Measuring Separability in Spatio-temporal Covariance Functions

Thaís C.O. Fonseca *and Mark F.J. Steel †

May 2, 2017

Abstract

In this work, we propose a measure of space-time dependence for general nonseparable (possibly nonstationary) covariance models. It is well known that nonseparable covariance functions are more realistic for modeling many geophysical and environmental processes. However, little is known about the strength of dependence in space-time that is achieved by the models proposed in the literature. We compute the proposed measure for various nonseparable models which reveals that some of these models generate a rather limited range of nonseparability in space-time. Moreover, spatio-temporal interaction parameters do not always have a monotonous relation to our measure of separability, as they often are not the only parameters affecting the degree of nonseparability obtained by the models.

Key words: Geostatistical modelling; Nonseparability; Nonstationarity.

*Department of Statistics, Federal University of Rio de Janeiro, CEP: 21945-970, Brazil, (Email: thais@im.ufrj.br)

†Department of Statistics, University of Warwick, Coventry, CV4 7AL, U.K. (Email: M.F.Steel@stats.warwick.ac.uk).

1 Introduction

Let $\{Z(s,t); s \in D \subset \Re^d, t \in T \subseteq \Re_+\}$ be a stationary spatiotemporal stochastic process with covariance function $Cov(Z(s_1,t_1), Z(s_2,t_2)) = C(s,t)$, where $s = s_1 - s_2$ and $t = t_1 - t_2$, $s_1, s_2 \in D, t_1, t_2 \in T$. If a spatiotemporal covariance function $C(s,t)$ is separable then it can be written as $C_1(s)C_2(t)$, where $C_1(s)$ and $C_2(t)$ are purely spatial and temporal covariance functions, respectively. In terms of the joint covariance function, a fully symmetric (Gneiting, 2002) covariance function is separable if and only if

$$C(s,t) = \frac{C(s,0)C(0,t)}{C(0,0)}, \tag{1}$$

where $s \in D, t \in T$ are spatial and temporal lags, respectively. The separability assumption is computationally very convenient but usually unrealistic in practice. There are some papers in the literature that propose tests for assessing separability such as Mitchell et al. (2005), Fuentes (2006), Li et al. (2007) and Crujeiras et al. (2010). However these tests are designed only to assess whether a covariance function is separable, and not for measuring the degree of nonseparability. The models proposed in Fonseca and Steel (2011) suggest a measure of nonseparability, that is always between 0 and 1, which is easily assessed as it is a simple function of the parameters in the model. Gneiting (2002) also proposed a valid covariance model that has a parameter measuring the degree of nonseparability that is always between 0 and 1. This model has been very often used in the literature to derive more general and flexible models in several applications. To cite a few, recently Apanasovich and Genton (2010) proposed a class of cross-covariance functions for multivariate random fields which considers the nonseparable proposal of Gneiting (2002) to model space-component dependence through latent dimensions. Also Genton (2007) considered Gneiting functions to derive separable approximations of space-time covariance matrices obtained from nonseparable functions. It is natural then to ask which are the degrees of nonseparability attained by these different models and whether their parameters are comparable. Moreover, it is important to consider how separability in general nonseparable models can be measured.

In this work we show that some of the models proposed in the literature do not achieve strong degrees of nonseparability and that the parameters do not always have clear interpretations. In particular, we investigate the nonseparable models proposed in Cressie and Huang (1999), Gneiting (2002) and Rodrigues and Diggle (2010). Furthermore, we illustrate how our proposal can be applied to nonstationary fields.

Firstly, define the function

$$\tilde{C}(s,t) = \frac{C(s,0)C(0,t)}{C(0,0)}, \tag{2}$$

where $s \in D$ and $t \in T$. If the spatiotemporal covariance function is separable then $\tilde{C}(s,t) = C(s,t)$ from (1). The intuitive idea is that if $\tilde{C}(s,t)$ is close to $C(s,t)$ then the covariance is roughly separable and if $\tilde{C}(s,t)$ is very different from $C(s,t)$ it is very nonseparable. Notice that $\tilde{C}(s,0) = C(s,0)$ and $\tilde{C}(0,t) = C(0,t)$, that is, the margins of $\tilde{C}(s,t)$ and $C(s,t)$ are the same.

So if the covariance function is separable then the ratio $R(s,t) = \frac{\tilde{C}(s,t)}{C(s,t)}$ is equal to 1. Rodrigues and Diggle (2010) define the concept of positive and negative nonseparability. De Iaco and Posa (2013) further investigate the properties of negatively and positively defined separability of space-time covariance functions. If a covariance is positively nonseparable then $R(s,t) < 1$ and if a covariance is negatively nonseparable then $R(s,t) > 1$. Since negative nonseparability is uncommon in practice we consider mostly positively nonseparable functions in what follows. In this case, the more nonseparable the model is, the closer the ratio $R(s,t)$ is to 0. This happens because the covariance function $C(s,t)$ decreases less rapidly in space at larger temporal lags.

In the following sections we further investigate these properties for some models proposed in the literature.

1.1 Example 1

Consider the isotropic nonseparable class proposed in Fonseca and Steel (2011) with covariance function given by

$$C(s,t) = \sigma^2 M_0(-\gamma_1(s) - \gamma_2(t))M_1(-\gamma_1(s))M_2(-\gamma_2(t)), \tag{3}$$

where, for instance, $\gamma_1(s) = ||s/a||^\alpha$ and $\gamma_2(t) = |t/b|^\beta$ with $\alpha \in (0,2]$ and $\beta \in (0,2]$. Let us assume that $M_0(x) = (1-x)^{-\lambda_0}$ and $M_i(x) = (1-x)^{-1}, i = 1,2$. For this class, the degree of nonseparability used in Fonseca and Steel (2011) is given by $c = \lambda_0/(\lambda_0+1)$. This is a benchmark measure that makes sense in this model and is motivated by the construction of the dependence in space and time through mixing. Since the class is isotropic, $C(s,t)$ is a function of $||s||$ and $|t|$ only.

By construction, this model always leads to positive (or zero) nonseparability in the sense that $R(s,t) \leq 1$ for all s and t. In particular, the ratio $R(s,t)$ is a function of 2 arguments $||s||, |t| \in \Re_+$ given by

$$R(s,t) = \frac{M_0(-\gamma_1(s))M_0(-\gamma_2(t))}{M_0(-\gamma_1(s)-\gamma_2(t))} = \left\{ \frac{1+\gamma_1(s)+\gamma_2(t)}{1+\gamma_1(s)+\gamma_2(t)+\gamma_1(s)\gamma_2(t)} \right\}^{\lambda_0}. \qquad (4)$$

Thus, $R(s,t) < 1$ when both $||s||$, $|t|$ and λ_0 are strictly positive. It is one when $\lambda_0 = 0$ (the separable case) and is also one on the margins where either one of $||s||$ or t is zero. Clearly, away from the margins the ratio $R(s,t)$ is a decreasing function of $(||s||, |t|)$ since $\gamma_1(s)$ and $\gamma_2(t)$ increase with $||s||$ and $|t|$.

