Spatial Statistics 40 (2020) 100405

Contents lists available at ScienceDirect

Spatial Statistics

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

A spatial concordance correlation coefficient with an application to image analysis

Ronny Vallejos [a,*], Javier Pérez [a], Aaron M. Ellison [b], Andrew D. Richardson [c,d]

[a] Departamento de Matemática, Universidad Técnica Federico Santa María, Avenida España 1680, Valparaíso, Chile
[b] Harvard Forest, Harvard University, Petersham, MA, USA
[c] School of Informatics, Computing and Cyber Systems, Northern Arizona University, USA
[d] Center for Ecosystem Science and Society, Northern Arizona University, USA

ARTICLE INFO

Article history:
Received 13 May 2019
Received in revised form 29 November 2019
Accepted 19 December 2019
Available online 30 December 2019

Keywords:
Bivariate Wendland covariance function
Concordance
Correlation
Lin's coefficient
Spatial correlation function

ABSTRACT

In this work we define a spatial concordance coefficient for second-order stationary processes. This problem has been widely addressed in a non-spatial context, but here we consider a coefficient that for a fixed spatial lag allows one to compare two spatial sequences along a 45° line. The proposed coefficient was explored for the bivariate Matérn and Wendland covariance functions. The asymptotic normality of a sample version of the spatial concordance coefficient for an increasing domain sampling framework was established for the Wendland covariance function. To work with large digital images, we developed a local approach for estimating the concordance that uses local spatial models on non-overlapping windows. Monte Carlo simulations were used to gain additional insights into the asymptotic properties for finite sample sizes. As an illustrative example, we applied this methodology to two similar images of a deciduous forest canopy. The images were recorded with different cameras but similar fields-of-view and within minutes of each other. Our analysis showed that the local approach helped to explain a percentage of the non-spatial concordance and provided additional information about its decay as a function of the spatial lag.

* Corresponding author.
E-mail address: ronny.vallejos@usm.cl (R. Vallejos).

https://doi.org/10.1016/j.spasta.2019.100405
2211-6753/

1. Introduction

In recent decades, concordance correlation coefficients have been developed in a variety of different contexts. For instance, in assay or instrument validation processes, the reproducibility of the measurements among trials or laboratories is of interest. When a new instrument is developed, it may be relevant to evaluate whether its performance is concordant with other, existing ones, or its results accord with a "gold standard". There are also situations in which one is interested in comparing two methods without a designated gold standard or target values (Lin et al., 2002). In the literature, this latter type of concordance has been tackled from different perspectives (Barnhart et al., 2007). Cohen (1968) discussed this problem in the context of categorical data. Schall and Williams (1996) and Lin (2000) performed similar studies in the context of bioequivalence.

One way to approach the concordance problem for continuous measurements is to construct a scaled summary index that can take on values between -1 and 1, analogous to a correlation coefficient. Using this approach, Lin (1989) suggested a concordance correlation coefficient (CCC) that evaluates the agreement between two continuous variables by measuring their joint deviation from a 45° line through the origin. There have been some extensions of this CCC that use several measuring instruments and techniques to evaluate the agreement between two instruments; these efforts have led to interesting graphical tools (Hiriote and Chinchilli, 2011; Stevens et al., 2017). In the context of goodness of fit, Vonesh et al. (1996) proposed a modified Lin's CCC for choosing models that have a better agreement between the observed and the predicted values. Recently Stevens et al. (2017) and Chodhary and Nagaraja (2017) developed the probability of agreement, and Leal et al. (2019) studied the local influence of the CCC and the probability of agreement considering both first- and second-order measures under the case-weight perturbation scheme. Atkinson and Nevill (1997) critiqued the CCC because any correlation coefficient is highly dependent on the measurement range. In general, therefore, CCC is used only when measuring ranges are comparable or when methods are on the same scale.

In this paper, we suggest an approach to assessing the agreement between two continuous responses when the observations of both variables have been georeferenced in space. We define a spatial CCC (SCCC) as a generalization of Lin's (1989) coefficient that measures the agreement between two spatial variables. For a fixed lag, our SCCC shares the same properties as the original CCC. For an increasing domain sampling scheme, (i.e., for an asymptotic method for which the two-dimensional domain increases), we establish the asymptotic normality of the sample SCCC for a bivariate Gaussian process with a Wendland covariance function. To facilitate the computation of the spatial concordance coefficient when image sizes are large, we developed a local approach for estimating it that uses local spatial models on non-overlapping windows. This approach constitutes a new way of thinking about concordance that has not been considered previously, especially for large digital images. Our approach also captures the decay of the SCCC as a function of the norm of the spatial lag. Monte Carlo simulations and numerical experiments with real datasets accompany the exposition of the methodological aspects. An image-analysis example is worked in detail to illustrate the fitting of a local SCCCs. We conclude with a summary of the main findings and an outline of problems to be tackled in future research.

2. Preliminaries and notation

Assume that X and Y are two continuous random variables such that the joint distribution of X and Y has finite second moments with means μ_X and μ_Y, variances σ_X^2 and σ_Y^2, covariance σ_{XY} and correlation coefficient between X and Y

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

The mean squared deviation of $D = X - Y$ is

$$\text{MSD} = \epsilon^2 = \mathbb{E}[D^2] = \mathbb{E}[(X - Y)^2].$$

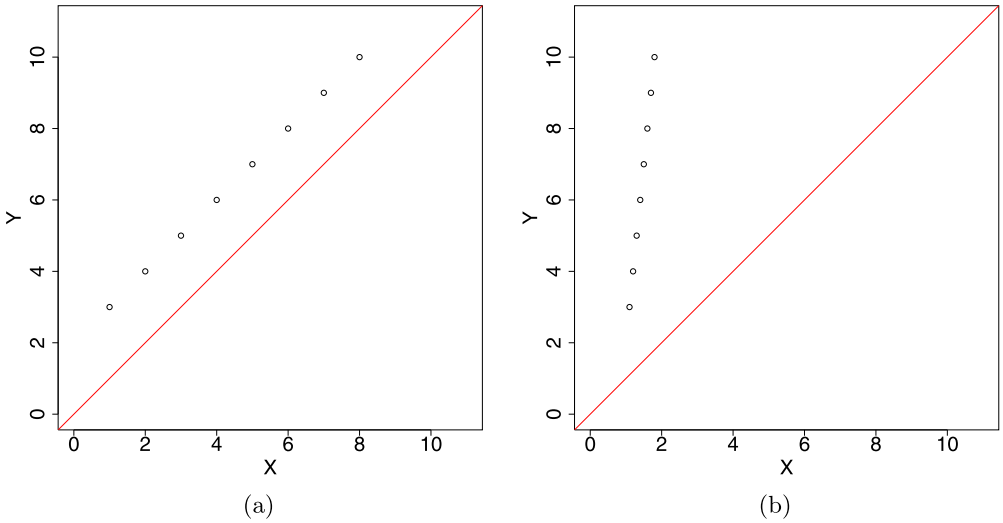


Fig. 1. Departure of two datasets defined on the plane from the 45° line intersecting the origin. (a) Perfect correlation with a shift effect results in relatively high concordance ($\widehat{\rho}_c = 0.7241$). (b) Perfect correlation but poor concordance ($\widehat{\rho}_c = 0.0340$).

It is straightforward to see that $\epsilon^2 = (\mu_X - \mu_Y)^2 + \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}$ and the sample counterpart satisfies $e^2 = (\overline{X} - \overline{Y})^2 + s_X^2 + s_Y^2 - 2s_{XY}$. Using this framework, Lin (1989) defined a CCC as:

$$\rho_c = 1 - \frac{\epsilon^2}{\epsilon^2|\rho = 0} = \frac{2\sigma_{XY}}{\sigma_X^2 + \sigma_Y^2 + (\mu_X - \mu_Y)^2}, \tag{1}$$

where $\epsilon^2|\rho = 0$ is the error, given that the correlation coefficient between X and Y is null. Lin (1989) pointed out that Eq. (1) is equal to one minus the ratio between the expected squared perpendicular deviation from the 45°-line intersecting the origin, and the expected squared perpendicular deviation from it when X and Y are uncorrelated.

The CCC satisfies the following properties:

1. $\rho_c = \alpha \cdot \rho$, where $\alpha = \frac{2}{w+1/w+v^2}$, $v = \frac{\mu_X - \mu_Y}{\sqrt{\sigma_X \sigma_Y}}$, and $w = \frac{\sigma_X}{\sigma_Y}$.
2. $|\rho_c| \leq |\rho| \leq 1$.
3. $\rho_c = 0$ if and only if $\rho = 0$.
4. $\rho_c = \rho$ if and only if $\sigma_Y = \sigma_X$ and $\mu_Y = \mu_X$.

It should be emphasized that ρ_c is always less than ρ because the first coefficient evaluates the degree to which pairs fall on the 45°-line intersecting the origin, but it is not the exact linear correlation between X and Y.

The sample estimate of ρ_c is given as:

$$\widehat{\rho}_c = \frac{2s_{XY}}{s_X^2 + s_Y^2 + (\overline{X} - \overline{Y})^2}.$$

To clarify the difference between the CCC and the correlation coefficient, two examples are illustrated in Fig. 1. Fig. 1(a) is a scatterplot for the case where there are different sample means but identical sample variances ($X = (1, 2, 3, \ldots, 8)$, $Y = (3, 4, 5, \ldots, 10)$); the correlation between X and Y coefficient is equal to 1 and $\widehat{\rho}_c = 0.7241$. Fig. 1(b) illustrates the case where both sample means and variances differ ($X = (1.1, 1.2, 1.3, \ldots, 1.8)$, $Y = (3, 4, 5, \ldots, 10)$). As in Fig. 1(a), there is a perfect correlation between X and Y ($\widehat{\rho} = 1$), but now there is much lower concordance X and Y ($\widehat{\rho}_c = 0.0340$).

Inference for this coefficient was addressed via Fisher's transformation. Lin (1989) proved that

$$\widehat{Z} = \frac{1}{2}\log\left(\frac{1+\widehat{\rho}_c}{1-\widehat{\rho}_c}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(Z, \sigma_Z^2), \text{ as } n \to \infty,$$

where

$$Z = \tanh^{-1}(\rho_c) = \frac{1}{2}\log\left(\frac{1+\rho_c}{1-\rho_c}\right)$$

and

$$\sigma_{\widehat{Z}}^2 = \frac{1}{n-2}\left[\frac{(1-\rho^2)\rho_c^2}{(1-\rho_c^2)\rho^2} + \frac{4v^2(1-\rho_c)\rho_c^3}{(1-\rho_c^2)^2\rho} + \frac{2v^4\rho_c^4}{(1-\rho_c^2)^2\rho^2}\right].$$

As a consequence of the asymptotic normality of the sample CCC, an approximate hypothesis testing problem of the form

$$\mathrm{H}_0 : \rho_c = \rho_0 \text{ versus } \mathrm{H}_1 : \rho_c \neq \rho_0$$

for a fixed ρ_0 can be constructed. Alternatively, an approximate confidence interval for ρ_c of the form

$$[\tanh(Z - z_{\alpha/2}\sigma_{\widehat{Z}}), \tanh(Z + z_{\alpha/2}\sigma_{\widehat{Z}})]$$

can be used, where $z_{\alpha/2}$ is the upper quantile of order $\alpha/2$ of the standard normal distribution. Applications and extensions of Lin's coefficient can be found in Lin et al. (2012), among others.

Although the notion of concordance and particularly Lin's coefficient have been generalized to different contexts, they also have been criticized. Atkinson and Nevill (1997) questioned Lin's coefficient for two reasons. First, these kind of methods are highly sensitive to sample heterogeneity. Second, its interpretation is problematic, especially when ρ_c is small. Lawrence and Chinchilli (1997) argued that most of the correlation coefficients largely depend on the analytical range, so that good concordance observed over a small range of measurements cannot be extrapolated to a similarly good concordance over a larger range of measurements. Three factors contribute to low values of ρ_c: small analytical range, imprecision, and inaccuracy (systematic bias) (Lawrence and Chinchilli, 1997). Because of these factors, study design plays a crucial role in the interpretation of Lin's coefficient. Measurement ranges should be compatible and they should be reported together with Lin's coefficient.

3. A spatial concordance coefficient and its properties

In this section we generalize Lin's coefficient for bivariate spatial processes. The main advantage of this extension is the fact that the new coefficient considers the existing spatial information of a georeferenced sample on the two-dimensional space. The way the coefficient takes into account the spatial association is through its dependence on a spatial lag $\boldsymbol{h} \in \mathbb{R}^2$, similarly to the covariance function or the variogram of a weakly stationary process. This extension preserves Lin's interpretation in the sense that for a particular spatial lag, the spatial concordance captures the departure from the 45° line passing through the origin. In particular, for isotropic processes the aim of this extension is to yield a plot of the spatial concordance coefficient versus the norm of the spatial lag. This will help to study the decay of the concordance as a function of the distance between the observations.

We start by extending Lin's CCC for bivariate second-order spatial processes for a general bivariate covariance function and a fixed lag in space.

