|}{\phi_t} \right)^{\alpha_t} \right\}^{1/\alpha_s}, \quad 20^{1/\beta} - 1 - \left(\frac{|u|}{\phi_t} \right)^{\alpha_t} > 0.$$

Using the estimates given in Table 2 and setting $u = 0$

Table 1

Mean and standard deviation (SD) of the estimates for the spatiotemporal process with an exponential separable and a non-separable Iacocesare covariance structure. The cases of positive and negative slope are included. "Percent valid" is the percentage of simulations for which estimates of ϕ_s were greater than 0 and less than the maximum size of the grid.

Covariance model	(N_S, N_T)		a_0	a_1	ϕ_s	ϕ_t	σ^2	α_s	α_t	β	Percent valid
		True	0.500	-0.100	6.676	1.000	0.100				
	$(20, 10)$	Average	0.510	-0.101	7.714	0.789	0.090				77.6
		SD	0.145	0.023	3.342	0.231	0.014				
	$(50, 10)$	Average	0.504	-0.101	7.944	0.952	0.097				74.0
Exponential		SD	0.070	0.011	3.816	0.126	0.008				
		True	0.500	0.100	6.676	1.000	0.100				
	$(20, 10)$	Average	0.511	0.098	7.437	0.830	0.090				82.4
		SD	0.145	0.024	2.974	0.226	0.014				
	$(50, 10)$	Average	0.506	0.099	7.548	0.940	0.096				84.4
		SD	0.063	0.010	3.706	0.113	0.007				
		True	0.500	-0.100	6.676	1.000	0.100	1.000	1.000	2.000	
	$(20, 10)$	Average	0.513	-0.102	7.250	0.987	0.091	1.891	2.180	2.534	62.0
		SD	0.132	0.019	2.475	0.887	0.009	2.246	2.106	1.730	
	$(50, 10)$	Average	0.506	-0.101	7.092	1.104	0.096	1.698	1.358	2.268	85.6
Iacocesare		SD	0.086	0.011	2.407	0.750	0.005	1.738	0.895	1.209	
		True	0.500	0.100	6.676	1.000	0.100	1.000	1.000	2.000	
	$(20, 10)$	Average	0.493	0.099	6.896	0.979	0.090	1.485	2.413	2.881	67.2
		SD	0.132	0.020	2.272	0.641	0.009	2.009	2.173	2.299	
	$(50, 10)$	Average	0.504	0.101	6.778	1.051	0.096	1.452	1.497	2.217	82.8
		SD	0.082	0.012	2.072	0.579	0.005	1.765	1.128	1.055	

(no time lag), the estimated practical range is 22.1 pixels. In addition, for $u = 8$ the practical range was approximately 0.

Finally, Figure 5 illustrates estimates of the PA of the G_{cc} between images as a function of spatial lag $\|\boldsymbol{h}\|$ for four different time lags (u) and four different maximum acceptable differences c. Regardless of the values of u and c, PA decreases with increasing $\|\boldsymbol{h}\|$, but the rate of decrease declines rapidly with increasing u. We note that if $u = 0$, we are working with a single image (not comparing two images at different times); in such a case PA computes the probability that the leaves have the same color at different spatial locations in that image. In contrast, when $u > 0$, we are comparing images through time, and indeed, the PA (i.e., color of a leaf either at the same location or at different locations) get worse as u increases. The PA is more sensitive to u than $\|\boldsymbol{h}\|$ because the mean of the process has a linear temporal trend. This is consistent with the hypothesis test described in Appendix C.3. In all cases the p-value is 0 (see Figure S15), hence the hypothesis that $\psi_c(\|\boldsymbol{h}\|, u)$ is equal to 0.95 is rejected, except for $u = 0$, where, for example, for $c = 0.04$ the null hypothesis is not rejected until $\|\boldsymbol{h}\| \approx 4$. The observed results imply that the timing of leaf senescence is changing (as indicated by lack of agreement in G_{cc} at a given location through time). In this case, because G_{cc} is increasing through time (Figure 4), leaves are staying greener later and senescence is delayed relative to previous years. We note that for $u > 1$, PA is practically constant as a function of $\|\boldsymbol{h}\|$, which we interpret to mean that the PA of the spatiotemporal processes separated by two or more years is not affected by temporal or spatial changes.