Figure 1 presents some 3-D graphs of $R(s,t)$ for this class where $c = 0$ indicates separability and c tending to one indicates strong nonseparability. Notice that for $c = 0$ this ratio is constant, as mentioned above. Also, for $c \to 1$ this ratio goes to zero. But how can we measure how much $\tilde{C}(s,t)$ and $C(s,t)$ resemble each other? That is, what measure should we use to reflect the discrepancy of these two functions of space and time?

2 Measuring separability in stationary nonseparable models

From the discussion in the previous section, we conclude that a sensible measure of nonseparability should measure the distance between the surface defined by $R(s,t)$ and the plane surface equal to one in the usable domain $D \times T$. Therefore we define a measure based on the volume contained between these two surfaces. Define

$$B = \int_D \int_T R(s,t) ds dt \qquad (5)$$

Note that in the case of a separable covariance function $B = B_s = \int_D \int_T 1 ds dt$. Then we define

$$v_0 = \frac{B_s - B_{ns}}{B_s} \qquad (6)$$

as a measure of separability such that $0 \leq v_0 \leq 1$. B_{ns} is the integral defined in (5) for the nonseparable model under study. This measure is the volume between the constant surface equal to 1 and the surface defined by $R(s,t) = \tilde{C}(s,t)/C(s,t)$ in the region $D \times T$. We divide by B_s in order to obtain values of v_0 that are always between 0 and 1.

In order to compute v_0 it is, however, necessary to define the regions of integration D and T. In the case of an isotropic covariance function we can write $R(s,t) = R^*(||s||, |t|)$ and $B = \int_0^{h_{10}} \int_0^{h_{20}} R^*(h_1, h_2) dh_1 dh_2$ and depending on the chosen values for (h_{10}, h_{20}) we obtain different values of v_0. In the choice of (h_{10}, h_{20}) we want to cover the region where the covariance is nonnegligible. On the other hand, if we take a very large region for integration, the integrand of B_{ns} will have very small values for large (h_1, h_2) so that B_{ns} will hardly be affected while $B_s = h_{10}h_{20}$ keeps increasing when the region of integration increases. Thus $v_0 = 1 - B_{ns}/h_{10}h_{20}$ will tend to one as we expand the region. Therefore, we need to define a reasonable value for (h_{10}, h_{20}) that is not too large nor too small, for the model in question. Define the correlation function $\rho(s,t) = C(s,t)/C(0,0)$. One way to deal with this issue is to find the minimum value of (h_{10}, h_{20})=$(||s_0||, |t_0|)$ that satisfies $\rho(s_0, 0) \leq \epsilon$ and $\rho(0, t_0) \leq \epsilon$ for a given ϵ. The reason to use the correlation margins instead of $\rho(s_0, t_0)$ is that there are several values of (h_{10}, h_{20}) that leads to $\rho(s_0, t_0) \leq \epsilon$, while with the margins (h_{10}, h_{20}) is uniquely defined by choosing ϵ.

2.1 Example 1 continued

By the construction of this class of models, nonseparability is meaningfully related to the parameter λ_0. As is clear from (4), $R(s,t)$ is a decreasing function of λ_0, that is, as the model becomes more nonseparable (as λ_0 increases) the ratio decreases and tends towards zero.

Let us now verify to what extent the generic measure v_0 in (6) accords with the measure c which is based on the specific construction of the model in this example. In particular, we compare v_0 and c for a grid of values of λ_0 (which defines c). We use numerical integration in R with A.C. Genz's Fortran ADAPT subroutine to do all of the calculations. In the specification of Section 1.1 we set the parameters as $\alpha = 1.5$, $a = 1$, $\beta = 1.5$ $b = 1$. Figures 2(a)-(c) plot the values of v_0 against c for 3 different regions of integration. Each region was found based on the margins $\rho(s_0, 0)$ and $\rho(0, t_0)$

of the separable model ($\lambda_0 = 0$) given a value of ϵ. It seems that v_0 and c give a very similar representation of the degree of separability. Notice that as expected, the values of v_0 will depend on the limits used for integration of the covariance function. In Figure 2(b) we see that it is possible to obtain v_0 very close to c, for the entire range $(0, 1)$. The region of integration increases as ϵ decreases. In Figure 2(a-c) $\epsilon = 0.200$ implies $||s_0|| = |t_0| = 2.41$, $\epsilon = 0.072$ implies $||s_0|| = |t_0| = 5.43$ and $\epsilon = 0.010$ implies $||s_0|| = |t_0| = 21.26$, respectively (in this example, $a = b$ and $\alpha = \beta$).

We perform a sensitivity study where we vary ϵ. As in this model we have one natural measure of nonseparability, c, we aim to choose ϵ that makes v_0 as close as possible to c. Given ϵ (and thus $(||s_0||, |t_0|)$) we compute v_0 for several values of c $(c_1, \ldots, c_r)$. Then we compute the difference $|c_i - v_0^{(i)}|, i = 1, \ldots, r$. The mean of this difference over all considered values of c_i is presented in Figure 3 for different parametrizations of the covariance model. Notice from this picture that the smoother the process is, the smaller ϵ has to be. Notice in Figure 3(b) that changing the value of the range parameters doesn't change the value of ϵ at all. This might be because we find different regions of integration for each combination of scales and these scales are clearly only affecting the range not the nonseparability. This reinforce that the role of parameters in our model are very clear. In Table 1 we present the value of ϵ that minimizes the mean difference $|c - v_0|$ for each configuration of the parameter set used in this study. For the parametrizations considered the values of ϵ that minimized the difference varied from 0.0457 (smoother process, configuration 7) to 0.1881 (rough process, configuration 6). In Figure 4 we present $\epsilon_m = arg\ min\{|c_i - v_0^{(i)}|\}$ for each value of c for some parametrizations of the covariance model (3). The smooth process (configuration 7) is the one that requires the lower values of ϵ for all values of c in the minimization of the difference. It seems that as c increases the value of ϵ has to be larger, except for values of c very close to 1. As a general rule we would recommend the use of $\epsilon = 0.0660$, which is the value that minimizes the difference $|c - v_0|$ for a process that is mean squared continuous.

In this example, we conclude that v_0 provides a useful measure of separability that can be very close to c, which is a measure of separability derived in a completely different manner by the construction of this particular model.