Definition 1. Let $\boldsymbol{Z}(\boldsymbol{s}) = (X(\boldsymbol{s}), Y(\boldsymbol{s}))^\top$ 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

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

where

$$C_X(\boldsymbol{h}) = \text{Cov}[X(\boldsymbol{s}), X(\boldsymbol{s}+\boldsymbol{h})],$$
$$C_Y(\boldsymbol{h}) = \text{Cov}[Y(\boldsymbol{s}), Y(\boldsymbol{s}+\boldsymbol{h})],$$
$$C_{XY}(\boldsymbol{h}) = C_{YX}(\boldsymbol{h}) = \text{Cov}[X(\boldsymbol{s}), Y(\boldsymbol{s}+\boldsymbol{h})].$$

Then the SCCC is defined as

$$\rho^c(\boldsymbol{h}) = 1 - \frac{\mathbb{E}[(X(\boldsymbol{s}+\boldsymbol{h}) - Y(\boldsymbol{s}))^2]}{\mathbb{E}[(X(\boldsymbol{s}+\boldsymbol{h}) - Y(\boldsymbol{s}))^2 | C_{XY}(\boldsymbol{0}) = 0]} = \frac{2C_{XY}(\boldsymbol{h})}{C_X(\boldsymbol{0}) + C_Y(\boldsymbol{0}) + (\mu_X - \mu_Y)^2}. \tag{2}$$

Some straightforward properties of this SCCC are:

1. For $\eta = \frac{2\sqrt{C_X(\boldsymbol{0})C_Y(\boldsymbol{0})}}{C_X(\boldsymbol{0})+C_Y(\boldsymbol{0})+(\mu_X-\mu_Y)^2}$, and $\rho_{XY}(\boldsymbol{h}) = \frac{C_{XY}(\boldsymbol{h})}{\sqrt{C_X(\boldsymbol{0})C_Y(\boldsymbol{0})}}$, it follows that
$$\rho^c(\boldsymbol{h}) = \eta \cdot \rho_{YX}(\boldsymbol{h}). \tag{3}$$
2. $|\rho^c(\boldsymbol{h})| \leq |\rho_{XY}(\boldsymbol{h})| \leq 1$.
3. $\rho^c(\boldsymbol{h}) = 0$ iff $\rho_{XY}(\boldsymbol{h}) = 0$.
4. $\rho^c(\boldsymbol{h}) = \rho_{XY}(\boldsymbol{h})$ iff $\mu_X = \mu_Y$ and $C_X(\boldsymbol{0}) = C_Y(\boldsymbol{0})$.

In the sequel, the estimation of parametric covariance functions will be relevant, so we first define them for parametric correlations. If $R(\boldsymbol{h}, \boldsymbol{\phi})$ is a correlation function with parameter vector $\boldsymbol{\phi}$, then a covariance function is defined as:

$$C_X(\boldsymbol{h}) = \sigma_X^2 R(\boldsymbol{h}, \boldsymbol{\phi}_X),$$
$$C_Y(\boldsymbol{h}) = \sigma_Y^2 R(\boldsymbol{h}, \boldsymbol{\phi}_Y),$$
$$C_{XY}(\boldsymbol{h}) = \rho_{XY}\sigma_X\sigma_Y R(\boldsymbol{h}, \boldsymbol{\phi}_{XY}). \tag{4}$$

The SCCC then can be written as:

$$\rho^c(\boldsymbol{h}) = \frac{2\sigma_X\sigma_Y}{\sigma_X^2 + \sigma_Y^2}\rho_{XY} R(\boldsymbol{h}, \boldsymbol{\phi}_{XY}). \tag{5}$$

This means that the SCCC can be seen as a corrected version of the correlation between processes $X(\cdot)$ and $Y(\cdot)$. Because we expect that $\rho^c(\|\boldsymbol{h}\|)$ in general will decrease as a function of $\|\boldsymbol{h}\|$, it is possible to quantify the relationship between the spatial concordance and Lin's coefficient for a fixed spatial lag. Moreover, the value of $\|\boldsymbol{h}\|$ for which $\rho^c(\|\boldsymbol{h}\|)$ is negligible when it decreases as a function of $\|\boldsymbol{h}\|$ can be quantified similarly to the correlation function. In this case, the range of SCCC is the value of the norm of $\boldsymbol{h}$ for which $\rho^c(\|\boldsymbol{h}\|)$ is zero.

Using similar arguments as in properties 1–4, the SCCC in (5) could be derived for specific parametric bivariate correlation functions. Here we provide two examples.

1. For a bivariate random field with mean $(\mu, \mu)^\top$ and the Matérn covariance function
$$C_X(\boldsymbol{h}) = \sigma_X^2 M(\boldsymbol{h}, \nu_X, a_X), \tag{6}$$
$$C_Y(\boldsymbol{h}) = \sigma_Y^2 M(\boldsymbol{h}, \nu_Y, a_Y), \tag{7}$$
$$C_{XY}(\boldsymbol{h}) = \rho_{XY}\sigma_X\sigma_Y M(\boldsymbol{h}, \nu_{XY}, a_{XY}), \tag{8}$$
where $M(\boldsymbol{h}, \nu, a) = (a\|\boldsymbol{h}\|)^\nu K_\nu(a\|\boldsymbol{h}\|)$, $K_\nu(\cdot)$ is a modified Bessel function of the second kind, $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^2$, and ρ_{XY} is the co-located correlation coefficient between X

and Y (defined by Gneiting et al., 2010). It follows that

$$\begin{aligned}\rho^c(\boldsymbol{h}) &= \frac{2C_{XY}(\boldsymbol{h})}{C_X(\mathbf{0}) + C_Y(\mathbf{0}) + (\mu-\mu)^2} \\ &= \frac{2\rho_{XY}\sigma_X\sigma_Y M(\boldsymbol{h}, \nu_{XY}, a_{XY})}{\sigma_X^2 M(\mathbf{0}, \nu_X, a_X) + \sigma_Y^2 M(\mathbf{0}, \nu_Y, a_Y)} \\ &= \frac{2\rho_{XY}\sigma_X\sigma_Y M(\boldsymbol{h}, \nu_{XY}, a_{XY})}{\sigma_X^2 + \sigma_Y^2} \\ &= \frac{2\sigma_X\sigma_Y M(\boldsymbol{h}, \nu_{XY}, a_{XY})}{\sigma_X^2 + \sigma_Y^2} \cdot \rho_{XY} \\ &= \eta(\boldsymbol{h}) \cdot \rho_{XY},\end{aligned}$$

where $\eta(\boldsymbol{h}) = \frac{2\sigma_X\sigma_Y M(\boldsymbol{h}, \nu_{XY}, a_{XY})}{\sigma_X^2+\sigma_Y^2}$.

A special case of the Matérn covariance function is when $\nu_{XY} = p + 1/2$. Then

$$M(\boldsymbol{h}, \nu_{XY}, a_{XY}) = M(\boldsymbol{h}, p+1/2, a_{XY}) = \exp(-a_{XY}\|\boldsymbol{h}\|)\sum_{k=0}^{p}\frac{(p+k)!}{(2p)!}\binom{p}{k}(2a_{XY}\|\boldsymbol{h}\|)^{p-k},$$

and the SCCC is

$$\rho^c(\boldsymbol{h}) = \frac{2\sigma_{XY}}{\sigma_X^2+\sigma_Y^2}\exp(-a_{XY}\|\boldsymbol{h}\|)\sum_{k=0}^{p}\frac{(p+k)!}{(2p)!}\binom{p}{k}(2a_{XY}\|\boldsymbol{h}\|)^{p-k}.$$

By choosing $p = 0$ and $M(\boldsymbol{h}, 1/2, a_{XY}) = \exp(-a_{XY}\|\boldsymbol{h}\|)$, the SCCC can be written in its simplest form:

$$\rho^c(\boldsymbol{h}) = \frac{2\sigma_{XY}}{\sigma_X^2+\sigma_Y^2}\exp(-a_{XY}\|\boldsymbol{h}\|).$$

For illustrative purposes, consider $\sigma_X = 1$, $\sigma_Y = 2$, $\sigma_{XY} = 1.8$, $a_{XY} = 1/2$ and $\nu = \nu_{XY} = \{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}\}$. Then for $\|\boldsymbol{h}\| \in \{0, 1, \ldots, 15\}$, the SCCCs for different values of the smoothing parameter ν and using the Matérn covariance function are illustrated in Fig. 2. The curves of the SCCC decay more rapidly to zero as ν increases.

2. For a bivariate Wendland–Gneiting covariance function (Daley et al., 2015) of the form

$$\begin{aligned}C_X(\boldsymbol{h}) &= \sigma_X^2 W(\boldsymbol{h}, b_X, \gamma_X), \\ C_Y(\boldsymbol{h}) &= \sigma_Y^2 W(\boldsymbol{h}, b_Y, \gamma_Y), \\ C_{XY}(\boldsymbol{h}) &= \rho_{XY}\sigma_X\sigma_Y W(\boldsymbol{h}, b_{XY}, \gamma_{XY}),\end{aligned} \tag{9}$$

where

$$W(\boldsymbol{h}, b, \gamma) = b^{\nu+2k+1}B(\nu+2k+1, \gamma+1)\tilde{\psi}_{\nu+\gamma+1,k}\left(\frac{\|\boldsymbol{h}\|}{b}\right),$$

$B(\cdot,\cdot)$ is the beta function, and $\tilde{\psi}_{v,k}$ is defined for $k \geq 1$ as

$$\tilde{\psi}_{v,k}(t) = \int_t^1 \frac{u(u^2-t^2)^{k-1}(1-u)_+^v}{B(2k, v+1)}du, \quad 0 \leq t \leq 1,$$

where $(x)_+ = \mathbf{1}_{(x\geq 0)}$ (Gneiting, 2002), the SCCC is

$$\rho^c(\boldsymbol{h}) = \frac{2\rho_{XY}\sigma_X\sigma_Y W(\boldsymbol{h}, b_{XY}, \gamma_{XY})}{\sigma_X^2+\sigma_Y^2+(\mu_X-\mu_Y)^2}, \quad \boldsymbol{h} \in \mathbb{R}^2.$$

In particular, considering $W(\boldsymbol{h}, b_{XY}, \gamma_{XY}) = p_k(\|\boldsymbol{h}\|)(1-\|\boldsymbol{h}\|/b_{XY})_+^l$, where $k = 1$, $l = v+\gamma+1$, $\gamma = 0$ and $b_{XY} > 0$,

$$\rho^c(\boldsymbol{h}) = \frac{2\rho_{XY}\sigma_X\sigma_Y\,(1+l\|\boldsymbol{h}\|/b_{XY})\,(1-\|\boldsymbol{h}\|b_{XY})_+^l}{\sigma_X^2+\sigma_Y^2+(\mu_X-\mu_Y)^2}. \tag{10}$$

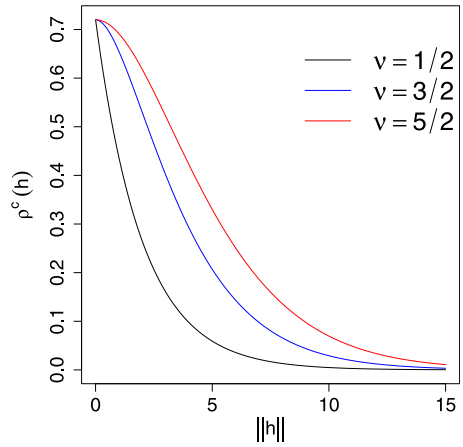


Fig. 2. $\rho^c(\boldsymbol{h})$ versus $\|\boldsymbol{h}\|$ for the Matérn covariance function and different values of the smoothness parameter ν.

4. Inference

In the previous section we showed that for several covariance structures, the spatial concordance correlation coefficient defined in Eq. (2) can be written as a product of the correlation function and a constant. Thus, we can consider plug-in estimators for the correlation function and the constant.

Let $\boldsymbol{Z}(\boldsymbol{s}) = (X(\boldsymbol{s}), Y(\boldsymbol{s}))^\top$, $\boldsymbol{s} \in D \subset \mathbb{R}^2$ be a Gaussian process with mean $\boldsymbol{\mu} = (\mu_X, \mu_Y)^\top$ and covariance function $\boldsymbol{C}(\boldsymbol{h})$, $\boldsymbol{s}, \boldsymbol{h} \in \mathbb{R}^2$. Then a sample (plug-in) estimate of the SCCC index (3) is

$$\widehat{\rho}^c(\boldsymbol{h}) = \widehat{\eta} \cdot \widehat{\rho}_{XY}(\boldsymbol{h}), \tag{11}$$

where $\widehat{\eta} = ((\widehat{a} + 1/\widehat{a} + \widehat{b}^2)/2)^{-1}$, $\widehat{a} = \left(\dfrac{\widehat{C}_{XX}(\mathbf{0})}{\widehat{C}_{YY}(\mathbf{0})}\right)^{1/2}$, $\widehat{b} = \dfrac{\widehat{\mu}_X - \widehat{\mu}_Y}{(\widehat{C}_{XX}(\mathbf{0})\widehat{C}_{YY}(\mathbf{0}))^{1/4}}$, and $\widehat{\mu}_X$, $\widehat{\mu}_Y$, $\widehat{C}_{XX}(\mathbf{0})$, and $\widehat{C}_{YY}(\mathbf{0})$ are the maximum likelihood (ML) estimates of μ_X, μ_Y, $C_{XX}(\mathbf{0})$, and $C_{YY}(\mathbf{0})$, respectively.

The asymptotic properties of an estimator like equation (11) have been studied in the literature for specific cases. Bevilaqua et al. (2015) studied the asymptotic properties of the ML estimator for a separable Matérn covariance model. They used a result provided by Mardia and Marshall (1984) in an increasing domain sampling framework: a process $\boldsymbol{Z}(\boldsymbol{s})$, with points $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ located in a rectangle $D_n \subset \Delta\mathbb{Z}^2$, for $0 < \Delta < \infty$, such that D_n satisfies the increasing condition $D_n \subset D_{n+1}$, for all n.