9. Discussion and future work

The probability of agreement has been generalized for the analysis of concordance between two georeferenced variables.

This extension possesses monotonic properties—the PA declines as a function of the norm of the spatial lag—and thus quantifies the effective (or practical) spatial range. Our spatial PA is meaningful for isotropic processes. The hypothesis testing developed in Section 6 allows the estimation of the PA as a function of the spatial range, and provides a way to rule out spatial agreement when the null hypothesis is rejected for all $\|h\|$. Our theoretical extension of the PA also works for spatiotemporal processes with a trend, allowing for the analysis of nonstationary processes in the mean.

The monotonic properties of the PA for finite sample size were supported by Monte Carlo simulation experiments, which also showed that the parameter estimates using the composite likelihood have small bias and low variance. The value of c plays a crucial role in the estimation and the PA is sensitive to the choice of it. In practice, the value of c needs to be scaled by the square root of the sill in order to account for the scale of the data.

The application presented in Section 8 illustrated that the PA can describe the change of a spatiotemporal variable in time while accounting for spatial and temporal dependence. In this particular case, the PA also provided information about the dependence of the trend on the spatial information of the past realizations of the process; such information may be of value in modeling the trend and developing such models should be a focus of future work. Although parameter estimates were not very sensitive to the choice of the covariance function, identification of the best covariance model (e.g., separable or non-separable, and types of each) can still be improved. If the trend cannot be modeled easily with a well-known function, prior exploration will be needed to characterize a parametric function that accurately identifies observed patterns. Nonparametric trend estimation (e.g., Strandberg et al., 2019) could also be used.
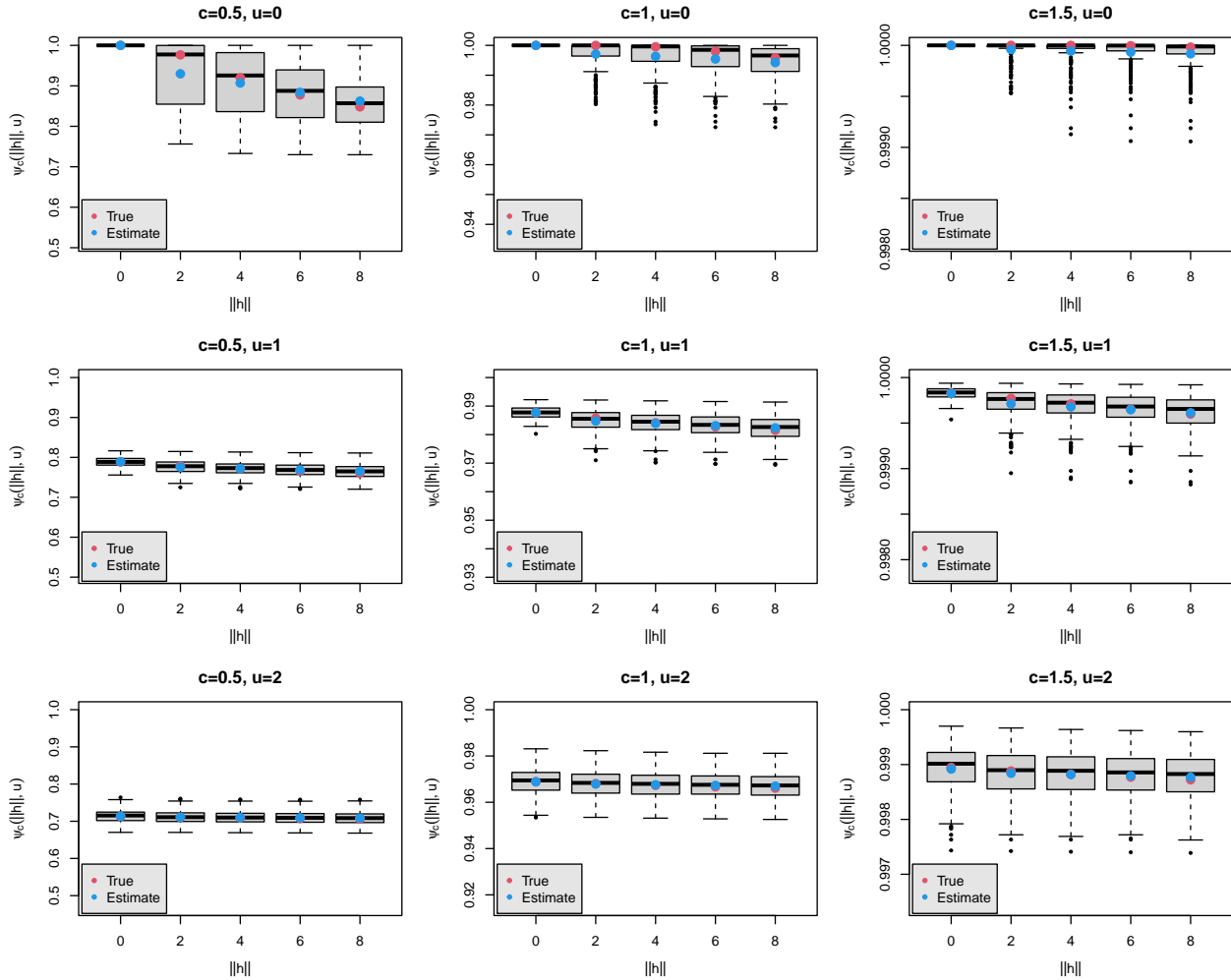
Figure 3. Estimates of the probability of agreement as a function of $\|\boldsymbol{h}\|$, u and c for a spatiotemporal Gaussian process with a positive linear trend and a non-separable Iacocesare covariance structure with fixed parameters given in Table 1, and for $N_s = 50$. Note differences in range limits of the y-axis among the nine panels.