Configuration	α	β	a	b	ϵ with $\min\{\text{mean}\lvert c - v_0\rvert\}$
1	1.5	1.5	1	1	0.0660
2	1.0	1.5	1	1	0.0864
3	2.0	1.5	1	1	0.0559
4	1.5	1.5	2	2	0.0660
5	1.5	1.5	0.5	0.3	0.0660
6	0.5	0.5	1	1	0.1881
7	2	2	1	1	0.0457
8	0.5	2	1	1	0.0864
9	1.5	0.5	1	1	0.1169
10	0.5	1	1	1	0.1474

Table 1: Value of ϵ that minimizes the mean difference for each configuration of the parameters in Example 1.

2.2 Example 2: The model of Gneiting (2002)

Gneiting (2002) proposed a valid space-time covariance model based on complete monotone functions. The model is given by

$$C(s,t) = \sigma^2 \frac{1}{\psi(|t|^2)^{d/2}} \varphi\left(\frac{||s||^2}{\psi(|t|^2)}\right), \tag{7}$$

where $\varphi(.)$ is a complete monotone function ($(-1)^n \varphi^{(n)}(x) \geq 0, \forall x$ and for $n = 0, 1, \ldots$) with $\varphi(0) = 1$ and $\psi(.)$ has a complete monotone derivative with $\psi(0) = 1$. Notice that this class is either separable or positively nonseparable as $R(s,t) = \varphi(||s||^2)/\{\varphi[||s||^2/\psi(|t|^2)]\} \leq 1$. This result follows from the fact that $\varphi(.)$ is a complete monotone function and $\psi(.)$ has a complete monotone derivative and is also mentioned in Rodrigues and Diggle (2010).

One example of this class would be $\varphi(x) = \exp\{-bx^\gamma\}, b > 0, 0, \gamma \leq 1$ and $\psi(x) = (ax^\alpha + 1)^{\beta_G}, a > 0, 0 < \alpha \leq 1, 0 \leq \beta_G \leq 1$. But this model is either nonseparable ($\beta_G > 0$) or separable with the covariance only varying over space (if $\beta_G = 0$ then $C(s,t) = \sigma^2\varphi(||s||^2)$). A related model proposed in Gneiting (2002) is the product of this covariance with the valid temporal covariance $(a|t|^{2\alpha}+1)^{-\delta}$. After reparametrization ($\tau = \delta + \beta_G d/2$) we obtain

$$C(s,t) = \sigma^2 (a|t|^{2\alpha}+1)^{-\tau} \exp\left\{ -\frac{b||s||^{2\gamma}}{(a|t|^{2\alpha}+1)^{\beta_G \gamma}} \right\}, \tag{8}$$

where $\beta_G \in [0,1]$ is a parameter measuring interaction in space-time: $\beta_G = 0$ means separability and the degree of nonseparability increases with β_G. We refer to this model as Gneiting's model 1. Gneiting (2002) suggests to parametrize in terms of τ in order to obtain a more easily interpretable interaction parameter β_G (but see our comments at the end of this subsection). For model (8), $R(s,t) = \exp\left\{ -b||s||^{2\gamma} \left[1 - \frac{1}{(a|t|^{2\alpha}+1)^{\beta_G \gamma}} \right] \right\}$. When $\beta_G = 0$ we have that $R(s,t) = 1$ and as β_G increases towards 1, $R(s,t)$ decreases. Note, however, that $R(s,t)$ does not decrease to zero as the nonseparability parameter $\beta_G \to 1$, unlike the situation in Example 1.

Another model presented in Gneiting (2002) uses $\varphi(t) = (2^{\nu-1}\Gamma(\nu))^{-1}(bt^{1/2})^{\nu} K_\nu(bt^{1/2})$, with K_ν the modified Bessel function of the second kind of order $\nu > 0$ and $\psi(t) = (at^\alpha + 1)^\beta, a > 0, 0 < \alpha \leq 1, 0 \leq \beta \leq 1$, where the resulting covariance is again constructed as the product of this covariance with the valid temporal covariance $(a|t|^{2\alpha}+1)^{-\delta}$. This leads to

$$C(s,t) = \sigma^2 (a|t|^{2\alpha}+1)^{-\tau} (2^{\nu-1}\Gamma(\nu))^{-1} \left(\frac{b||s||}{(a|t|^{2\alpha}+1)^{\beta_G/2}} \right)^{\nu} K_\nu \left(\frac{b||s||}{(a|t|^{2\alpha}+1)^{\beta_G/2}} \right), \tag{9}$$

with $\beta_G \in [0,1]$ again the parameter measuring interaction in space-time. We refer to this model as Gneiting's model 2. If $\nu = 0.5$ we obtain the covariance (8). Notice that for model (9), $R(s,t) = (a|t|^{2\alpha}+1)^{\beta_G/2\nu} K_\nu(b||s||)/K_\nu\left(\frac{b||s||}{(a|t|^{2\alpha}+1)^{\beta_G/2}} \right)$. For $\beta_G = 0$ we have that $R(s,t) = 1$ and $R(s,t)$ decreases as a function of β_G. However, as β_G increases towards 1, $R(s,t)$ again does not tend to zero.

For both Gneiting models, we computed the value of v_0 for this model for several values of β_G. We set $\sigma^2 = 1$, $a = 1$, $\gamma = 0.5$, $b = 1$ and for Model 1 we take $\alpha = 0.8$ and $\tau = (1,2,3)$, whereas for Model 2 we adopt $\alpha = (0.2, 0.5, 0.8)$ and $\nu = (0.2, 2, 4)$. The values of v_0 are presented in Figures 5 (Model 1) and 7 (Model 2) for 3 values of ϵ. Notice that in this case, the margins $\rho(s_0, 0)$ and $\rho(0, t_0)$ do not depend on β_G. Thus given a value of ϵ the region of integration is the same for different values of β_G.

The effect of $R(s,t) > 0$ even for $\beta_G = 1$ is obvious for both Gneiting models. The clear consequence is that B_{ns} in (5) does not tend to zero and $v_0 = 1 - B_{ns}/h_{10}h_{20}$ does not tend to one as β_G grows. In other words, even for very large values of β_G we do not obtain values of v_0 close to one unless ϵ is very small (and thus h_{10} and h_{20} are very large). Rather, for β_G varying in a very large range, the degree of nonseparability (as measured by v_0) varies very little. This means that very different values of β_G generate very similar degrees of nonseparability. In addition, the achievable range of nonseparability depends on other parameters in the model. In particular, for Model 1 the parameter τ greatly influences the degree of nonseparability as illustrated in Figure 5. If we parametrize in terms of the original parameter δ (*i.e.* avoid the reparametrization implicit in (8)), then we obtain the situation described in Figure 6. Interestingly, the influence of δ on v_0 is pretty close to that of τ in the other parametrization, with slightly less influence for the (perhaps more relevant) case with $\epsilon = 0.001$ where the model can cover a reasonable range of nonseparability. Thus, the use of our criterion can give us a benchmark for comparing parametrizations in terms of how clearly defined the role of any separability parameter is.