Using this theorem and the delta method, we can establish the following result for the Wendland–Gneiting model:

Theorem 1. *Let $\boldsymbol{Z}(\boldsymbol{s})$, $\boldsymbol{s} \in D \subset \mathbb{R}^2$ be a bivariate Gaussian spatial process with mean $\mathbf{0}$ and covariance function given by*

$$\boldsymbol{C}_X(\boldsymbol{h}) = \sigma_X^2 \left(1 + (\nu+1)\frac{\|\boldsymbol{h}\|}{b_X}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_X}\right)_+^{\nu+1},$$

$$\boldsymbol{C}_Y(\boldsymbol{h}) = \sigma_Y^2 \left(1 + (\nu+1)\frac{\|\boldsymbol{h}\|}{b_Y}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_Y}\right)_+^{\nu+1},$$

$$\boldsymbol{C}_{XY}(\boldsymbol{h}) = C_{YX}(\boldsymbol{h}) = \rho_{XY}\sigma_X\sigma_Y \left(1 + (\nu+1)\frac{\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1},$$

for $\nu > 0$ fixed. Define $\boldsymbol{\theta} = (\sigma_X^2, \sigma_Y^2, \rho_{XY}, b_{XY})^\top$ and denote the ML estimator of $\boldsymbol{\theta}$ as $\widehat{\boldsymbol{\theta}}_n$. Then,

$$\left(\nabla g(\boldsymbol{\theta})^\top \boldsymbol{F}_n(\boldsymbol{\theta})^{-1} \nabla g(\boldsymbol{\theta})\right)^{-1/2} (g(\widehat{\boldsymbol{\theta}}_n) - g(\boldsymbol{\theta})) \xrightarrow{D} \mathcal{N}(0,1), \ \text{as } n \to \infty,$$

in an increasing domain sense, where

$$g(\boldsymbol{\theta}) = \frac{2\rho_{XY}\sigma_X\sigma_Y \left(1 + (\nu+1)\frac{\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1}}{\sigma_X^2 + \sigma_Y^2},$$

$\boldsymbol{F}_n(\boldsymbol{\theta})^{-1}$ is the Fisher information matrix with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\theta}_n$,

$$\nabla g(\boldsymbol{\theta}) = \begin{pmatrix} \dfrac{\sigma_Y \rho_{XY}(\sigma_Y^2 - \sigma_X^2)\left(1 + (\nu+1)\frac{\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1}}{\sigma_X(\sigma_X^2 + \sigma_Y^2)^2} \\ \dfrac{\sigma_X \rho_{XY}(\sigma_X^2 - \sigma_Y^2)\left(1 + (\nu+1)\frac{\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1}}{\sigma_Y(\sigma_X^2 + \sigma_Y^2)^2} \\ \dfrac{2\sigma_X\sigma_Y \left(1 + (\nu+1)\frac{\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1}}{\sigma_X^2 + \sigma_Y^2} \\ \dfrac{2\sigma_X\sigma_Y\rho_{XY} f(b_{XY})}{\sigma_X^2 + \sigma_Y^2} \end{pmatrix},$$

and $f(b_{XY}) = \left(-\frac{(\nu+1)\|\boldsymbol{h}\|}{b_{XY}^2}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1} + \left(1 + \frac{(\nu+1)\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu} \frac{(\nu+1)\|\boldsymbol{h}\|}{b_{XY}^2}.$

Proof. See Appendix. □

5. A local approach

When the sizes of the images for which correspondence is to be assessed are large, it can be difficult to find a single model fitting reasonably well to the entire image(s). This has been investigated in the literature for autoregressive processes defined on the plane in the context of image restoration and segmentation (for examples, see Bustos et al., 2009).

Here we describe a local approach for a bivariate process of the form $\boldsymbol{Z}(\boldsymbol{s})$, $\boldsymbol{s} \in D \subset \mathbb{R}^2$, where the observations are located over a rectangular grid of size $n \times m$. The extension to an $l \in \mathbb{N}$-variate process is natural when $l > 2$. In this framework, we assume that the whole domain D can be divided into p sub-windows D_i, such that $\cup_{i=1}^n D_i = D$, for $i = 1, \ldots, p$. Then we define p processes of the form $\boldsymbol{Z}_i(\boldsymbol{s}) = (\boldsymbol{X}_i(\boldsymbol{s}), \boldsymbol{Y}_i(\boldsymbol{s}))^\top$, $\boldsymbol{s} \in D_i$, where each process has a covariance function given by

$$\begin{aligned} C_{X,i}(\boldsymbol{h}) &= \sigma_{X,i}^2 R(\boldsymbol{h}, \boldsymbol{\phi}_{X,i}), \\ C_{Y,i}(\boldsymbol{h}) &= \sigma_{Y,i}^2 R(\boldsymbol{h}, \boldsymbol{\phi}_{Y,i}), \\ C_{XY,i}(\boldsymbol{h}) &= \rho_{XY,i}\sigma_{X,i}\sigma_{Y,i} R(\boldsymbol{h}, \boldsymbol{\phi}_{XY,i}), \quad i = 1, \ldots, p, \end{aligned} \tag{12}$$

where $R(\boldsymbol{h}, \boldsymbol{\phi})$ is a correlation function with parameter vector $\boldsymbol{\phi}$, and $\boldsymbol{Z}_i(\cdot)$ and $\boldsymbol{Z}_j(\cdot)$ are supposed to be independent (Ojeda et al., 2010). Then for each local process $\boldsymbol{Z}_i(\cdot)$ we define the local SCCC $\rho_i^c(\cdot)$ using the theory developed in Section 3:

$$\rho_i^c(\boldsymbol{h}) = \frac{2\sigma_{X,i}\sigma_{Y,i}}{\sigma_{X,i}^2 + \sigma_{Y,i}^2} \rho_{XY,i} R(\boldsymbol{h}, \boldsymbol{\phi}_{XY,i}). \tag{13}$$

Based on the local coefficients $\rho_i^c(\cdot)$, we suggest two global SCCCs. The first one is the average of the p local coefficients:

$$\rho_1(\boldsymbol{h}) = \frac{1}{p}\sum_{i=1}^{p}\rho_i^c(\boldsymbol{h}). \tag{14}$$

The second one considers the average of each parameter in the correlation function such that the global coefficient is

$$\rho_2(\boldsymbol{h}) = \frac{2\overline{\sigma}_X\overline{\sigma}_Y}{\overline{\sigma}_X^2+\overline{\sigma}_X^2}\overline{\rho}_{XY}R(\boldsymbol{h},\overline{\boldsymbol{\phi}}_{XY}), \tag{15}$$

$\overline{\sigma}_X = \frac{1}{p}\sum_{i=1}^{p}\sigma_{X,i}$, and similarly for $\overline{\sigma}_Y$, $\overline{\rho}_{XY}$, and $R(\boldsymbol{h},\overline{\boldsymbol{\phi}}_{XY})$. As a result, we have two global coefficients of spatial concordance depending on averages: the first one is the average of the local coefficients and the second one is a plug-in of the parameter averages.

When process $\boldsymbol{Z}(\boldsymbol{s})$ have been observed in the sites $\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n$ and all the local coefficients have been computed, the sample versions of $\rho_1(\cdot)$ and $\rho_2(\cdot)$ are

$$\widehat{\rho}_1(\boldsymbol{h}) = \frac{1}{p}\sum_{i=1}^{p}\widehat{\rho}_i^c(\boldsymbol{h}),$$

$$\widehat{\rho}_2(\boldsymbol{h}) = \frac{2\widehat{\overline{\sigma}}_X\widehat{\overline{\sigma}}_Y}{\widehat{\overline{\sigma}}_X^2+\widehat{\overline{\sigma}}_Y^2}\widehat{\overline{\rho}}_{XY}R(\boldsymbol{h},\widehat{\overline{\boldsymbol{\phi}}}_{XY}),$$

where $\widehat{\overline{\sigma}}_X$, $\widehat{\overline{\sigma}}_Y$, $\widehat{\overline{\rho}}_{XY}$, and $R(\boldsymbol{h},\widehat{\overline{\boldsymbol{\phi}}}_{XY})$ are the means of the ML estimators of the parameters defined in Eq. (15).

Considering an increasing domain sampling scheme, the asymptotic normality of $\widehat{\rho}_1(\boldsymbol{h})$ is straightforward. Indeed, let $\boldsymbol{Z}_i(\boldsymbol{s}) = (X_i(\boldsymbol{s}), Y_i(\boldsymbol{s}))^\top$, $\boldsymbol{s}\in D_i$, be a bivariate process with correlation structure given by (12). Define the parameter vector $\boldsymbol{\theta}^i = (\rho_{XY,i}, \sigma_{X,i}, \sigma_{Y,i}, \sigma_{XY,i}, \boldsymbol{\phi}_{XY,i})^\top$ associated with $\boldsymbol{Z}_i(\boldsymbol{s})$ and denote the ML estimator of $\boldsymbol{\theta}^i$ as $\widehat{\boldsymbol{\theta}}_n^i$. If the covariance satisfies the Mardia and Marshall (1984) conditions, then

$$\widehat{\boldsymbol{\theta}}_n^{\,i} \xrightarrow{D} \mathcal{N}(\boldsymbol{\theta}^i, \boldsymbol{F}_n^i(\boldsymbol{\theta}^i)^{-1}),$$

where $\boldsymbol{F}_n^i(\boldsymbol{\theta}^i)$ is the covariance matrix of $\widehat{\boldsymbol{\theta}}_n^{\,i}$. Then for $g(\boldsymbol{\theta}) = \rho_i^c(\boldsymbol{h})$,

$$\left(\nabla g(\boldsymbol{\theta}^i)^\top \boldsymbol{F}_n^i(\boldsymbol{\theta}^i)^{-1}\nabla g(\boldsymbol{\theta}^i)\right)^{-1/2}(g(\widehat{\boldsymbol{\theta}}^i_n) - g(\boldsymbol{\theta}^i)) \xrightarrow{D} \mathcal{N}(\boldsymbol{0},1).$$

Assuming that $\boldsymbol{Z}_i(\boldsymbol{s})$ and $\boldsymbol{Z}_j(\boldsymbol{s})$ are independent for all $i\neq j$,

$$\widehat{\rho}_1(\boldsymbol{h}) = \frac{1}{p}\sum_{i=1}^{p}\widehat{\rho}_i^c(\boldsymbol{h}) \xrightarrow{D} \mathcal{N}\left(\rho_i^c(\boldsymbol{h}), \frac{1}{p^2}\sum_{i=1}^{p}\nabla g(\boldsymbol{\theta}^i)^\top \boldsymbol{F}_n^i(\boldsymbol{\theta}^i)^{-1}\nabla g(\boldsymbol{\theta}^i)\right).$$

6. Monte Carlo simulations

We used Monte Carlo simulations to explore the properties of the SCCC, $\rho^c(\cdot)$, for finite samples sizes. The performance of the ML estimates were then analyzed with respect to the true values of the coefficient. We generated 500 replicates from a Gaussian random field sampled on a regular lattice of size 20×20 inside the region $[-\frac{3}{2},\frac{3}{2}]^2$. Each replicate was generated from a bivariate Gaussian random field with mean zero and Wendland–Gneiting covariance function given in Eq. (9). In each case, we estimated the parameters of the covariance function using ML and used them to compute the SCCC given in Eq. (10). Four set of parameters were considered to assess the performance of the SCCC when these parameters vary:

1. Case 1: $\sigma_X=\sigma_Y=1$, $\rho_{XY}=-0.15$, $b_X=0.5$, $b_Y=0.4$, $b_{XY}=0.35$.
2. Case 2: $\sigma_X=\sigma_Y=1$, $\rho_{XY}=0.25$, $b_X=1.2$, $b_Y=0.9$, $b_{XY}=1$.

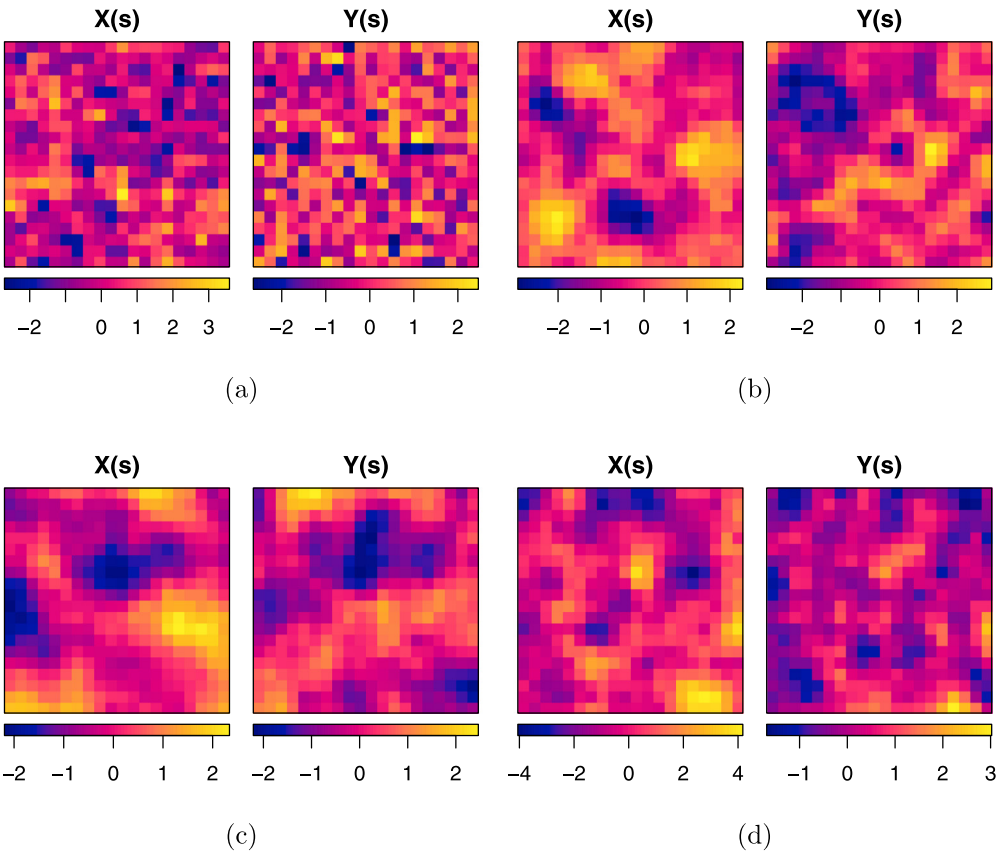


Fig. 3. Realization of a Gaussian random field with bivariate Wendland–Gneiting correlation function. (a) Case 1; (b) Case 2; (c) Case 3; (d) Case 4.

3. Case 3: $\sigma_X = \sigma_Y = 1$, $\rho_{XY} = 0.3$, $b_X = 1.8$, $b_Y = 1.4$, $b_{XY} = 1.5$.
4. Case 4: $\sigma_X = 2$, $\sigma_Y = 0.5$, $\rho_{XY} = 0.45$, $b_X = 1.1$, $b_Y = 1$, $b_{XY} = 0.9$.

Note that in all cases $\nu = 4$ and $k = 1$ and that Case 4 includes different marginal variances.

In Fig. 3 we show a realization of the random field for each case.

Table 1 summarizes the estimates of the Wendland–Gneiting model for the simulations. The estimates were obtained using ML estimation and the values reported are the average over 500 simulation runs. In each case the value of the standard error is in parenthesis under the estimation. In general we observed reasonable estimates of all parameters and small standard errors.