Table 2

Initial values and parameter estimates for the linear temporal trend and the Iacocesare covariance structure.

	$\widehat{a}_0$	$\widehat{a}_1$	$\widehat{\alpha}_s$	$\widehat{\alpha}_t$	$\widehat{\beta}$	$\widehat{\phi}_s$	$\widehat{\phi}_t$	$\widehat{\sigma}^2$
Initial	0.50000	0.01000	1.00000	1.00000	2.00000	6.67616	1.00000	0.00100
Estimate	0.38453	0.00258	1.06297	0.95371	1.94867	6.53170	2.08316	0.00044

The computational efficiency of the composite likelihood method used in Sections 7 and 8 deserves attention. The computational implementations used in this article worked well for images of the order of size 50×50; we had to rasterize and downscale the original images used in Section 8 by two orders of magnitude before we could estimate the relevant parameters. Overcoming memory limitations to enable parameter estimation from much larger images (more than 10^6 pixels) will require innovative parallelizable algorithms.

We also note that the Monte Carlo simulations and the application were developed and illustrated using spatial processes (images) defined on regular grids. However, there is no

apparent reason that the proposed methods could not also be applied to irregularly spaced spatial data.

The normality assumption of the bivariate process is a challenging aspect because the available data consists of a single realization of a bivariate Gaussian process, as is typically assumed in spatial statistics. Based on this information, only the marginal normal distributions can be examined using standard statistical tools. However, this analysis does not provide useful information for assessing the bivariate normality of the vector $(X(\boldsymbol{s}), Y(\boldsymbol{s}))^{\top}$. There are extensions in the literature of the Kolmogorov-Smirnov test for the marginal distributions when the data are spatially correlated (see for instance, Pardo-Igúzquiza and Sowd, 2004). If the variance of