2.3 Example 3: The model of Cressie and Huang (1999)

Consider example 3 of Cressie and Huang (1999) where the covariance function is valid and positively nonseparable and given by

$$C(s,t) = \sigma^2 \frac{b|t|^2 + 1}{((b|t|^2 + 1)^2 + a||s||^2)^{(d+1)/2}}, \tag{10}$$

$s \in \Re^d$, $t \in \Re$. This model does not have a single parameter that is responsible for the degree of nonseparability in the model. What kind of nonseparability is generated by the parameters (a, b)? We consider $d = 2$, set $\sigma^2 = 1$ and vary (a, b) in order to assess the degree of nonseparability generated by this model according to our proposed measure. In this situation we find the region of integration based on the margins $\rho(s_0, 0)$ and $\rho(0, t_0)$ of the parametrization under study, that will depend on (a, b) (using $\epsilon = 0.2, 0.066, 0.005$). The values of v_0 for various values of a and b and ϵ are presented in Figure 8. (a, b) are clearly just decay parameters. Given a value of ϵ or a given region of integration, it is possible to compute v_0 and check which is the achieved degree of nonseparability implied by (a, b). This might be a way of comparing the degree of nonseparability achieved by several different models.

In this model, the ranges in space and time are confounded with the degree of nonseparability since the parameters determine both range and the degree of nonseparability at the same time. In conclusion, the parameters of this model are not easily interpretable in terms of nonseparability.

2.4 Example 4: The model of Rodrigues and Diggle (2009)

Rodrigues and Diggle (2010) propose a valid nonseparable convolution-based model given by

$$C(s,t) = \sigma^2 \int \frac{\rho(h-t)\rho(h)}{\rho(h-t)^\lambda + \rho(h)^\lambda} \exp\left\{ -\frac{\rho(h-t)^\lambda \rho(h)^\lambda}{\rho(h-t)^\lambda + \rho(h)^\lambda} (||s||/\tau)^2 \right\} dh. \tag{11}$$

For $\lambda = 0$ the covariance is separable, and positive nonseparability is generated by $\lambda > 0$ whereas negative nonseparability is possible with these models, corresponding to negative values of λ.

Negative nonseparability implies that $R(s,t) > 1$ and thus $B_{ns} > h_{10}h_{20}$. As v_0 just represents the standardized volume between $R(s,t)$ for the nonseparable case and the constant surface equal to 1, we will define $v_0 = (B_{ns} - B_s)/B_{ns}$ in case of negative nonseparability.

We compute the value of v_0 using $\rho(d) = \exp\{-|d|/\phi\}$, $\phi = (0.5, 1, 2)$, $\sigma^2 = 1$ and $\tau = (\sqrt{0.1}, \sqrt{0.5}, \sqrt{2})$. The values of v_0 for different parametrizations are plotted in Figure 9. When λ is negative, v_0 is monotonously increasing in $|\lambda|$ and the entire range of nonseparability is achieved. However, there is no monotonous relation between positive λ and v_0 since the degree of nonseparability decreases with λ for large positive values of λ. Thus, different λ can lead to the same v_0 and also the achievable range of positive nonseparability is quite restricted (despite the small value of ϵ). This renders the parameter λ a rather unreliable measure of separability (with only negative values conveying a clear meaning). In addition, the effect of the other two covariance parameters on nonseparability is substantial and very nonlinear (especially for τ).

3 Measuring separability in nonstationary nonseparable models

Now we extend this idea to nonstationary nonseparable covariance models. Consider the case of nonstationarity in space. Let $\{Z(s,t); s \in D \subset \Re^d, t \in T \subset \Re_+\}$ be a spatiotemporal random field with covariance function $Cov(Z(s_1,t_1),Z(s_2,t_2)) = C(s_1,s_2,t)$, where $t = t_1 - t_2$, $s_1, s_2 \in D, t_1, t_2 \in T$. If $C(s_1,s_2,t)$ is separable then it can be written as $C_1(s_1,s_2)C_2(t)$, where $C_1(s_1,s_2)$ and $C_2(t)$ are purely spatial and temporal covariance functions respectively.

In terms of the joint covariance function, a separable fully symmetric nonstationary covariance function can be expressed as

$$C(s_1,s_2,t) = \frac{C(s_1,s_2,0)C(s_1,s_1,t)}{C(s_1,s_1,0)}, \tag{12}$$

where $s_1, s_2 \in D$, $t \in T$ are spatial locations and a temporal lag, respectively. Notice that taking s_2 as the "common" location, rather than s_1 does not affect matters. This follows from the fact that if $C(s_1,s_2,t)$ is separable then $C(s_1,s_2,t) = C_1(s_1,s_2)C_2(t)$.

We can then define $\tilde{C}(s_1,s_2,t)$ as the right hand side of (12) so that if the spatiotemporal process is separable then $\tilde{C}(s_1,s_2,t) = C(s_1,s_2,t)$. Define

$$B = \int_D \int_D \int_T \frac{\tilde{C}(s_1,s_2,t)}{C(s_1,s_2,t)} dt ds_1 ds_2, \tag{13}$$

which is now a function of 3 variables (s_1,s_2,t). Then

$$v_0 = \frac{B_s - B_{ns}}{B_s} \tag{14}$$

is a measure of separability such that $0 \leq v_0 \leq 1$. Here $B_s = \int_D \int_D \int_T 1 ds_1 ds_2 dt$ and B_{ns} is the integral defined in (13) for the nonseparable model under study.