The ML estimates of the parameters of the Wendland–Gneiting covariance function had low bias and standard errors, and agreed with previously published results (e.g., Bevilaqua et al., 2019). Using these estimates, we computed the SCCC in each case for $0 < \|\boldsymbol{h}\| < 2$. This length for $\|\boldsymbol{h}\|$ was enough to observe the decay of the SCCC in each case even though the maximum distance between two points in the region was $3\sqrt{2}$ (Fig. 4).

The mean square errors of the estimates shown in (Table 1) were all less than $1.6485 \cdot 10^{-6}$, 0.0001, $6.5292 \cdot 10^{-5}$, and $9.3775 \cdot 10^{-6}$, respectively, for cases 1–4. $\rho^c(\boldsymbol{h})$ *versus* $\|\boldsymbol{h}\|$ and $\widehat{\rho}^c(\boldsymbol{h})$ *versus* $\|\boldsymbol{h}\|$ are plotted in Fig. 4; the true coefficient is drawn with a continuous line. In addition, for each value of $\|\boldsymbol{h}\|$, we included 90% confidence intervals based on quantiles 5 and 95 of the simulation values. The estimates of the SCCC were reasonably well-behaved but worsened when $\|\boldsymbol{h}\|$ was close

Table 1
Theoretical values, parameter estimates and standard errors (in parentheses) for the four cases considered in the simulation study in which we generated realizations of a random Gaussian random field with a bivariate Wendland–Gneiting covariance function.

Case 1	$\sigma_1^2 = 1$	$\sigma_2^2 = 1$	$\rho_{12} = -0.15$	$b_1 = 0.5$	$b_{12} = 0.35$	$b_2 = 0.4$
	0.9906 (0.0896)	0.9995 (0.0715)	−0.1493 (0.0507)	0.4994 (0.0369)	0.3707 (0.1772)	0.4511 (0.1451)
Case 2	$\sigma_1^2 = 1$	$\sigma_2^2 = 1$	$\rho_{12} = 0.25$	$b_1 = 1.2$	$b_{12} = 0.9$	$b_2 = 1$
	1.0039 (0.2034)	1.0011 (0.1706)	0.2439 (0.1274)	1.2137 (0.1876)	1.0711 (0.3884)	0.9528 (0.3172)
Case 3	$\sigma_1^2 = 1$	$\sigma_2^2 = 1$	$\rho_{12} = 0.3$	$b_1 = 1.8$	$b_{12} = 1.4$	$b_2 = 1.5$
	0.9901 (0.1744)	1.0003 (0.00922)	0.3092 (0.3892)	1.7973 (0.5574)	1.4829 (0.2836)	1.4811 (0.2483)
Case 4	$\sigma_1^2 = 2$	$\sigma_2^2 = 0.5$	$\rho_{12} = 0.45$	$b_1 = 1.1$	$b_{12} = 0.9$	$b_2 = 1$
	1.9998 (0.3810)	0.4960 (0.0798)	0.4379 (0.0955)	1.1073 (0.1641)	0.9347 (0.4215)	1.1606 (0.3298)

to zero, as is typical of lag-dependent spatial functions computed over a rectangular grid where the minimum distance between coordinates is fixed. We also ran Monte Carlo simulations using the bivariate Matérn covariance function; the results were similar. The estimate of ρ_{XY} was better for the Matérn case in terms of the mean square error. With either covariance function, however, the estimates of ρ_{XY} affected the estimates of the SCCC.

For the same region used in the previous Monte Carlo simulation, we computed the asymptotic variance of $\widehat{\rho}^c(\cdot)$. For $0 < \|\boldsymbol{h}\| < 2$, all variances were less than 0.006, and the largest discrepancies between cases 1–4 were seen near the origin.

For the local approach, we generated 100 replicates from a bivariate Gaussian random field sampled on a regular lattice of size 100×100 in the region $[0, 100]^2$, with mean equal to zero and Wendland–Gneiting covariance given in Eq. (9). We then split each process into 25 square sub-images of size 10×10, and estimated the parameters of the covariance function for each sub-image using ML. We also computed the local SCCCs given by Eqs. (14) and (15). Fig. 5 illustrates box-plots of the estimates of σ_X, σ_Y, ρ_{XY} and b_{XY} for the 25 sub-images for case 3, where the true parameters were $\nu = 4$, $k = 1$, $\sigma_X = \sigma_Y = 1$, $\rho_{XY} = 0.3$, $b_X = 1.8$, $b_Y = 1.4$, and $b_{XY} = 1.5$. Although the estimates did vary, the average values of all sub-images were reasonably close to the true values. Finally, we computed $\widehat{\rho}_1(\boldsymbol{h})$ and $\widehat{\rho}_2(\boldsymbol{h})$ given by Eqs. (14) and (15), and compared them with the global SCCC (Fig. 6)

To observe the performance of the estimates of SCCC for an increasing domain framework, we did another Monte Carlo simulation study. Five hundred simulation runs from a zero-mean Gaussian random field with a Wendland–Gneiting covariance function with parameters $\nu = 4$, $k = 1$, $\sigma_X = \sigma_Y = 1$, $\rho_{XY} = 0.3$, $b_X = 1.8$, $b_Y = 1.4$, and $b_{XY} = 1.5$, were generated. We ran the simulations for the following six regions: $[1, 4]^2$, $[1, 8]^2$, $[1, 12]^2$, $[1, 16]^2$, $[1, 20]^2$ and $[1, 24]^2$, with 16, 64, 144, 256, 400 and 576 points respectively. Then in each simulation run, the ML estimations of the parameters of the SCCC were computed together with 90% confidence intervals based on the 5^{th} and 95^{th} quantiles. Estimates improve as the domain increases, yielding thinner confidence intervals (Fig. 7).

To gain insight into the computational time required for computing $\widehat{\rho}^c(\cdot)$ for the covariance functions used in this work, we ran similar simulations with different window sizes. We ran 100 simulations for window sizes = 8×8, 12×12, 16×16, and 20×20. In each, $\widehat{\rho}^c(\cdot)$ was computed for the Matérn and Wendland–Gneiting covariance functions. All computations were done using an HP ProLiant DL380G9 server, equipped with a 2x Intel Xeon E5-2630 v3 2.40 GHz processor, 128 GB DDR4 2.133 Ghz RAM, and 512 GB SSD storage.

Time to run each simulation increased exponentially with window size (Fig. 8). Although the time required to compute the Wendland–Gneiting covariance function was always smaller than the time to compute the Matérn covariance function, for real images it is not feasible to compute

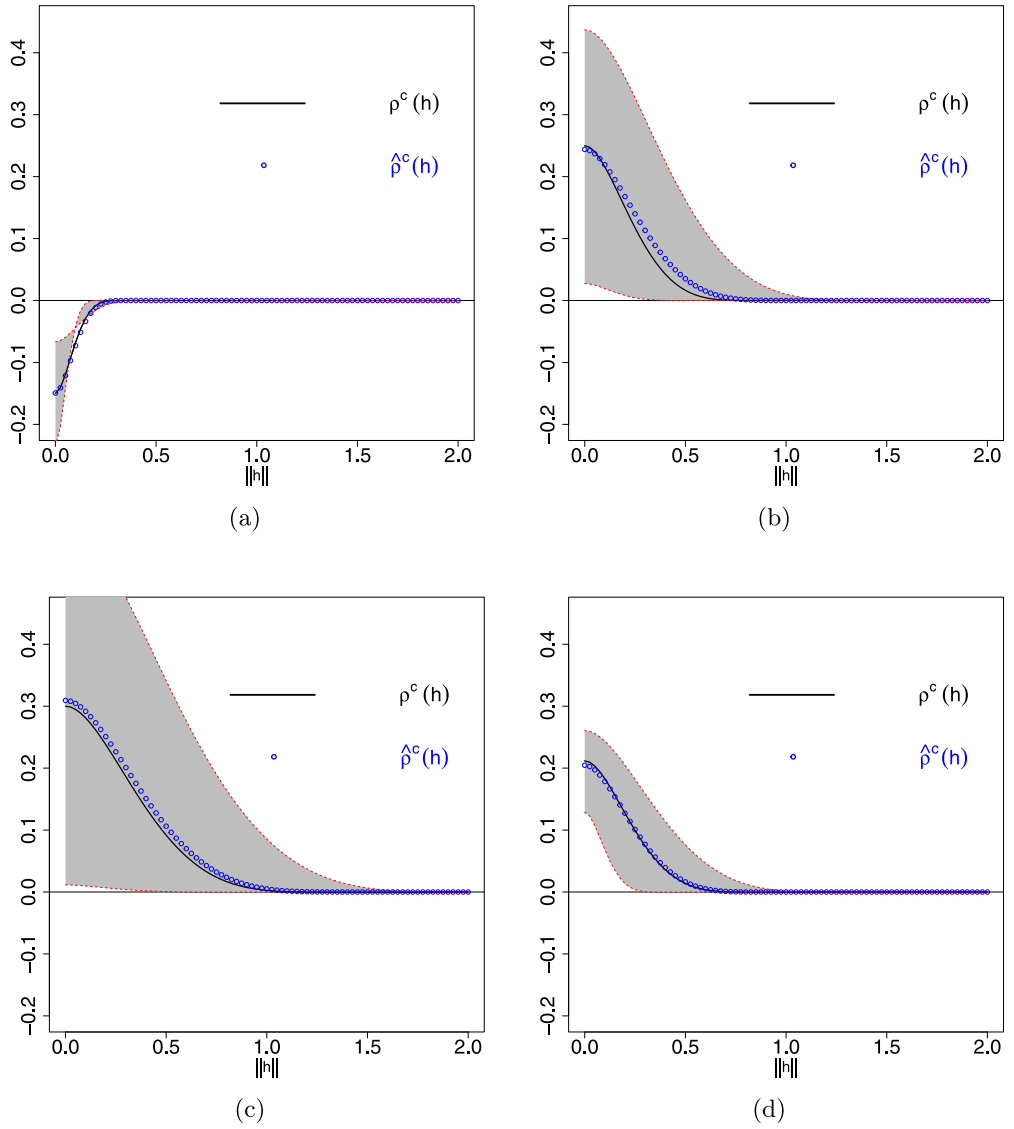


Fig. 4. Theoretical coefficients (solid lines) and estimates (circles) for the distinct sets of parameters,with a 90% confidence interval based on quantiles 5 and 95. (a) Case 1; (b) Case 2; (c) Case 3; (d) Case 4.

$\widehat{\rho}^{c}(\cdot)$, at least using an interpreted language such as R, which we used here. It is expected to find this kind of drawback for the Matérn covariance function because of the complexity of the model, the number of parameters, and the large amount of data. The compact support covariance functions (as in the Wendland model) reduces the computational burden (Wendland, 1995).

This result further supports the use of the local approach we presented in Section 5, but we plan to explore ways to optimize and accelerate the computation of $\widehat{\rho}^{c}(\cdot)$.

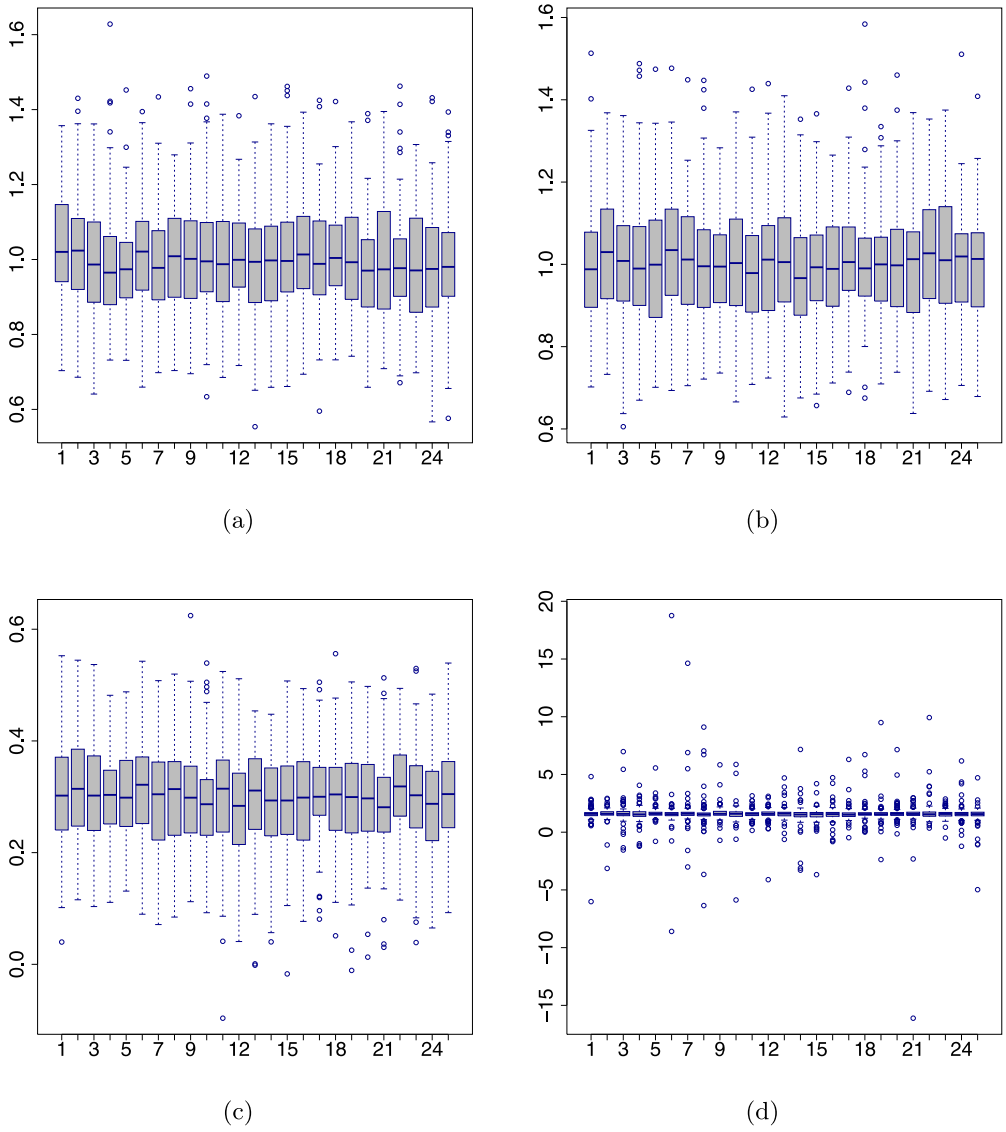


Fig. 5. Box-plot for the estimates of the covariance parameters in each of the 25 sub-images. (a) σ_X; (b) σ_Y; (c) ρ_{XY} and (d) b_{XY}.