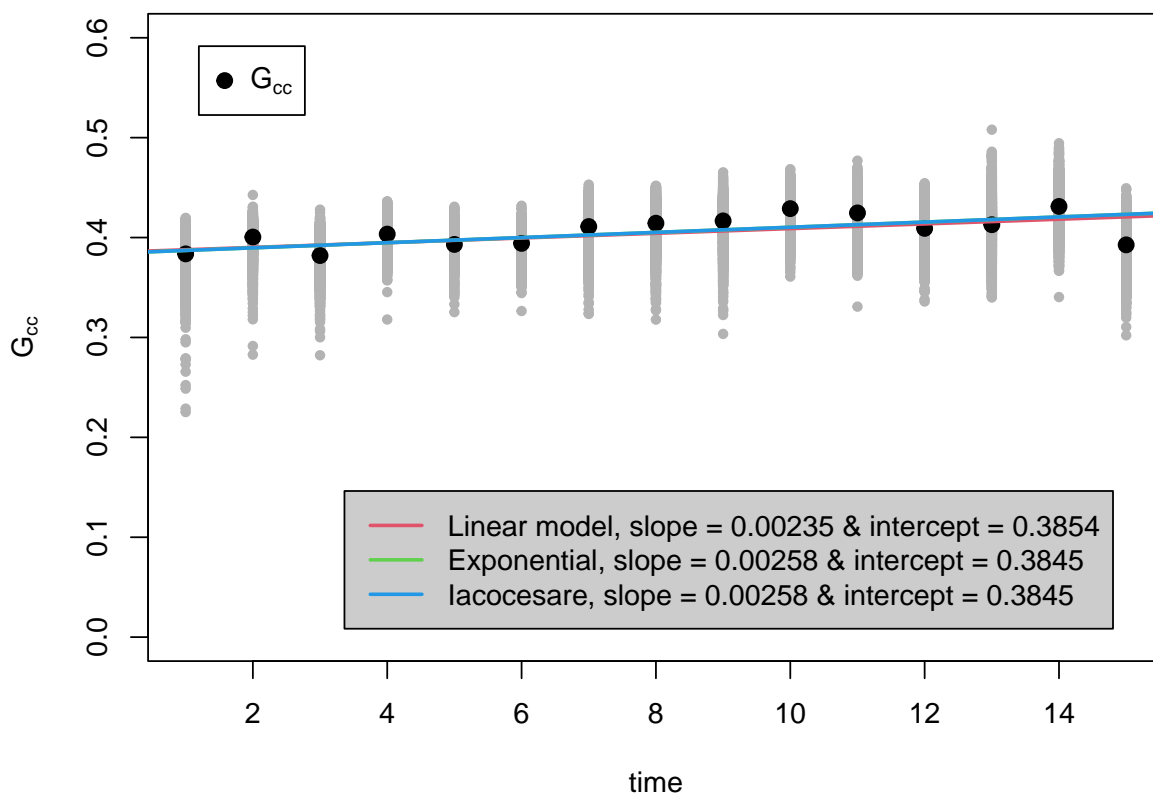
Figure 4. Green chromatic indices (G_{cc}) of each of the rasters in the 15 downscaled images (grey symbols) and the mean G_{cc} estimated for the entire region of interest (ROI; masked area in Figure 1D) by the PhenoCam network. Shown are the slopes and intercepts of the temporal trend in the G_{cc} by the PhenoCam network (Richardson et al., 2018) (red line), and using our spatiotemporal PA with separable (green), and non-separable (blue) spatial covariances.

$\widehat{\psi}_c$ is large, the normal approximation would not be suitable for values near the boundaries. An alternative nonparametric estimator could be implemented for ψ_c.

For practical applications of the PA we recommend comprehensive exploratory data analysis to determine the feasibility of using linear trends in time and to measure the sensitivity of the covariance structure under different parametric models.

Finally, a natural but unexplored extension of the present work is the definition of the PA for spatiotemporal marked point processes. Because the randomness in point processes is in the location (as well as in the marks) and repeated observations of a given point process will yield a new set of locations, it is of theoretical interest to estimate the PA of point patterns that evolve through time. Such estimation would have immediate applicability to ecological spatial datasets, which are predominantly samples of point processes, not rasters (e.g., Plant, 2019)

Acknowledgements

References

Anderson, T.W. (2003). An Introduction to Multivariate Statistical Analysis, 3rd Edition. Wiley, New York.