3.1 Example 5

Consider the nonstationary and nonseparable model given by

$$C(s,s',t) = \int_D K(s-w)K(s'-w)\sigma^2(w)M_0(-\gamma_1(s-s';w)-\gamma_2(t))M_1(-\gamma_1(s-s';w))M_2(-\gamma_2(t))dw, \tag{15}$$

where $K(.)$ is a convolution kernel and we mix through $w \in D$. This is a nonstationary version of the model by Fonseca and Steel (2011) in Example 1, where we use continuous mixtures of locally stationary processes, as in Fuentes et al. (2005). If the kernel $K(.)$ decreases rapidly and the parameters $\sigma^2(w)$ and $\gamma_1(.;w)$ vary slowly then we have local stationarity. However, these parameters are allowed to vary across the whole spatial domain resulting in a nonstationary process. For instance, here we define $\gamma_1(d;w) =$

$||d/a(w)||^{\alpha(w)}$. Using (12), we obtain

$$\frac{\tilde{C}(s,s',t)}{C(s,s',t)} = \frac{\int_D K(s-w)K(s'-w)\sigma^2(w)M_1(-\gamma_1(s-s';w))M_0(-\gamma_1(s-s';w))M_0(-\gamma_2(t))dw}{\int_D K(s-w)K(s'-w)\sigma^2(w)M_1(-\gamma_1(s-s';w))M_0(-\gamma_1(s-s';w)-\gamma_2(t))dw}. \tag{16}$$

In practice, the covariance function in (15) can be approximated on a grid by

$$C(s,s',t) = \sum_{i=1}^{m} K(s-w_i)K(s'-w_i)\sigma^2(w_i)M_0(-\gamma_1(s-s';w_i)-\gamma_2(t))M_1(-\gamma_1(s-s';w_i))M_2(-\gamma_2(t)) \tag{17}$$

We consider the kernel function $K(x) = \frac{2}{\pi h^2}(1-||x/h||)_+$ where h is the bandwidth, which will be chosen depending on the grid of values for $w \in D$. As in the stationary case, the value of v_0 will depend on the range chosen for integration. We perform a sensitivity study to identify a reasonable value of ϵ such that the difference (when compared to c, which is still a natural separability measure in this model) is as small as possible. In this example, the parameters that depend on the spatial location are defined as

$$\sigma^2(w) = \sigma_0^2 + \sigma_1^2 w_1,\ a(w) = a_0 + a_1 w_2,\ \alpha(w) = \alpha_0 + \alpha_1 w_1. \tag{18}$$

We take $\gamma_2(t) = |t|^\beta$ and use $\lambda_0 = (0.0, 0.1, 0.2, 0.5, 1, 2, 5, 10, 100)'$ implying $c = (0.00, 0.09, 0.17, 0.33, 0.50, 0.67, 0.83, 0.91, 0.99)'$.

We would expect the construction argument to carry over from the stationary model in Example 1 so that separability is well measured by c. However, the nonstationarity induced by mixing over processes with parameters that may vary with location requires us to investigate whether it is again close to the general nonseparability measure v_0. Figure 10 (a-d) presents the values of v_0 for different values of $a(w) = a_0 + a_1 w$ and for $\sigma^2(w) = 0.5$ and $\alpha(w) = 1.5$ and bandwidth $h = 3$. The region of integration used in the four configurations was the same and obtained using $\epsilon = 0.03$ for the (stationary) configuration of Figure 10(a). With bandwidth $h = 5$ the results were essentially the same. For Figure 10(a) we have a stationary process for which v_0 is almost identical to c. The first impression would be that v_0 and c become more disparate when the model becomes more nonstationary but this can be counteracted. In Figure 11 we have the same covariance as in Figure 10 (b), but now we vary the region of integration using different values of ϵ. Throughout, values of ϵ refer to the stationary case. We are able to recover a close fit by increasing the region of integration ($\epsilon = 0.01$).

The results for several configurations of this model for the mean absolute difference between c and v_0 are presented in Figure 12. The values of ϵ that minimize this difference usually vary between 0.01 and 0.04. As a benchmark we would suggest the use of ϵ around 0.03, which gives resonable results for most of the configurations studied here. This benchmark might, of course, be sensitive to the specific model chosen.

4 Conclusions

In this paper we presented a general measure of nonseparability in space and time that can be easily implemented and used for any nonseparable (possibly nonstationary) covariance model. The proposed measure is a sensible measure of nonseparability based on the intuitive concept of a direct comparison between separable and nonseparable covariance structures. It behaves as expected for situations where another natural separability measure is available (Examples 1 and 5), but is generally applicable. This is useful for classes where there is no specific parameter responsible for measuring separability (as in Example 3) or when there is such a parameter but its interpretation is not clear (as in Examples 4). It also clarifies situations where the available range of nonseparability is limited (as in Examples 2-4). The measure also usefully flags up situations where several parameters are important in determining the degree of nonseparability as we illustrated with Gneiting's model 1 in Example 2. Finally, it can provide a way of comparing different parametrizations of covariance structures in terms of the role and interpretability of the parameters driving nonseparability.

References

Apanasovich, T. V. and Genton, M. G. (2010). "Cross-covariance functions for multivariate random fields based on latent dimensions." *Biometrika*, 17, 15–30.

Cressie, N. and Huang, H.-C. (1999). "Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions." *Journal of the American Statistical Association*, 94, 448, 1330–1340.

Crujeiras, R. M., Fernández-Casal, R., and González-Manteiga, W. (2010). "Nonparametric Tests for Separability of Spatio-Temporal Processes." *Environmetrics*, 21, 382–399.

De Iaco, S. and Posa, D. (2013). "Positive and negative non-separability for space–time covariance models." *Journal of Statistical Planning and Inference*, , 143, 378–391.

Fonseca, T. C. O. and Steel, M. F. J. (2011). "A General Class of Nonseparable Space-time Covariance Models." *Environmetrics*, 22, 2, 224–242.

Fuentes, M. (2006). "Testing for Separability of Spatial-Temporal Covariance Functions." *Journal of Statistical Planning and Inference*, 136, 447–466.

Fuentes, M., C, L., Davis, J. M., and Lackmann, G. M. (2005). "Modeling and Predicting Complex Space-Time Structures and Patterns of Coastal Wind Fields." *Environmetrics*, 16, 449–464.

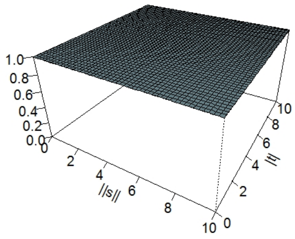
Genton, M. G. (2007). "Separable approximations of space-time covariance matrices." *Environmetrics*, , 18, 681–695.

Gneiting, T. (2002). "Nonseparable, Stationary Covariance Functions for Space-Time Data." *Journal of the American Statistical Association*, 97, 458, 590–600.

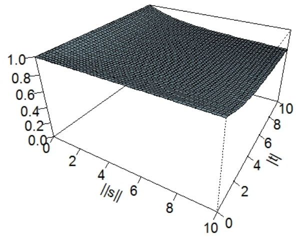
Li, B., Genton, M. G., and Sherman, M. (2007). "A Nonparametric Assessment of Properties of Space-Time Covariance Functions." *Journal of the American Statistical Association*, 102, 736–744.

Mitchell, M. W., Genton, M. G., and Gumpertz, M. L. (2005). "Testing for separability of space-time covariances." *Environmetrics*, 16, 819–831.