7. An application

7.1. Motivation

Our application derives from ecology. In order to track the seasonality ("phenology") of vegetation in different ecosystems, digital cameras have been deployed to record high-frequency images of the canopy at hundreds of research sites around the world (Richardson, 2018). From each image, color-channel information (e.g., RGB [red-green-blue] values of each pixel) are extracted and converted to a suite of "vegetation indices" derived from linear or nonlinear transformations

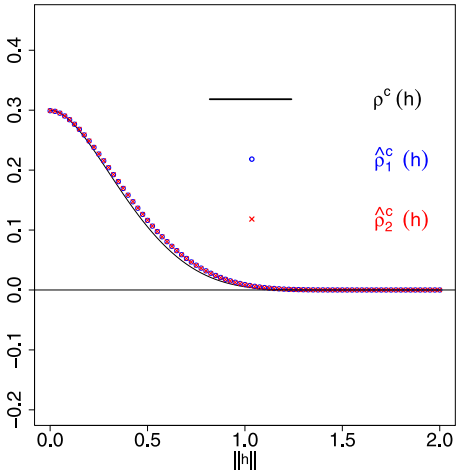


Fig. 6. Local SCCCs estimates (red and blue lines) and global theoretical SCCC (black) for a Gaussian spatial process with Wendland–Gneiting covariance function. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of the RGB or other color spaces (Sonnentag et al., 2012; Mizunuma et al., 2014; Toomey et al., 2015; Nguy-Robertson et al., 2016). These indices have been used to identify the timing of seasonal phenomena such as leaf-out, senescence, and abscission, and to monitor how these phenomena are changing in response to ongoing climatic change (Sonnentag et al., 2012). However, different cameras may render the same scene differently because of the specifics of the imaging sensor being used (e.g., CCD, CMOS) and researchers have used a wide range of different cameras because of considerations including trade-offs between cost and image quality. Changes in scene illumination (e.g., caused by time-of-day or cloud cover) also may impact the resulting image. Although previous research has shown that diurnal, seasonal, and weather-related changes in illumination can have large effects on estimates of average color (or color index) for the whole image or a region of interest (Sonnentag et al., 2012), spatial information has not been incorporated previously in these estimates.

7.2. Imagery

We focus here on comparing two jpeg images taken of the same scene on 20 October 2010 by two different cameras (Figs. 9(a) and 9(b)). These images were taken with, respectively, an outdoor StarDot NetCam XL 3MP camera with a 2048 $\times$ 1636-pixel CMOS sensor (Fig. 9(a)) and an outdoor Axis 223M camera with a 1600 $\times$ 1200-pixel CCD sensor (Fig. 9(b)). These images were selected from the image archive associated with an experiment, analyzed and reported on previously by Sonnentag et al. (2012), in which images, color time series, and phenological transition dates from eleven different cameras were compared. Although the two images we use here are of the same scene and were taken at the same time, they are not identical. For example, both cameras were pointing due north with an $\approx 20^\circ$ tilt angle, but image displacement occurred because the cameras were mounted at different positions on a fixed platform. The resolution and overall field-of-view also differed because of different sensor sizes and lens characteristics. Sonnentag et al. (2012) compared color information averaged across a small "region of interest" in the images. Here, we work with the entire images after correction for differences of field-of-view and displacement.

To account for differences in field-of-view and displacement, the two images were first manually cropped using tools in IrfanView (version 4.38; Skiljan 2014) to equivalent areas and aspect ratios. The resulting images had 2023 $\times$ 1444 pixels for the higher-resolution one taken with the StarDot

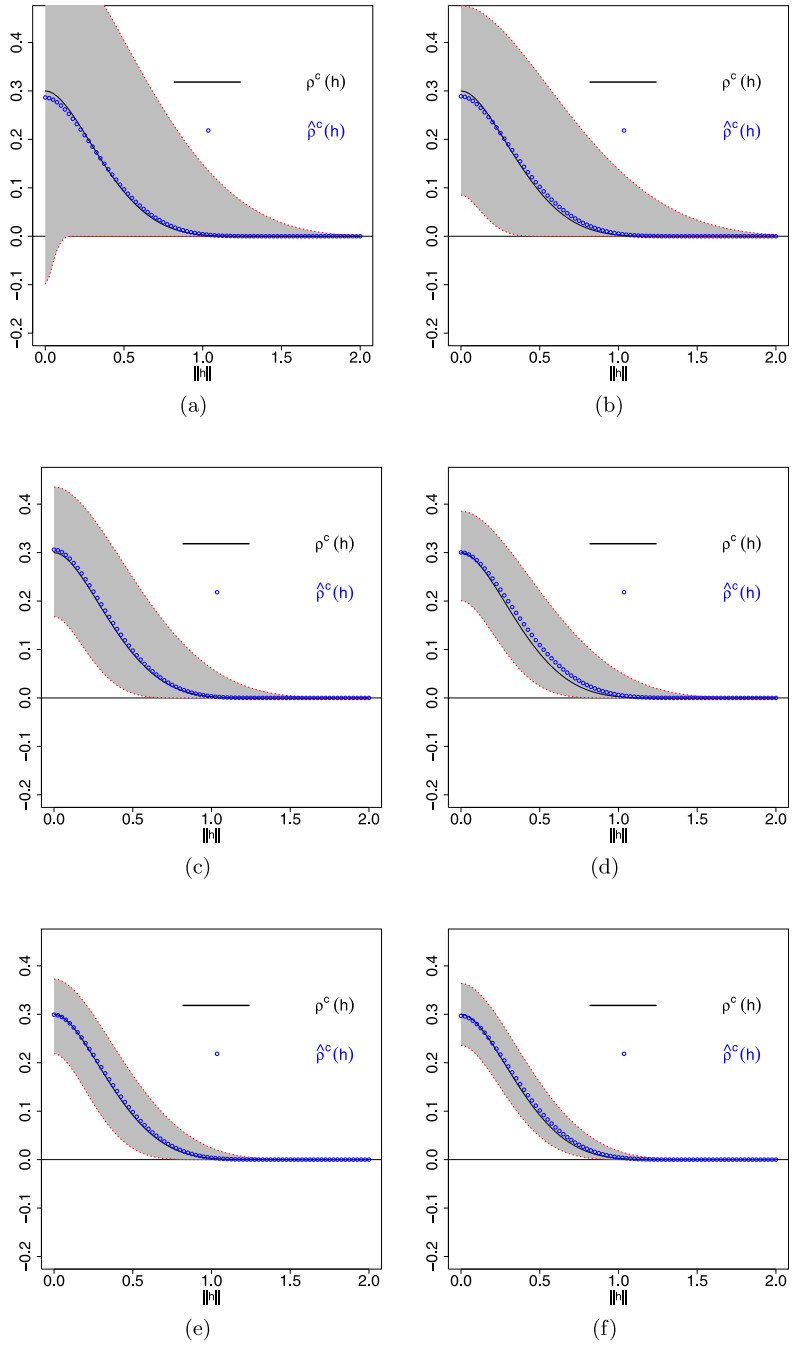


Fig. 7. Theoretical SCCC (black line) and estimates (blue circles) for the distinct regions, with a confidence interval of 90%. (a) 4 × 4 grid ; (b) 8 × 8 grid; (c) 12 × 12 grid; (d) 16 × 16 grid; (e) 20 × 20 grid ; (f) 24 × 24 grid.

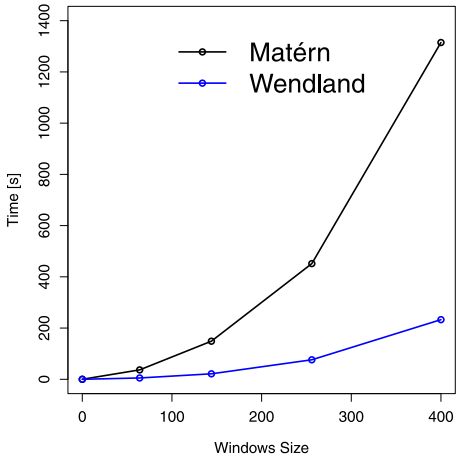


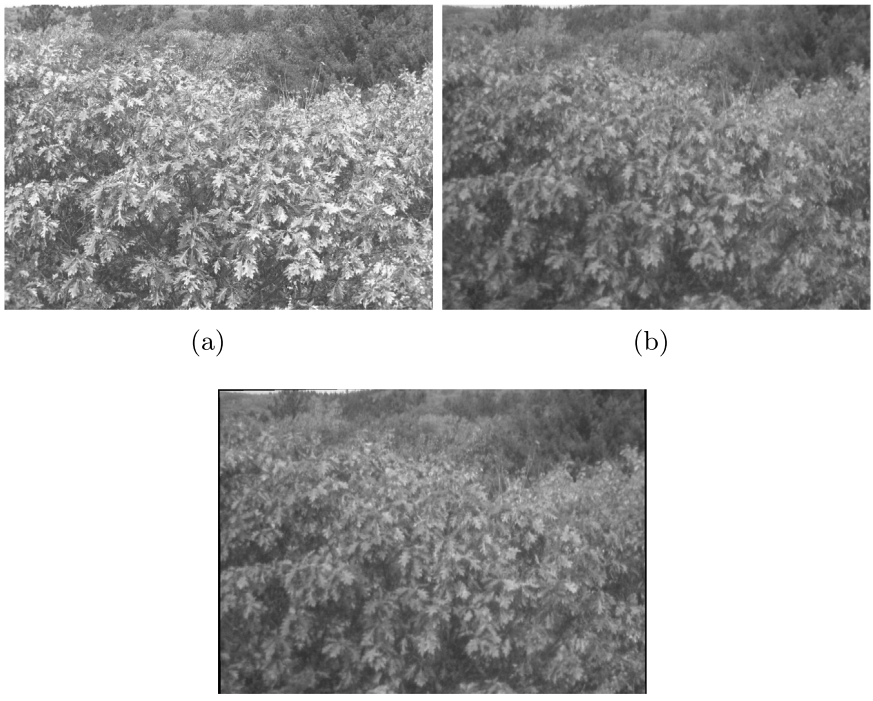
Fig. 8. Computational time in seconds to compute $\widehat{\rho}^{c}(\cdot)$ for the Matérn and Wendland–Gneiting covariance functions.

Fig. 9. Two images taken by adjacent cameras of the same site at Harvard Forest. (a): Image taken with an outdoor StarDot NetCam XL 3MP camera. (b): Image taken with an outdoor Axis 223M camera. The dominant tree species (foreground) is red oak (*Quercus rubra*), and there is some white pine (*Pinus strobus*) in the upper right corner.

camera and 1297 × 922 pixels for the lower-resolution one taken with the Axis camera. The higher-resolution image was then resized and down-sampled in IrfanView so that it had the same number of pixels as the lower-resolution image (Figs. 10(a) and 10(b)). These two images were loaded into the `R` software system (version 3.51; R Core Team, 2018) using the `load.image` function in the `imager` package (Urbanek, 2014) and transformed either to gray-scale using the `grayscale` function in the same package (Figs. 11(a) and 11(b)) or to green chromatic coordinates ($g_{cc} = \frac{G}{R+G+B}$), which normalizes for brightness (Gillespie et al., 1987) (Figs. 12(a) and 12(b)). For both the gray-scale and g_{cc} images, the lower-resolution image (Figs. 11(b) and 12(b), respectively) was then coordinate-registered to the higher-resolution image (Figs. 11(a) and 12(a), respectively) using the `R` package `RNiftyReg` and a linear (affine) transformation with 12 degrees of freedom (Clayton et al., 2018). Spatial concordance was assessed between the resampled higher-resolution images (Fig. 11(a) or 12(a)) and the coordinate-registered lower-resolution images (Fig. 11(c) or 12(c)).

(a) (b)

Fig. 10. The two images from Harvard Forest after cropping to equivalent views and resampling to equivalent pixel dimensions. (a): Image taken with an outdoor StarDot NetCam XL 3MP camera (Fig. 9(a)); (b): Image taken with an outdoor Axis 223M camera (9(b)).



(a) (b)

(c)

Fig. 11. The two cropped and resampled images (Figs. 10(a) and 10(b)) converted to grayscale, and the coordinate registration of the second image with respect to the first. (a): Image taken with an outdoor StarDot NetCam XL 3MP camera (Fig. 10(a)); (b): Image taken with an outdoor Axis 223M camera (Fig. 10(b)); (c): Image (b) registered to image (a).

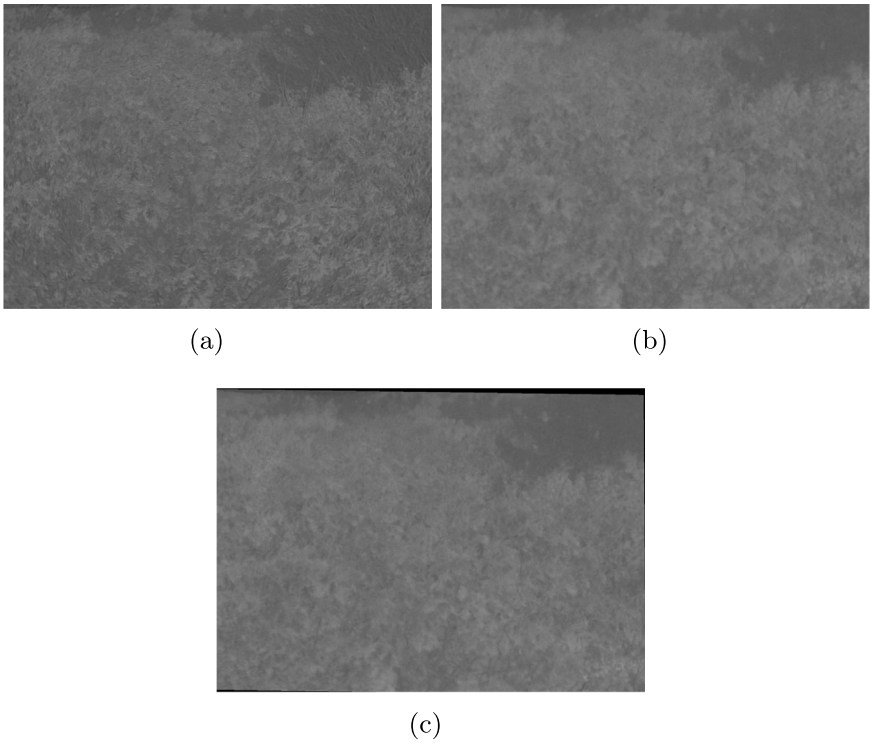


Fig. 12. Cropped and resampled images (Figs. 10(a) and 10(b)) corrected for brightness using the green chromatic coordinate (g_{cc}), and the coordinate registration of the second image with respect to the first. (a): g_{cc} for Fig. 10(a); (b): g_{cc} for Fig. 10(b); (c): Image (b) registered to image (a).