Barichivich, J., Briffa, K. R., Osborn, T. J., Melvin, T. M., and Caesar, J. (2012). Thermal growing season and timing of biospheric carbon uptake across the Northern Hemisphere. Global Biogeochemical Cycles, 26, GB4015.

Bevilacqua, M,, Faouzi, T., Furrer, R., and Porcu E. (2019). Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics. The Annals of Statistics, 47, 828–856.

Cressie, N. (1993) *Statistics for Spatial Data.* New York: Wiley.

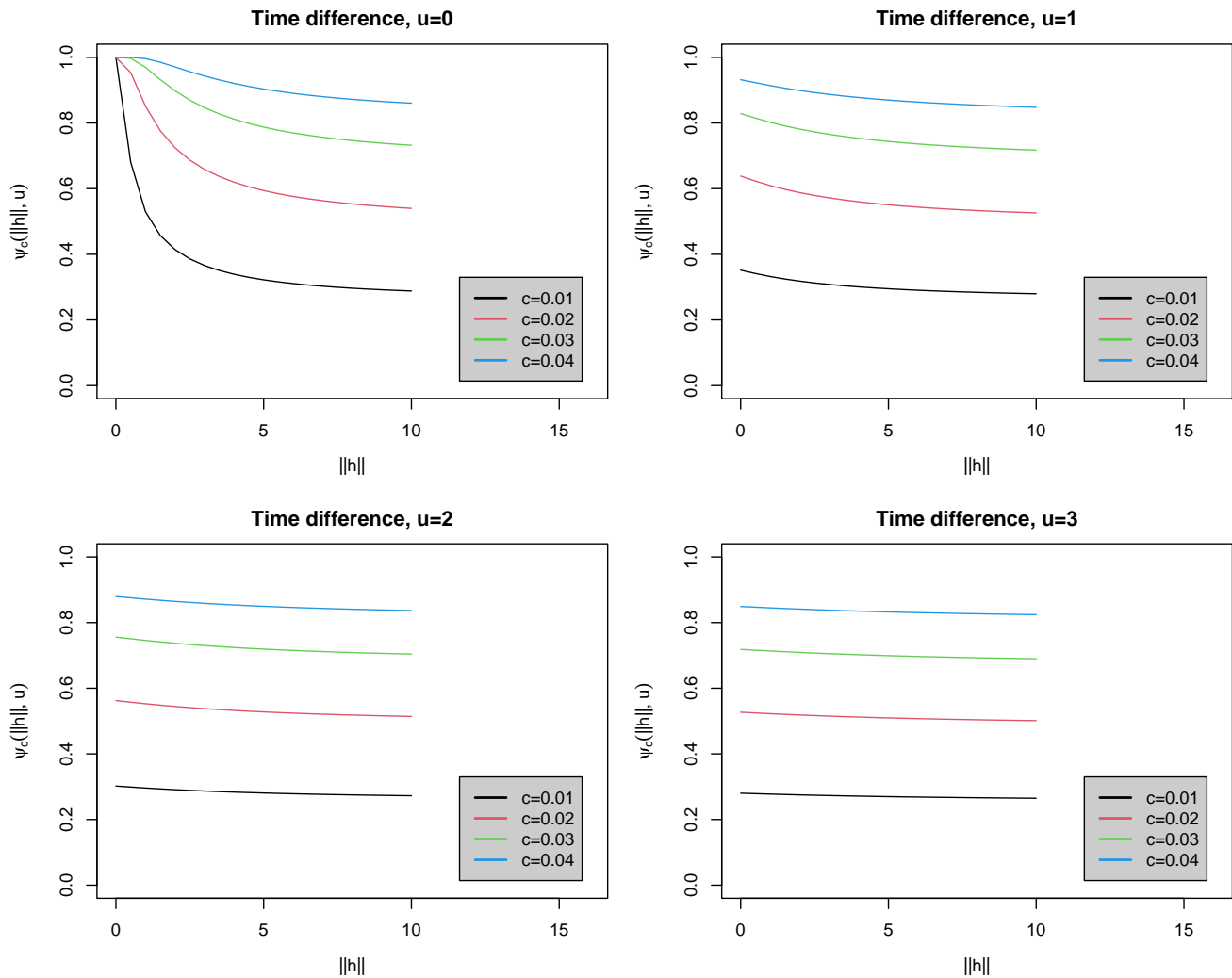**Figure 5.** Probability of agreement as a function of $\|\boldsymbol{h}\|$ for four time lags u with different maximum acceptable differences c.