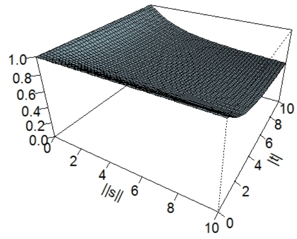
Rodrigues, A. and Diggle, P. J. (2010). "A class of convolution-based models for spatio-temporal processes with non-separable covariance structure." *Scandinavian Journal of Statistics*, 37, 4, 553–567.



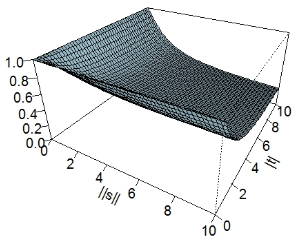
(a) $c = 0$ (separable model).



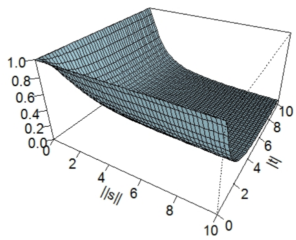
(b) $c = 0.09$.



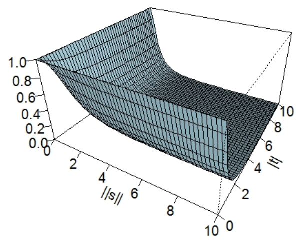
(c) $c = 0.17$.



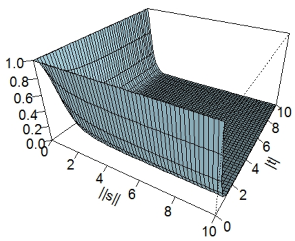
(d) $c = 0.33$.



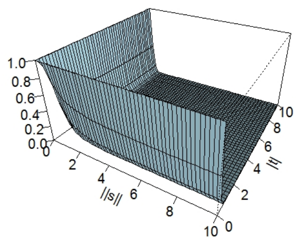
(e) $c = 0.50$.



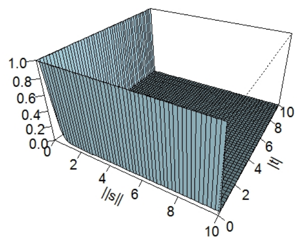
(f) $c = 0.67$.



(g) $c = 0.83$.

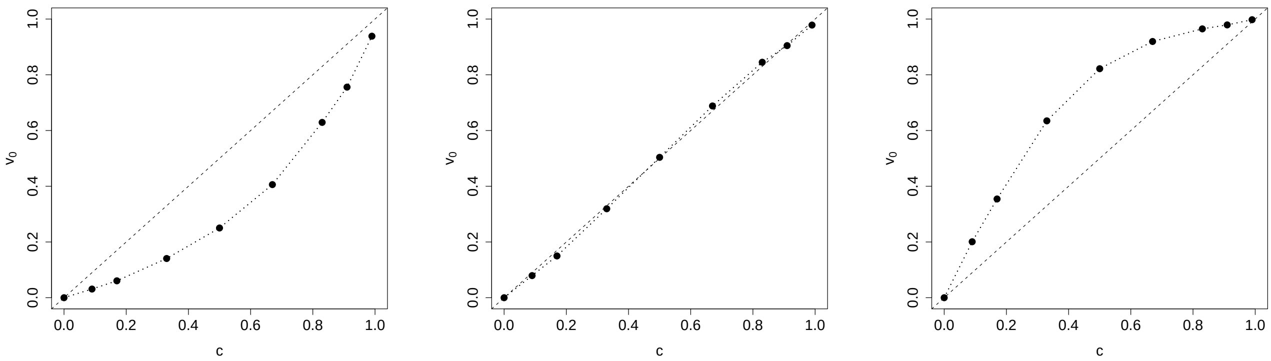


(h) $c = 0.91$.



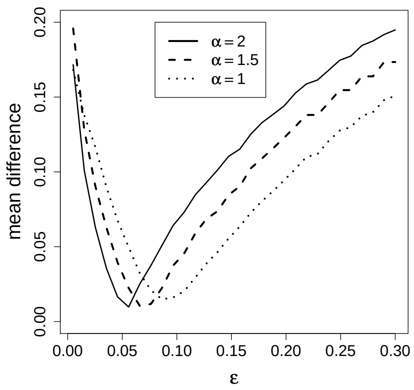
(i) $c = 0.99$.

Figure 1: 3-D graph of $\tilde{C}(s,t)/C(s,t)$ for different values of the nonseparability parameter c in Example 1.

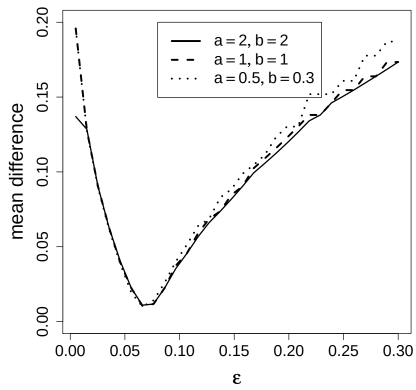


(a) $\epsilon = 0.200$.   (b) $\epsilon = 0.072$.   (c) $\epsilon = 0.010$.

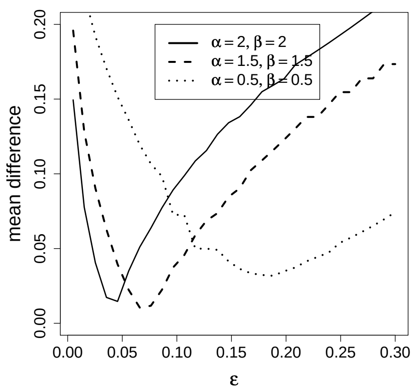
Figure 2: Proposed measure of separability (v_0) against the measure of separability (c) derived from the model in Example 1. The straight line represents $c = v_0$.



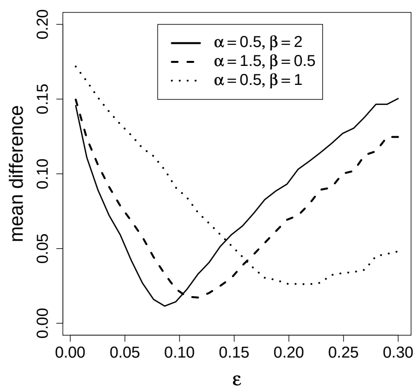
(a) Varying α, $a = b = 1$ and $\beta = 1.5$.



(b) Varying a, $\alpha = \beta = 1.5$.



(c) Varying α and β, $a = b = 1$.



(d) Varying α and β, $a = b = 1$.

Figure 3: ϵ against mean difference $|c - v_0|$ for several parametrisations of the covariance model (3). For all configurations $\sigma^2 = 1$.