7.3. Estimating concordance

For each pair of images, we first calculated Lin's (1989) CCC. We then calculated the SCCC as described in Section 5. We calculated the local concordance coefficient $\rho_i(\cdot)$ in small (12 × 12-pixel) non-overlapping windows. To fit the local model to each small window, we used a Gaussian process $\boldsymbol{Z}(\boldsymbol{s}) = (X(\boldsymbol{s}), Y(\boldsymbol{s}))^\top$, $\boldsymbol{s} \in \mathbb{R}^2$, with mean $(\mu_X, \mu_Y)^\top$ and the covariance functions described in Eqs. (6)–(9). We used the function `GeoFit` in the `R` package `GeoModels` (Bevilacqua and Morales-Oñate, 2018) to compute the ML estimators of the parameters involved in the models. For computational efficiency, the Matérn and Wendland–Gneiting covariances were estimated for a randomly-selected set of 1,467 20 × 20-pixel sub-images; the model to be used was selected based on the Akaike and Bayesian Information Criteria (AIC and BIC, respectively). In general, the AIC and BIC coefficients were smaller for estimates using the Matérn covariance than for the Wendland–Gneiting covariance, and so we used the Matérn model even though it took somewhat more time to use it to compute the local estimators. Finally, the global SCCCs for each pair of images were estimated using Eqs. (14) and (15).

7.4. Estimates of concordance

Lin's coefficient was $\rho_c = 0.1334$ for the grayscale images (Fig. 11(a) *vs.* 11(c)) and $\rho_c = 0.2450$ for the g_{cc}-indexed images (Fig. 12(a) *vs.* 12(c)). In Fig. 13 we plot Lin's coefficient and the two global coefficients as a function of the spatial norm. We observed a rapid decay of $\widehat{\rho}_2(\cdot)$ and a slower decay of $\widehat{\rho}_1(\cdot)$. The decay was related to the way in which the estimates were computed

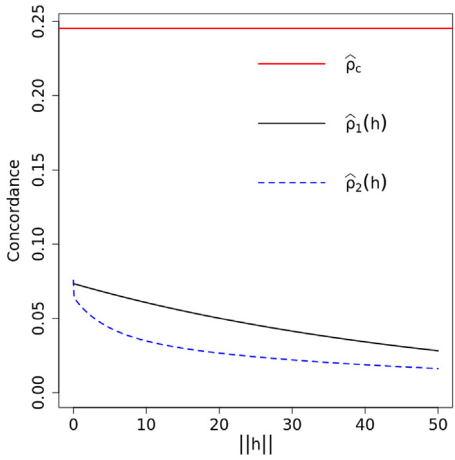


Fig. 13. Global concordance coefficients and Lin's CCC for the g_{cc}-indexed images.

for each window: $\widehat{\rho}_1(\cdot)$ is a coefficient obtained by plugging in the average of the parameters in the concordance function, but $\widehat{\rho}_2(\cdot)$ is the average of the concordance using all possible windows.

For $\|\boldsymbol{h}\| = \mathbf{0}$, we observed that the SCCC was approximately one-third ($0.08/0.245 \times 100 \approx 0.33$) of Lin's CCC. This suggests that Lin's CCC overestimated the spatial concordance between these two images, and implies that it would be inappropriate to use it for modeling spatial data.

It is also worth mentioning that the low SCCC between two images like these – high-resolution *versus* low-resolution registered to high-resolution – implies that even at small lags there is substantial information lost in the post-processing of the lower-resolution image. Thus, one should be cautious in trying to up-scale information from either lower-resolution imagery or similar (e.g., historical) datasets. We are unlikely to be able to extract more information from an image than is actually contained within it.

We also note that the development of hypothesis testing about the SCCC using Theorem 1 requires the computation of the asymptotic variance, which could be complicated for an arbitrary ν. Resampling methods provide alternatives that could help estimate the asymptotic variance of the sample version of the SCCC. Castillo-Páez et al. (2019) present recent work along these lines in a spatial framework.

The images and all the code used in this paper are available from the Environmental Data Initiative doi: 10.6073/pasta/4a975798efc9e099104a1e4391f73574.

8. Discussion

With the work presented herein, we have extended the standard methodology for estimating concordance into the spatial domain. Our approach consisted in defining a new coefficient that preserves the interpretation of Lin's (1989) concordance correlation coefficient (CCC) for two spatial variables and for a fixed spatial lag. Our new spatial concordance correlation coefficient (SCCC) compares the correlation between two spatial variables with respect to their fit to a 45° line that passes through the origin. The properties of Lin's (1989) CCC are inherited by our SCCC. The ML estimator of our SCCC for the Wendland–Gneiting covariance function is asymptotically normal for an increasing domain sampling scheme. We defined a local SCCC and established its asymptotic normality for the sample version. From the local SCCC, we derived two estimates for the overall SCCC, one based on the average of the p local coefficients and the other based on the average of the parameters in the correlation function. Deriving the global SCCC from local coefficients estimated in small non-overlapping windows is computationally more efficient and permits the

estimation of spatial concordance for large images that are used commonly in a wide range of applications.

The Monte Carlo simulation study presented in Section 6 revealed that for the Matérn and Wendland–Gneiting covariance functions, the sample version of the SCCC produced accurate estimates of the SCCC that decreased with distance (spatial lag). However, the time required to compute SCCC grows exponentially with window size, implying that for a large image size it is unfeasible to compute $\widehat{\rho}(\cdot)$ using an interpreted language like R. Although we are exploring ways to improve computational efficiency, the local approach introduced here (Section 5) appears to be a straightforward way to estimate SCCC for large images.

The camera comparison experiment conducted by Sonnentag et al. (2012) found that images recorded with a variety of different camera makes and models, all mounted on the top of the same canopy access tower and with a similar field of view, varied in visual appearance, including color balance, saturation, contrast, and brightness. These differences can be attributed to internal differences in sensor design and image processing, and external factors such as lighting. However, Sonnentag et al. (2012) also found that when simple normalized indices were calculated from the image data, and the emphasis was placed on the seasonality – rather than absolute magnitude – of those indices, the phenological information derived from the imagery was extremely similar across all cameras. Notably, their analysis focused on information about the average color across a large "region of interest" drawn across the canopy (Sonnentag et al., 2012). Although this approach is widely used (Richardson, 2018) and it has the advantage of enabling integration across multiple individuals or species that may comprise a typical forest canopy, it lacks spatial information.

The SCCC we developed and presented here summarizes and accounts for the spatial information in the images, permitting more rigorous characterization of agreement between high-resolution digital images recorded by different sensors. Other applications include using images from different satellite remote-sensing platforms as part of ongoing efforts to harmonize, for example, imagery with different spatial resolution, spectral sensitivity, and angular characteristics (e.g., Landsat–Sentinel efforts: Claverie et al., 2018). Calculation of concordance statistics before and after sensor harmonization could provide critical and objective information about the success of different harmonization methods. There also could be potential applications in the fusion of remotely-sensed data obtained at different spatiotemporal resolutions, such as MODIS with its 500-m spatial resolution and daily temporal resolution, and Landsat with its 30-m spatial resolution and 16-day temporal resolution (Gao et al., 2015). Another application in ecology in which the SCCC coefficient could be used is for improving census information or better estimating detection probability of mobile organisms captured on camera traps. The aim could be to try to match animal coat or skin markings to reliably identify a re-sighting of the same individual *versus* different ones. In environmental sciences it is also of interest to test the agreement between different sensors that measure air quality in polluted cities. For example, Santiago, Chile has only 12 air-quality monitoring stations (The World Air Quality Index project: http://aqicn.org/contact/es/). The need for new stations makes relevant the spatial concordance analysis between different types of sensors in a calibration context.

An important part of the analysis of agreement between two images using the SCCC is the preprocessing applied to the images before computing the concordance index. Most filters deteriorate the spatial information contained in both images. Concordance indices calculated after image processing capture only relationships between the remaining patterns and could yield low concordance values that underestimate the real spatial agreement. The question of what kind of filters preserve spatial concordance is left as an open problem to be tackled in future research as, to the best of our knowledge, this problem has not yet been treated in the image-processing literature.

Another important aspect in modeling the spatial concordance between images is their subdivision into small windows to estimate the SCCC using a local approach. We are aware that when fitting each window using a different Matérn model, there is no continuity of the smoothing parameters. Therefore, two contiguous processes could have very different estimates of the smoothing parameters, producing a discontinuity at their edges. One way to approach this problem is to use the local likelihood to define a smoothing function in the lines (Anderes and Stein, 2011). With respect to the estimation of the bivariate Matérn model, we observed that when all parameters were estimated,

the ML estimates could be highly inefficient. Thus, we agree with Gneiting et al. (2010) that the best predictive models tend to be the most parsimonious, subject to retention of key characteristics, such as the negative dependency between the two process components (see also Makridakis and Taleb, 2009). Finally, the selection of the spatial lag $\boldsymbol{h}$ is crucial if there is no information about which direction to prefer. One way to approach this issue is to consider an average of spatial concordance values corresponding to the four cardinal directions with respect to the origin (Vallejos et al., 2015), i.e., for $\boldsymbol{h}_1 = (1, 0)$ and $\boldsymbol{h}_2 = (0, 1)$, we can define the average concordance coefficient as

$$\rho^c_{AV}(\boldsymbol{h}_1, \boldsymbol{h}_2) := \frac{1}{4}\left[\rho^c(\boldsymbol{h}_1) + \rho^c(-\boldsymbol{h}_1) + \rho^c(\boldsymbol{h}_2) + \rho^c(-\boldsymbol{h}_2)\right].$$

9. Future work

Several related theoretical and applied problems arise from the methodology suggested in this article that would be fruitful directions for future research. First, SCCC could be applied to images taken at two points in time by the same camera. The decay of the SCCC as a function of the norm would be expected to be similar to that seen in Fig. 13 for each sequential pair of images. Another approach for dealing with the same problem would be to consider a sequence of n images taken with the same camera to be a spatiotemporal process. Then, the SCCC and its estimation properties could be studied in that context. This generalization of the SCCC would have applications in, for example, spatiotemporal analysis of satellite images taken weeks, months, or years apart as a way of characterizing patterns of landscape change.

Acknowledgments

This work has been partially supported by the AC3E, UTFSM, Chile, under grant FB-0008, and by grants to A.D.R. from the US National Science Foundation (EF-1065029, EF-1702697, DEB-1237491) and the United States Geological Survey (G10AP00129). This is a contribution from the Harvard Forest Long-term Ecological Research (LTER) site, supported most recently by the US National Science Foundation (DEB 18-32210). R.Vallejos acknowledges support from the MATH-AmSud Regional Program (20-MATH-03).

Appendix

Mardia and Marshall theorem.

Let $\left\{Y(\boldsymbol{s}) : \boldsymbol{s} \in D \subset \mathbb{R}^d\right\}$ be a Gaussian random field such that $Y(\cdot)$ is observed on $D_n \subset D$. It is assumed that D_n is a non-random set satisfying $\|\boldsymbol{s} - \boldsymbol{t}\| \geq \gamma > 0$ for all $\boldsymbol{s}, \boldsymbol{t} \in D_n$. This ensures that the sampling set is increasing as n increases. Denote $\boldsymbol{Y} = (Y(\boldsymbol{s}_1), \ldots, Y(\boldsymbol{s}_n))^\top$ and assume that $\mathbb{E}[\boldsymbol{Y}] = \boldsymbol{X}\boldsymbol{\beta}$, $\mathrm{cov}(Y(\boldsymbol{t}), Y(\boldsymbol{s})) = \sigma(\boldsymbol{t}, \boldsymbol{s}; \boldsymbol{\theta})$, $\boldsymbol{X}$ is $n \times p$ with $\mathrm{rank}(\boldsymbol{X}) = p$, $\boldsymbol{\beta} \in \mathbb{R}^p$, and $\boldsymbol{\theta} \in \Theta$, where Θ is an open set of $\in \mathbb{R}^q$. Let $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$ be the covariance matrix of $\boldsymbol{Y}$ such that the ij-th element of $\boldsymbol{\Sigma}$ is $\sigma_{ij} = \sigma(\boldsymbol{s}_i, \boldsymbol{s}_j; \boldsymbol{\theta})$. We can estimate $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ using ML, by maximizing

$$L = L(\boldsymbol{\beta}, \boldsymbol{\theta}) = k - \frac{1}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}), \tag{A.1}$$

where k is a constant.

Let $\boldsymbol{L}_n^{(1)} = \nabla L = (\boldsymbol{L}_{\boldsymbol{\beta}}^\top, \boldsymbol{L}_{\boldsymbol{\theta}}^\top)^\top$ and

$$\boldsymbol{L}_n^{(2)} = \begin{pmatrix} \boldsymbol{L}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \boldsymbol{L}_{\boldsymbol{\beta}\boldsymbol{\theta}} \\ \boldsymbol{L}_{\boldsymbol{\theta}\boldsymbol{\beta}} & \boldsymbol{L}_{\boldsymbol{\theta}\boldsymbol{\theta}} \end{pmatrix}$$

be the gradient vector and Hessian matrix, respectively, obtained from Eq. (A.1). Let $\boldsymbol{F}_n = -\mathbb{E}[\boldsymbol{L}_n^{(2)}]$ be the Fisher information matrix with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$. Then, $\boldsymbol{F}_n = \mathrm{diag}(\boldsymbol{F}_n(\boldsymbol{\beta}), \boldsymbol{F}_n(\boldsymbol{\theta}))$, where $\boldsymbol{F}_n(\boldsymbol{\beta}) = -\mathbb{E}[\boldsymbol{L}_{\boldsymbol{\beta}\boldsymbol{\beta}}]$ and $\boldsymbol{F}_n(\boldsymbol{\theta}) = -\mathbb{E}[\boldsymbol{L}_{\boldsymbol{\theta}\boldsymbol{\theta}}]$.