Daley, D. J., Porcu, E., and Bevilacqua, M. (2015). Classes of compactly supported covariance functions for multivariate random fields. Stochastic Environmental Research and Risk Assessment, 29(4), 1249–1263

de Castro, M., and Galea, M. (2021). Bayesian inference for the pairwise probability of agreement using data from several measurement systems. Quality Engineering 33, 571-580.

Gneiting, T. (2002). Compactly supported correlation functions. Journal of Multivariate Analysis 83, 493–508.

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

De Iaco, S., Myers, D.E., Posa, D., (2002). Nonseparable space–time covariance models: some parametric families. Mathematical Geology 34, 23–42.

Gill, A. L., Gallinat, A. S., Sanders-DeMott, R., Rigden, A. J., Gianotti, D. J. S., Mantooth, J. A., and Templer, P.

H. (2015). Changes in autumn senescence in northern hemisphere deciduous trees: a meta-analysis of autumn phenology studies. Annals of Botany, 116, 875–888.

Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., et al. (2014). Net carbon uptake has increased through warming-induced changes in temperate forest phenology. Nature Climate Change 4, 598–604.

Leal, C., Galea, M., and Osorio, F. (2019). Assessment of local influence for the analysis of agreement. Biometrical Journal 61, 955–972.

Le Quéré, C., Raupach, M. R., Canadell, J. G., Marland, G., et al.(2009). Trends in the sources and sinks of carbon dioxide. Nature Geoscience, 2, 831–836.

Lin, L. (1989). A concordance correlation coefficient to evaluate reproducibility. Biometrics 45, 225–268.

Lin, L., Hedayat, A., Sinha, B., and Yang, M. (2002). Statistical methods in assessing agreement: models, issues, and tools. Journal of the American Statistical Association 97, 257–270.

Lin, L., Hedayat, A.S., and Wu, W. (2012). Statistical Tools for Measuring Agreement. Springer, New York.

Mardia, K. and Marshall, R. (1984). Maximum likelihood of models for residual covariance in spatial regression. Biometrika 71, 135–146.

Matérn, B. (1986). Spatial variation, second edition. Springer, New York.

Moon, M., Richardson, A. D., O'Keefe, J., and Friedl, M. A. (2022). Senescence in temperate broadleaf trees exhibits species-specific dependence on photoperiod versus thermal forcing. Agricultural and Forest Meteorology 322, 109026.

Pardo-Igúzquiza, E., Sowd, P. (2004). Normality tests for spatially correlated data. Mathematical Geology 36, 659–681.

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., and Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. Global Change Biology, 25, 1922–1940.

Plant, R. E. (2019). Spatial Data Analysis in Ecology and Agriculture Using R, 2nd edition. CRC Press, Florida.

Ponnet, J., Van Oirbeck, R., and Verdonck, T. (2021). Concordance probability for insurance pricing models. Risks 9, 178.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Richardson A. (2021). PhenoCam images and canopy phenology at the Harvard Forest EMS Tower since 2008. Harvard Forest Data Archive: HF158 (v.14).

Richardson, A. D., Jenkins, J. P., Braswell, B. H., Hollinger, D. Y., Ollinger, S. V., and M.-L. Smith. (2007). Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. Oecologia, 152, 323-334.

Richardson, A. D., Hufkens, K., Milliman, T., et al. (2018). Tracking vegetative phenology across diverse North American biomes using PhenoCam imagery. Scientific Data 5, 180028.

Seyednasrollah, B., Young, A. M., Hufkens, K., Milliman, T., Friedl, M. A., Frolking, S., and Richardson, A. D. (2019). Tracking vegetation phenology across diverse biomes using PhenoCam imagery: The PhenoCam Dataset v2.0. Scientific Data 6, 222.