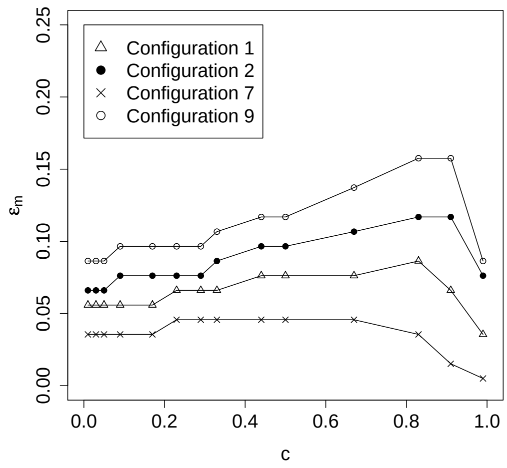
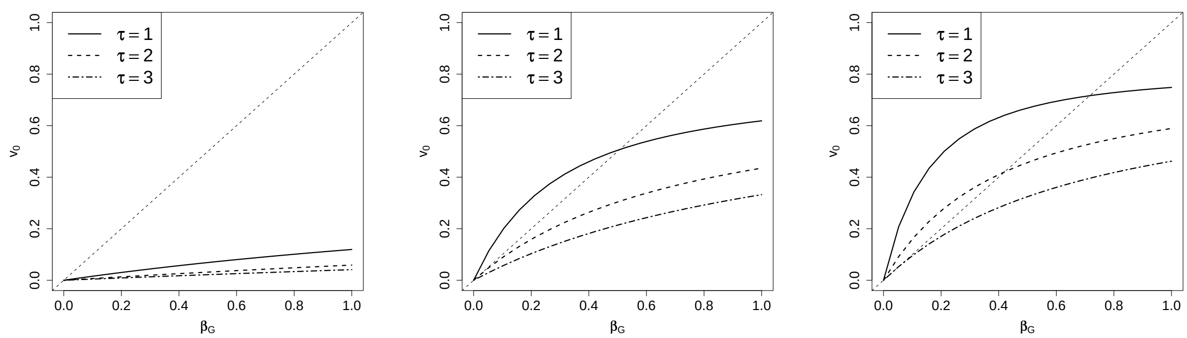
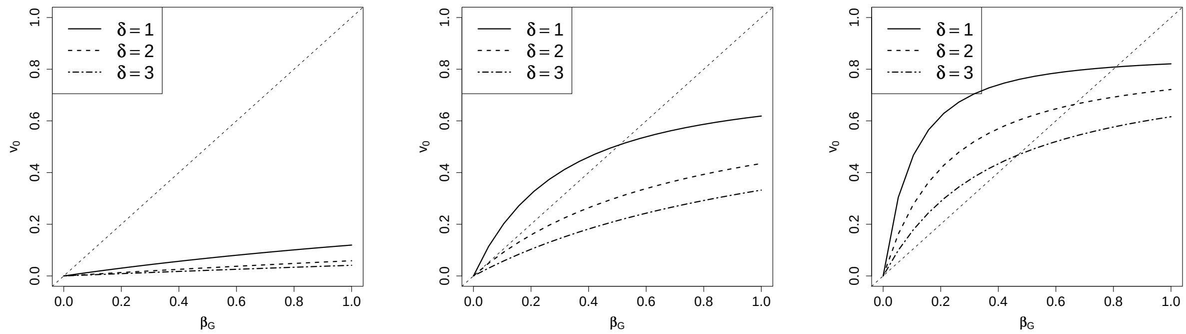


Figure 4: $\epsilon_m = arg\ min\{|c - v_0|\}$ against c for several parametrisations of the covariance model (3).



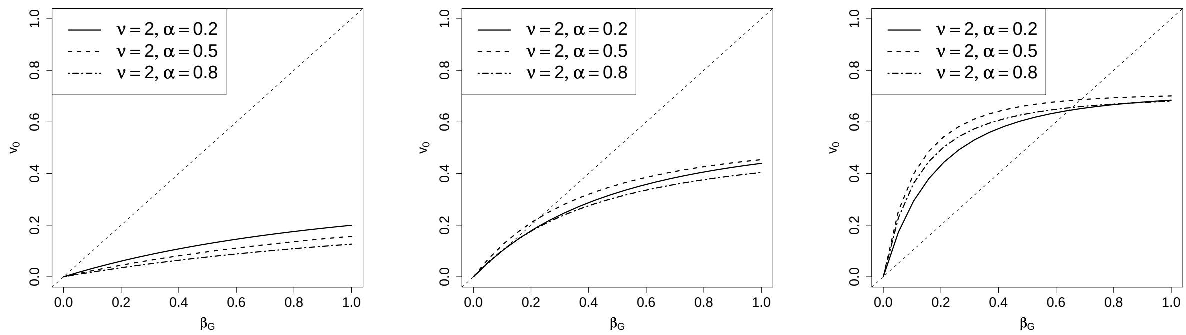
(a) $\epsilon = 0.30$.   (b) $\epsilon = 0.02$.   (d) $\epsilon = 0.001$.

Figure 5: Proposed measure of separability (v_0) against the separability parameter in Gneiting's model 1 in (8). The straight line represents $\beta_G = v_0$.

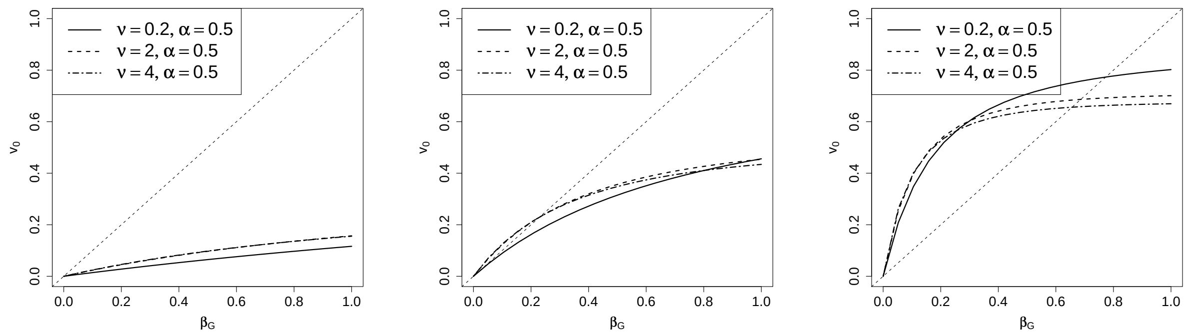


(a) $\epsilon = 0.30$.   (b) $\epsilon = 0.02$.   (d) $\epsilon = 0.001$.

Figure 6: Proposed measure of separability (v_0) against the separability parameter in Gneiting's model 1 parameterized in terms of δ. The straight line represents $\beta_G = v_0$.