For a twice differentiable covariance function $\sigma(\cdot,\cdot;\boldsymbol{\theta})$ on Θ with continuous second derivatives, Mardia and Marshall (1984) provided sufficient conditions on $\boldsymbol{\Sigma}$ and $\boldsymbol{X}$ such that the limiting distribution of $(\widehat{\boldsymbol{\beta}}_n^\top, \widehat{\boldsymbol{\theta}}_n^\top)^\top$ is normal, per the following:

Theorem. Let $\lambda_1 \leq \cdots \leq \lambda_n$ be the eigenvalues of $\boldsymbol{\Sigma}$, and let those of $\boldsymbol{\Sigma}_i = \dfrac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}$ and $\boldsymbol{\Sigma}_{ij} = \dfrac{\partial^2 \boldsymbol{\Sigma}}{\partial \theta_i \partial \theta_j}$ be λ_k^i and λ_k^{ij}, $k = 1, \ldots, n$, such that $|\lambda_1^i| \leq \cdots \leq |\lambda_n^i|$ and $|\lambda_1^{ij}| \leq \cdots \leq |\lambda_n^{ij}|$ for $i, j = 1, \ldots, q$. Suppose that as $n \to \infty$

(i) $\lim \lambda_n = C < \infty$, $\lim |\lambda_n^i| = C_i < \infty$ and $\lim |\lambda_n^{ij}| = C_{ij} < \infty$ for all $i, j = 1, \ldots, q$.
(ii) $\|\boldsymbol{\Sigma}_i\|^{-2} = \mathcal{O}(n^{-\frac{1}{2}-\delta})$ for some $\delta > 0$, for $i = 1, \ldots, q$.
(iii) For all $i, j = 1, \ldots, q$, $a_{ij} = \lim \left[t_{ij}/(t_{ii} t_{jj})^{\frac{1}{2}} \right]$ exists, where $t_{ij} = \operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_j\right)$ and $\boldsymbol{A} = (a_{ij})$ is nonsingular.
(iv) $\lim(\boldsymbol{X}^\top\boldsymbol{X})^{-1} = \boldsymbol{0}$.

Then, $(\widehat{\boldsymbol{\beta}}_n^\top, \widehat{\boldsymbol{\theta}}_n^\top)^\top \xrightarrow{\mathcal{L}} \mathcal{N}\left((\boldsymbol{\beta}^\top, \boldsymbol{\theta}^\top)^\top, \boldsymbol{F}_n^{-1}\right)$ as $n \to \infty$, in an increasing domain sense.

Proof of Theorem 1. The proof consists of verifying the Mardia and Marshall (1984) conditions. In Theorem 1, $\mathbb{E}[\boldsymbol{Z}(\boldsymbol{s})] = \boldsymbol{0}$; thus the fourth condition in Mardia and Marshall's 1984 theorem (above), $\lim(\boldsymbol{X}^\top\boldsymbol{X})^{-1} = \boldsymbol{0}$, is trivially satisfied. Satisfying the first three conditions is somewhat more complex.

For the first two conditions, we start by considering ν to be fixed. Then

$$\boldsymbol{C}_X(\boldsymbol{h}) = \sigma_X^2 \left(1 + (\nu+1)\frac{\|\boldsymbol{h}\|}{b_X}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_X}\right)_+^{\nu+1},$$

$$\boldsymbol{C}_Y(\boldsymbol{h}) = \sigma_Y^2 \left(1 + (\nu+1)\frac{\|\boldsymbol{h}\|}{b_Y}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_Y}\right)_+^{\nu+1},$$

$$\boldsymbol{C}_{XY}(\boldsymbol{h}) = C_{YX}(\boldsymbol{h}) = \rho_{XY}\sigma_X\sigma_Y \left(1 + (\nu+1)\frac{\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1}.$$

Let us consider an increasing domain scenario for process $\boldsymbol{Z}(\boldsymbol{s})$, with points $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ located in a rectangle $D_n \subset \Delta\mathbb{Z}^d$, for $0 < \Delta < \infty$, and $D_n \subset D_{n+1}$, for all n.

Define the distance matrix $\boldsymbol{H}_n = \left[H_{lq}\right]_{l=q=1}^n$, where $H_{lq} = \|\boldsymbol{s}_l - \boldsymbol{s}_q\|$, and $\|\cdot\|$ denotes the Euclidean norm. Then the covariance matrix of $(\boldsymbol{Z}(\boldsymbol{s}_1)^\top, \ldots, \boldsymbol{Z}(\boldsymbol{s}_n)^\top)^\top$ can be written as

$$\boldsymbol{\Sigma}_n(\boldsymbol{\theta}) = \begin{pmatrix} \sigma_X^2 & \sigma_X\sigma_Y\rho_{XY} \\ - & \sigma_Y^2 \end{pmatrix} \otimes \boldsymbol{\Gamma}_n,$$

where $\boldsymbol{\Gamma}_n = \left[\left(1 + \dfrac{(\nu+1)H_{lq}}{b_{XY}}\right)\left(1 - \dfrac{H_{lq}}{b_{XY}}\right)_+^{\nu+1}\right]_{l=q=1}^n$ and $\boldsymbol{\theta} = (\sigma_X^2, \sigma_Y^2, \rho_{XY}, b_{XY})^\top$. Taking derivatives, we obtain

$$\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_X^2} = \begin{pmatrix} 1 & \dfrac{\sigma_Y\rho_{XY}}{2\sigma_X} \\ - & 0 \end{pmatrix} \otimes \boldsymbol{\Gamma}_n, \qquad \frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_Y^2} = \begin{pmatrix} 0 & \dfrac{\sigma_X\rho_{XY}}{2\sigma_Y} \\ - & 1 \end{pmatrix} \otimes \boldsymbol{\Gamma}_n,$$

$$\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial b_{XY}} = \begin{pmatrix} \sigma_X^2 & \sigma_X\sigma_Y\rho_{XY} \\ - & \sigma_Y^2 \end{pmatrix} \otimes \boldsymbol{S}_n, \qquad \frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \rho_{XY}} = \begin{pmatrix} 0 & \sigma_X\sigma_Y \\ - & 0 \end{pmatrix} \otimes \boldsymbol{\Gamma}_n,$$

where $\boldsymbol{S}_n$ is given by

$$\boldsymbol{S}_n = \frac{\partial \boldsymbol{\Gamma}_n}{\partial b_{XY}} = \left[\frac{(\nu+1)H_{lq}}{b_{XY}^2}\left(1 - \frac{H_{lq}}{b_{XY}}\right)_+^{\nu}\left(-\left(1 - \frac{H_{lq}}{b_{XY}}\right)_+ + \left(1 + \frac{(\nu+1)H_{lq}}{b_{XY}}\right)\right)\right]_{l=q=1}^n.$$

For any matrix norm, the spectral radius $\lambda_{\max}\{\boldsymbol{A}\}$ of an $n \times n$ matrix $\boldsymbol{A}$ satisfies $\lambda_{\max}\{\boldsymbol{A}\} \leq \|\boldsymbol{A}\|$. Then, considering the norm $\|\cdot\|_\infty$, we have

$$\begin{aligned}\lambda_{\max}\{\boldsymbol{\Gamma}_n\} \leq \|\boldsymbol{\Gamma}_n\|_\infty &= \max_l \sum_{q=1}^{n} \left| \left(1 + \frac{(\nu+1)H_{lq}}{b_{XY}}\right) \left(1 - \frac{H_{lq}}{b_{XY}}\right)_+^{\nu+1} \right| \\ &= \sup_{1\leq l\leq n} \sum_{q=1}^{n} \left| \left(1 + \frac{(\nu+1)H_{lq}}{b_{XY}}\right) \left(1 - \frac{H_{lq}}{b_{XY}}\right)_+^{\nu+1} \right| \\ &< \sum_{s\in\Delta\mathbb{Z}^d} \left(1 + \frac{(\nu+1)\|\boldsymbol{s}\|}{b_{XY}}\right) \left(1 - \frac{\|\boldsymbol{s}\|}{b_{XY}}\right)_+^{\nu+1}.\end{aligned}$$

One can check that

$$\int_{s\in\mathbb{R}^d} \left(1 + \frac{(\nu+1)\|s\|}{b_{XY}}\right) \left(1 - \frac{\|\boldsymbol{s}\|}{b_{XY}}\right)_+^{\nu+1} ds < \infty.$$

Thus $\sup_n \lambda_{\max}\{\boldsymbol{\Gamma}_n\} < \infty$, which implies that $\sup_n \lambda_{\max}\{\boldsymbol{\Sigma}_n(\boldsymbol{\theta})\} < \infty$. Because $\boldsymbol{\Gamma}_n$ is positive definite, $\lambda_i\{\boldsymbol{\Gamma}_n\} > 0$, $i = 1, \ldots, n$. In particular, $\lambda_{\min}\{\boldsymbol{\Gamma}_n\} > 0$, so $\inf_n \lambda_{\min}\{\boldsymbol{\Gamma}_n\} > 0$ and $\inf_n \lambda_{\min}\{\boldsymbol{\Sigma}_n\} > 0$. Further,

$$\sup_n \lambda_{\max}\left\{\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_X^2}\right\} = \sup_n \lambda_{\max}\left[\begin{pmatrix} 1 & \frac{\sigma_Y\rho_{XY}}{2\sigma_X} \\ - & 0 \end{pmatrix} \otimes \boldsymbol{\Gamma}_n\right] < \infty, \text{ for } \frac{\sigma_Y\rho_{XY}}{2\sigma_X} < \infty.$$

Similarly,

$$\sup_n \lambda_{\max}\left\{\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_Y^2}\right\}, \sup_n \lambda_{\max}\left\{\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \rho_{XY}}\right\} < \infty.$$

Moreover, $\lambda_{\max}\{\boldsymbol{S}_n\} \leq \|\boldsymbol{S}_n\|_\infty < \infty$ because of the form of the polynomial in $\boldsymbol{s} \in \mathbb{R}^d$ and the compact support in b_{XY}. Then, for $\sigma_X^2, \sigma_Y^2, \sigma_X\sigma_Y\rho_{XY} < \infty$,

$$\sup_n \lambda_{\max}\left\{\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial b_{XY}}\right\} < \infty.$$

This implies that,

$$\sup_n \lambda_{\max}\left\{\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \theta_c}\right\} < \infty, \quad c = 1, 2, 3, 4.$$

The second derivatives are:

$$\frac{\partial^2 \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_X^2 \partial \sigma_Y^2} = \begin{pmatrix} 0 & \frac{\rho_{XY}}{4\sigma_X\sigma_Y} \\ - & 0 \end{pmatrix} \otimes \boldsymbol{\Gamma}_n, \qquad \frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_X^2 \partial b_{XY}} = \begin{pmatrix} 1 & \frac{\sigma_Y\rho_{XY}}{2\sigma_X} \\ - & 0 \end{pmatrix} \otimes \boldsymbol{S}_n,$$

$$\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_X^2 \partial \rho_{XY}} = \begin{pmatrix} 0 & \frac{\sigma_Y}{2\sigma_X} \\ - & 0 \end{pmatrix} \otimes \boldsymbol{\Gamma}_n, \qquad \frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_X^4} = \begin{pmatrix} 0 & -\frac{\sigma_Y\rho_{XY}}{4\sigma_X^3} \\ - & 0 \end{pmatrix} \otimes \boldsymbol{\Gamma}_n,$$

$$\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_Y^2 \partial b_{XY}} = \begin{pmatrix} 0 & \frac{\sigma_X\rho_{XY}}{2\sigma_Y} \\ - & 1 \end{pmatrix} \otimes \boldsymbol{S}_n, \qquad \frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_Y^2 \partial \rho_{XY}} = \begin{pmatrix} 0 & \frac{\sigma_X}{2\sigma_Y} \\ - & 0 \end{pmatrix} \otimes \boldsymbol{\Gamma}_n,$$

$$\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \sigma_Y^4} = \begin{pmatrix} 0 & -\frac{\sigma_X\rho_{XY}}{4\sigma_Y^3} \\ - & 0 \end{pmatrix} \otimes \boldsymbol{\Gamma}_n, \qquad \frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial b_{XY} \partial \rho_{XY}} = \begin{pmatrix} 0 & \sigma_X\sigma_Y \\ - & 0 \end{pmatrix} \otimes \boldsymbol{S}_n,$$

$$\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial b_{XY}^2} = \begin{pmatrix} \sigma_X^2 & \sigma_X\sigma_Y\rho_{XY} \\ - & \sigma_Y^2 \end{pmatrix} \otimes \boldsymbol{SS}_n, \qquad \frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \rho_{XY}^2} = 0,$$

where $\boldsymbol{SS}_n = \dfrac{\partial \boldsymbol{S}_n}{\partial b_{XY}}$.

Because $\sup_n \lambda_{\max}\{\mathbf{0}\} < \infty$, the compact support of $\boldsymbol{SS}_n$ in b_{XY}, and the previous results, $\lambda_{\max}\{\boldsymbol{SS}_n\} \leq \|\boldsymbol{SS}_n\|_\infty < \infty$. Then, for $\sigma_X^2, \sigma_Y^2, \sigma_X\sigma_Y < \infty$,

$$\sup_n \lambda_{\max}\left\{\frac{\partial^2 \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial b_{XY}^2}\right\} < \infty.$$

In addition,

$$\left\|\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \theta_i}\right\|_\infty \leq \left\|\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \theta_i}\right\| \leq \sqrt{n}\left\|\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \theta_i}\right\|_\infty.$$

This satisfies the first two conditions of Mardia and Marshall's theorem.