Sherman, M. (2011). Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties. Wiley, United Kingdom.

Stein, M. L. (1999). Interpolation of Spatial Data: Some Theory for Kriging. Springer, New York.

Stevens, N. T., and Anderson-Cook, C. M. (2017). Comparing the reliability of related populations with the probability of agreement. Technometrics 59, 371–380.

Stevens, N. T., and Lu, L. (2020). Comparing Kaplan-Meier curves with the probability of agreement. Statistics in Medicine 39, 4621–4635.

Stevens, N. T., Lu, L., and Anderson-Cook, C. M., and Rigdon. S. E. (2020). Bayesian probability of agreement for comparing survival or reliability functions with parametric lifetime regression models. Quality Engineering 32, 312–332.

Stevens, N. T., Steiner, S. H., and MacKay, R. J. (2017). Assessing agreement between two measurement systems: An alternative to the limits of agreement approach. Statistical Methods in Medical Research 26, 2487–2504.

Stevens, N. T., Steiner, S. H., and MacKay, R. J. (2018). Comparing heteroscedastic measurement systems with the probability of agreement. Statistical Methods in Medical Research 27, 3420–3435.

Strandberg, J., de Luna, S., and Mateu, J. (2019). Prediction of spatial functional random processes: comparing functional and spatio-temporal kriging approaches. Stochastic Environmental Research and Risk Assessment 33, 1699–1719.

Taffé, P. (2022). Use of clinical tolerance limits for assessing agreement. Statistical Methods in Medical Research 32, 195–2006.

Yang, P., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., Xu, B., Shi, J., and Dickinson, R. (2013). The role of satellite remote sensing in climate change studies. Nature Climate Change 3, 875–883.

Appendix

Proof of Theorem 1

Without loss of generality, we assume that the Gaussian process has mean $\mathbf{0}$ and $h = \|\boldsymbol{h}\|$ in (5). First notice that

$$\psi_c(h) = \Phi\left(\frac{c}{\sigma_D(h)}\right) - \Phi\left(-\frac{c}{\sigma_D(h)}\right) = 2\Phi\left(\frac{c}{\sigma_D(h)}\right) - 1.$$

Now, let $h_1 > 0$ and $h_2 > 0$ be such that $h_1 \leqslant h_2$. By hypothesis, $\sigma_D(h_1) \leqslant \sigma_D(h_2)$, then $c/\sigma_D(h_2) \leqslant c/\sigma_D(h_1)$ because $c > 0$. Finally, since $\Phi(\cdot)$ is an increasing function, then

$$\Phi\left(\frac{c}{\sigma_D(h_2)}\right) \leqslant \Phi\left(\frac{c}{\sigma_D(h_1)}\right)$$
$$\iff \quad 2\Phi\left(\frac{c}{\sigma_D(h_2)}\right) - 1 \leqslant 2\Phi\left(\frac{c}{\sigma_D(h_1)}\right) - 1.$$

Therefore, $\psi_c(h_2) \leqslant \psi_c(h_1)$ for all $h_1 \leqslant h_2$. $\qquad\square$

Proof of Lemma 1 Let $\widehat{\sigma}_D(\boldsymbol{h}; \boldsymbol{\theta}) = g(\widehat{\boldsymbol{\theta}})$. Applying a Taylor expansion of order 1 for $g(\widehat{\boldsymbol{\theta}})$ around $\boldsymbol{\theta}$, we have that $\widehat{\sigma}_D(\boldsymbol{h}; \boldsymbol{\theta}) \approx g(\boldsymbol{\theta}) + \nabla g(\boldsymbol{\theta})^\top (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$. Then, $\mathrm{E}(\widehat{\sigma}_D(\boldsymbol{h}; \boldsymbol{\theta})) \approx g(\boldsymbol{\theta})$ and $\mathrm{var}(\widehat{\sigma}_D(\boldsymbol{h}; \boldsymbol{\theta})) \approx \nabla g(\boldsymbol{\theta})^\top \boldsymbol{V_\theta} \nabla g(\boldsymbol{\theta})$. Now, because $g(\boldsymbol{\theta}) = \sqrt{\sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})}$, then $\nabla g(\boldsymbol{\theta}) = \nabla \sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})/\{2g(\boldsymbol{\theta})\}$, where the i-th element of $\nabla \sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})$ is given by $\partial \sigma_D^2(\boldsymbol{h}; \boldsymbol{\theta})/\partial \theta_i = \partial C_X(\mathbf{0}; \boldsymbol{\theta})/\partial \theta_i + \partial C_Y(\mathbf{0}; \boldsymbol{\theta})/\partial \theta_i - 2\partial C_{XY}(\boldsymbol{h}; \boldsymbol{\theta})/\partial \theta_i$. $\qquad\square$