For the third condition, we consider $\boldsymbol{A} = [a_{ij}]_{i=j=1}^p$, with $a_{ij} = \left\{\frac{t_{ij}}{(t_{mm}t_{nn})^{1/2}}\right\}$, and $t_{ij} = \operatorname{tr}\left\{\boldsymbol{\Sigma}_n(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \theta_i}\boldsymbol{\Sigma}_n(\boldsymbol{\theta})^{-1}\frac{\partial \boldsymbol{\Sigma}_n(\boldsymbol{\theta})}{\partial \theta_j}\right\}$ for all $i, j = 1, \ldots, p$; we prove that $\boldsymbol{A}$ is non singular.

Notice that

$$\boldsymbol{T} = [t_{ij}]_{i=j=1}^4 = \begin{pmatrix} \frac{n(\rho_{XY}^2-2)}{4\sigma_X^4(\rho_{XY}^2-1)} & \frac{n\rho_{XY}^2}{4\sigma_X^2\sigma_Y^2(\rho_{XY}^2-1)} & \frac{1}{2\sigma_X^2}\operatorname{tr}\{\boldsymbol{A}_n\} & \frac{n\rho_{XY}}{2\sigma_X^2(\rho_{XY}^2-1)} \\ - & \frac{n(\rho_{XY}^2-2)}{4\sigma_Y^4(\rho_{XY}^2-1)} & \frac{1}{2\sigma_Y^2}\operatorname{tr}\{\boldsymbol{A}_n\} & \frac{n\rho_{XY}}{2\sigma_Y^2(\rho_{XY}^2-1)} \\ - & - & \operatorname{tr}\{[\boldsymbol{A}_n]^2\} & \frac{\rho_{XY}}{\rho_{XY}^2-1}\operatorname{tr}\{\boldsymbol{A}_n\} \\ - & - & - & \frac{n(\rho_{XY}^2+1)}{(\rho_{XY}^2-1)^2} \end{pmatrix},$$

with $\boldsymbol{A}_n = \{\boldsymbol{\Gamma}_n^{-1} \circ \boldsymbol{S}_n\}$ where the operator $\circ$ denotes the matrix Hadamard product.

Then,

$$\boldsymbol{A} = \begin{pmatrix} 1 & \frac{\rho_{XY}^2}{\rho_{XY}^2-2} & \frac{\operatorname{tr}(\boldsymbol{A}_n)}{\left(\frac{n(\rho_{XY}^2-2)\operatorname{tr}([\boldsymbol{A}_n]^2)}{\rho_{XY}^2-1}\right)^{1/2}} & \frac{\rho_{XY}}{\left(\frac{(\rho_{XY}^2-2)(\rho_{XY}^2+1)}{\rho_{XY}^2-1}\right)^{1/2}} \\ - & 1 & \frac{\operatorname{tr}(\boldsymbol{A}_n)}{\left(\frac{n(\rho_{XY}^2-2)\operatorname{tr}([\boldsymbol{A}_n]^2)}{\rho_{XY}^2-1}\right)^{1/2}} & \frac{\rho_{XY}}{\left(\frac{(\rho_{XY}^2-2)(\rho_{XY}^2+1)}{\rho_{XY}^2-1}\right)^{1/2}} \\ - & - & 1 & \frac{-\rho_{XY}\operatorname{tr}(\boldsymbol{A}_n)}{(n\operatorname{tr}([\boldsymbol{A}_n]^2)(\rho_{XY}^2+1))^{1/2}} \\ - & - & - & 1 \end{pmatrix}. \tag{A.2}$$

For matrix $\boldsymbol{A}$ in Eq. (A.2),we have extended the result established by Bevilaqua et al. (2015). Thus $\boldsymbol{A}$ is positive definite. By Mardia and Marshall's Theorem the ML estimator of $\boldsymbol{\theta} = (\sigma_X^2, \sigma_Y^2, \rho_{XY}, b_{XY})^\top$ is asymptotically normal with variance $\boldsymbol{F}_n(\boldsymbol{\theta})^{-1}$.

Eq. (10) implies that

$$\rho^c(\boldsymbol{h}) = g(\boldsymbol{\theta}) = \frac{2\rho_{XY}\sigma_X\sigma_Y\left(1+(\nu+1)\frac{\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1-\frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1}}{\sigma_X^2+\sigma_Y^2}.$$

Fixing $\nu > 0$, noting that $g(\cdot)$ is a continuously differentiable function for $\sigma_X \neq 0$ and $\sigma_Y \neq 0$, and using the multivariate delta method for $g(\cdot)$ we obtain

$$\left(\nabla g(\boldsymbol{\theta})^\top \boldsymbol{F}_n(\boldsymbol{\theta})^{-1}\nabla g(\boldsymbol{\theta})\right)^{-1/2}(g(\boldsymbol{\theta}_n) - g(\boldsymbol{\theta})) \xrightarrow{D} N(0,1),$$

where

$$\nabla g(\boldsymbol{\theta}) = \begin{pmatrix} \dfrac{\sigma_Y \rho_{XY}(\sigma_Y^2 - \sigma_X^2)\left(1 + (\nu+1)\dfrac{\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \dfrac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1}}{\sigma_1(\sigma_X^2 + \sigma_Y^2)^2} \\ \dfrac{\sigma_X \rho_{XY}(\sigma_X^2 - \sigma_Y^2)\left(1 + (\nu+1)\dfrac{\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \dfrac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1}}{\sigma_Y(\sigma_X^2 + \sigma_Y^2)^2} \\ \dfrac{2\sigma_X\sigma_Y\left(1 + (\nu+1)\dfrac{\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \dfrac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1}}{\sigma_X^2 + \sigma_Y^2} \\ \dfrac{2\sigma_X\sigma_Y\rho_{XY} f(b_{XY})}{\sigma_X^2 + \sigma_Y^2} \end{pmatrix},$$

$$f(b_{XY}) = \left(-\frac{(\nu+1)\|\boldsymbol{h}\|}{b_{XY}^2}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu+1} + \left(1 + \frac{(\nu+1)\|\boldsymbol{h}\|}{b_{XY}}\right)\left(1 - \frac{\|\boldsymbol{h}\|}{b_{XY}}\right)_+^{\nu} \frac{(\nu+1)\|\boldsymbol{h}\|}{b_{XY}^2},$$

$$\boldsymbol{F}_n^{-1}(\boldsymbol{\theta}) = \begin{pmatrix} \dfrac{\sigma_X^4(n\mathrm{tr}(\boldsymbol{B}^2) - C)}{nC} & \dfrac{\sigma_X^2\sigma_Y^2([\mathrm{tr}(\boldsymbol{B})]^2 - 2\rho_{XY}^2 C)}{nC} & -\dfrac{\sigma_X^2\mathrm{tr}(\boldsymbol{B})}{C} & -\dfrac{\sigma_X^2\rho_{XY}(\rho_{XY}^2 - 1)}{n} \\ - & \dfrac{\sigma_2^4(n\mathrm{tr}(\boldsymbol{B}^2) - C)}{nC} & -\dfrac{\sigma_Y^2\mathrm{tr}(\boldsymbol{B})}{C} & -\dfrac{\sigma_Y^2\rho_{XY}(\rho_{XY}^2 - 1)}{n} \\ - & - & \dfrac{n}{C} & 0 \\ - & - & - & \dfrac{(\rho_{XY}^2 - 1)^2}{n} \end{pmatrix},$$

$\boldsymbol{B} = \left(\boldsymbol{R}^{-1}\dfrac{\partial \boldsymbol{R}}{\partial \boldsymbol{\phi}}\right)$, $C = n\mathrm{tr}(\boldsymbol{B}^2) - [\mathrm{tr}(\boldsymbol{B})]^2$, and $\boldsymbol{R} = [R(\boldsymbol{h}, \boldsymbol{\phi})]_{i,j=1}^n$. □

References

Anderes, E.B., Stein, M.L., 2011. Local likelihood estimation for nonstationary random fields. J. Multivariate Anal. 102, 506–520.

Atkinson, G., Nevill, A., 1997. Comments on the use of concordance correlation to assess the agreement between two variables. Biometrics 52, 775–778.

Barnhart, H.X., Haber, M.J., Line, L.I., 2007. An overview on assessing agreement with continuous measurements. J. Biopharm. Statist. 17, 529–569.

Bevilacqua, M., Morales-Oñate, V., 2018. Geomodels: a package for geostatistical gaussian and non gaussian data analysis. https://vmoprojs.github.io/GeoModels-page/. R package version 1.0.3-4.

Bevilaqua, M., Faouzi, T., Furrer, R., Porcu, E., 2019. Estimation and prediction using generalized wendland covariance functions under fixed effects asymptotics. Ann. Statist. 47, 828–856.

Bevilaqua, M., Vallejos, R., Velandia, D., 2015. Assessing the significance of the correlation between the components of a bivariate gaussian random field. Environmetrics 26, 545–556.

Bustos, O., Ojeda, S., and, R.V., 2009. Spatial ARMA models and its applications to image filtering. Braz. J. Probab. Stat. 23, 141–165.

Castillo-Páez, S., Fernández-Casal, R., García-Soidán, P., 2019. A nonparametric bootstrap method for spatial data. Comput. Statist. Data Anal. 137, 1–15.

Chodhary, P., Nagaraja, H., 2017. Measuring Agreement, Models, Methods, and Applications. Wiley, New York.

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S.V., Justice, C., 2018. The harmonized landsat and sentinel-2 surface reflectance data set. Remote Sens. Environ. 219, 145–161.

Clayton, J., Modat, M., Presles, B., Anthopoulos, T., Daga, P., 2018. Package RNiftyReg. https://cran.r-project.org/package=RNiftyReg.

Cohen, J., 1968. Weighted kappa: nominal scale agreement with provision for scale disagreement or partial credit. Psycol. Bull. 70, 213–220.

Daley, D., Porcu, E., Bevilacqua, M., 2015. Classes of compactly supported covariance functions for multivariate random fields. Stoch. Environ. Res. Risk Assess. 29, 1249–1263.

Gao, F., Hilker, T., Zhu, X., Anderson, M., Masek, J., Wang, P., Yang, Y., 2015. Fusing landsat and MODIS data for vegetation monitoring. IEEE Geosci. Remote Sens. Mag. 3, 47–60.

Gillespie, A., Kahle, A., Walker, R., 1987. Color enhancement of highly correlated images. 2. Channel ratio and chromaticity transformation techniques. Remote Sens. Environ. 22, 343—365.
Gneiting, T., 2002. Compactly supported correlation functions. J. Multivariate Anal. 83, 493–508.
Gneiting, T., Kleiber, W., Schlather, M., 2010. Matérn cross-covariance functions for multivariate random fields. J. Amer. Statist. Assoc. 105, 1167–1177.
Hiriote, S., Chinchilli, V.M., 2011. Matrix-based concordance correlation coefficient for repeated measures. Biometrics 67, 1007–1016.
Lawrence, I., Chinchilli, V., 1997. Rejoinder to the letter to the editor from Atkinson and Nevill.
Leal, C., Galea, M., Osorio, F., 2019. Assessment of local influence for the analysis of agreement. Biom. J. 61, 955–972.
Lin, L., 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics 45, 255–268.
Lin, L., 2000. Total deviation index for measuring individual agreement: with application in lab performance and bioequivalence. Stat. Med. 19, 255–270.
Lin, L., Hedayat, A., Sinha, B., Yang, M., 2002. Statistical methods in assessing agreement: models, issues, and tools. J. Amer. Statist. Assoc. 97, 257–270.
Lin, L., Hedayat, A.S., Wu, W., 2012. Statistical Tools for Measuring Agreement. Springer Science & Business Media.
Makridakis, S., Taleb, N., 2009. Living in a world of low levels of predictability. Int. J. Forecast. 25, 840–844.
Mardia, K.V., Marshall, T.J., 1984. Maximum likelihood estimation of models for residual covariance in spatial regression. Biometrika 19, 135–146.
Mizunuma, T., Mencuccini, M., Wingate, L., Ogee, J., Nichol, C., et al., 2014. Sensitivity of colour indices for discriminating leaf colours from digital photographs. Methods Ecol. Evol. 5 (10), 1078–1085.
Nguy-Robertson, A.L., Buckley, E.M.B., Suyker, A.S., Awada, T.N., 2016. Determining factors that impact the calibration of consumer-grade digital cameras used for vegetation analysis. Int. J. Remote Sens. 37, 3365–3383.
Ojeda, S., Vallejos, R., Bustos, O., 2010. A new image segmentation algorithm with applications to image inpainting. Comput. Statist. Data Anal. 54, 2082–2093.
R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/.
Richardson, A.D., 2018. Tracking seasonal rhythms of plants in diverse ecosystems with digital camera imagery. New Phytol. 00, http://dx.doi.org/10.1111/nph.15591.
Schall, R., Williams, R.L., 1996. Towards a practical strategy for assessing individual bioequivalence. J. Pharmacokinet. Biopharm. 24, 133–149.
Skiljan, I., 2014. Irfanview. https://www.irfanview.com/.
Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A.M., Friedl, M., Braswell, B.H., Milliman, T., O'Keefe, J., Richardson, A.D., 2012. Digital repeat photography for phenomenal research in forest ecosystems. Agricult. Forest Meteorol. 152, 159–177.
Stevens, N.T., Steiner, S.H., MacKay, R.J., 2017. Assessing agreement between two measurement systems: an alternative to the limits of agreement approach. Stat. Methods Med. Res. 26, 2487–2504.
Toomey, M., Friedl, M.A., Frolking, S., Hufkens, K., Klosterman, S., et al., 2015. Greenness indices from digital cameras predict the timing and seasonal dynamics of canopy-scale photosynthesis. Ecol. Appl. 25, 99–115.
Urbanek, S., 2014. Package jpeg. https://cran.r-project.org/package=jpeg.
Vallejos, R., Mallea, A., Herrera, M., Ojeda, S., 2015. A multivariate geostatistical approach for landscape classification from remotely sensed image data. Stoch. Environ. Res. Risk Assess. 29, 369—378.
Vonesh, E.F., Chinchilli, V.M., Pu, K., 1996. Goodness of fit in generalized nonlinear mixed-effect models. Biometrics 52, 572–587.
Wendland, H., 1995. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. Adv. Comput. Math. 4 (1), 389–396.