Proof of Theorem 4 Denote $\sigma_D = \sigma_D(\boldsymbol{h}, \boldsymbol{\theta})$, $\widehat{\sigma}_D = \sigma_D(\boldsymbol{h}, \widehat{\boldsymbol{\theta}})$, $\psi_c(\boldsymbol{h}; \mu_D, \boldsymbol{\theta}) = \psi_c(\boldsymbol{h}; \mu_D, \sigma_D)$, and $\psi_c(\boldsymbol{h}; \widehat{\mu}_D, \widehat{\boldsymbol{\theta}}) = \psi_c(\boldsymbol{h}; \widehat{\mu}_D, \widehat{\sigma}_D)$. Let $V_{\sigma_D}(\boldsymbol{h}) = \mathrm{var}(\sigma_D(\boldsymbol{h}; \widehat{\boldsymbol{\theta}}))$ given in (19). By Lemma 1, we have that $\sigma_D(\boldsymbol{h}; \widehat{\boldsymbol{\theta}})$ is consistent, and $\{\sigma_D(\boldsymbol{h}; \widehat{\boldsymbol{\theta}}) - \sigma_D(\boldsymbol{h}; \boldsymbol{\theta})\}/\sqrt{V_{\sigma_D}(\boldsymbol{h})} \xrightarrow{d} \mathcal{N}(0, 1)$. Now, using the delta method, it follows that

$$\psi_c(\boldsymbol{h}; \widehat{\mu}_D, \widehat{\sigma}_D) \approx \psi_c(\boldsymbol{h}; \mu_D, \sigma_D) + \tau_1(\widehat{\mu}_D - \mu_D) + \tau_2(\widehat{\sigma}_D - \sigma_D),$$
$$\tag{A.1}$$

where

$$\tau_1 = \left. \frac{\partial \psi_c(\boldsymbol{h}; \mu_D, \sigma_D)}{\partial \mu_D} \right|_{(\widehat{\mu}_D, \widehat{\sigma}_D)}$$

$$= \frac{1}{\sqrt{2\pi}\widehat{\sigma}_D} \left[\exp\left\{ -\frac{(c + \widehat{\mu}_D)^2}{2\widehat{\sigma}_D^2} \right\} - \exp\left\{ -\frac{(c - \widehat{\mu}_D)^2}{2\widehat{\sigma}_D^2} \right\} \right],$$

and

$$\tau_2 = \left. \frac{\partial \psi_c(\boldsymbol{h}; \mu_D, \sigma_D)}{\partial \sigma_D} \right|_{(\widehat{\mu}_D, \widehat{\sigma}_D)}$$

$$= -\frac{1}{\sqrt{2\pi}\widehat{\sigma}_D} \left[\exp\left\{ -\frac{(c + \widehat{\mu}_D)^2}{2\widehat{\sigma}_D^2} \right\} \frac{(c + \widehat{\mu}_D)}{\widehat{\sigma}_D} \right.$$

$$\left. + \exp\left\{ -\frac{(c - \widehat{\mu}_D)^2}{2\widehat{\sigma}_D^2} \right\} \frac{(c - \widehat{\mu}_D)}{\widehat{\sigma}_D} \right].$$

Applying expected value and variance in both sides of (A.1), the result follows.   □