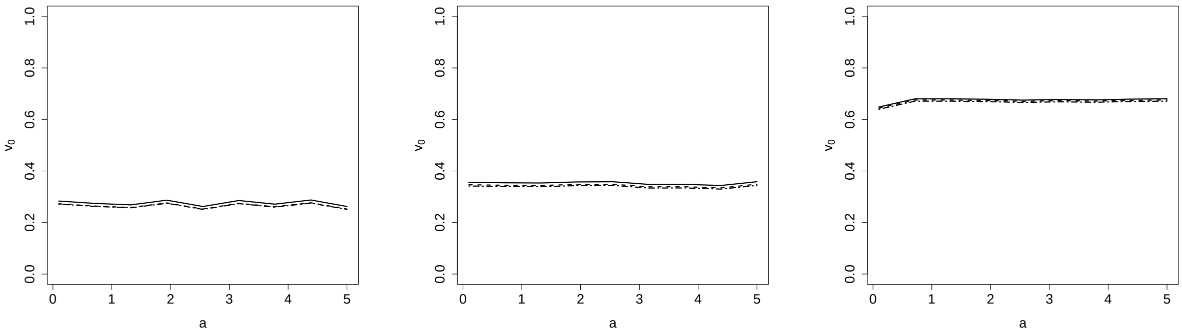


(a) $\epsilon = 0.30$.   (b) $\epsilon = 0.07$.   (c) $\epsilon = 0.005$.



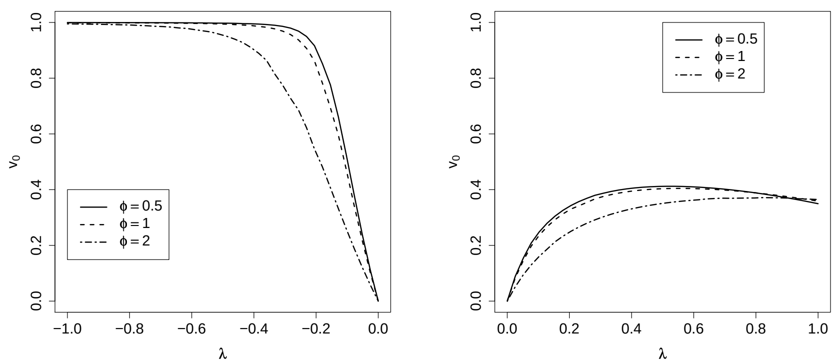
(d) $\epsilon = 0.30$.   (e) $\epsilon = 0.07$.   (f) $\epsilon = 0.005$.

Figure 7: Proposed measure of separability (v_0) against the separability parameter in Gneiting's model 2 in (9). The straight line represents $\beta_G = v_0$.

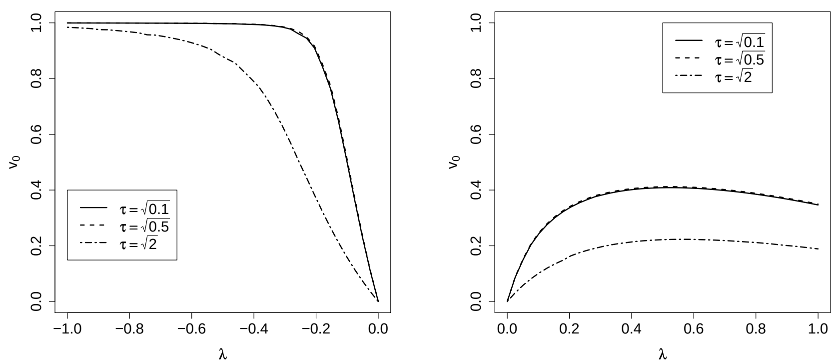


(a) $\epsilon = 0.2$ (b) $\epsilon = 0.066$ (c) $\epsilon = 0.005$

Figure 8: Proposed measure of separability (v_0) for different configurations of parameters (a, b) in model (10).

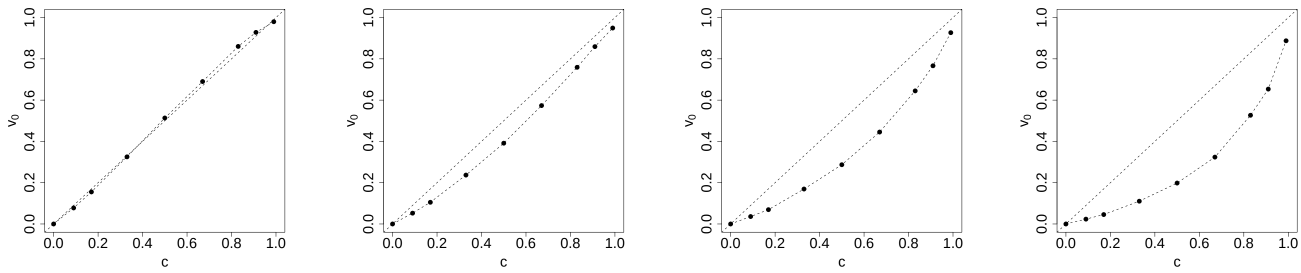


(a) Negative λ, $\tau = \sqrt{0.5}$ (b) Positive λ, $\tau = \sqrt{0.5}$



(c) Negative λ, $\phi = 0.5$ (d) Positive λ, $\phi = 0.5$

Figure 9: Proposed measure of separability (v_0) against the separability parameter λ in model (11) using $\epsilon = 0.001$ and $\rho(d) = \exp\{-|d|/\phi\}$.



(a) $a(w) = 0.5 + 0.0w$. (b) $a(w) = 0.5 + 0.2w$. (c) $a(w) = 0.5 + 0.5w$. (d) $a(w) = 0.5 + 1w$.

Figure 10: Proposed measure of separability (v_0) against c for different values of the range parameter $a(w)$, $w \in (0, 5]$ in Example 5. Throughout $\sigma^2(w) = 0.5$ and $\alpha(w) = 1.5$. The straight line represents $c = v_0$. The region of integration is defined by $\epsilon = 0.03$ in the model of case (a).

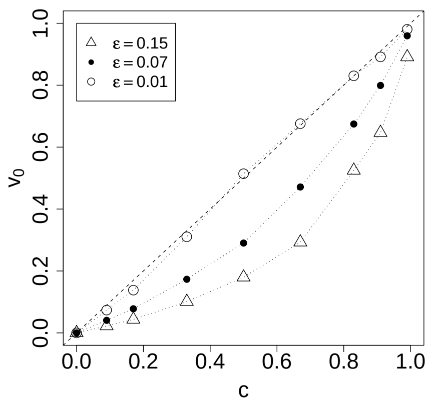
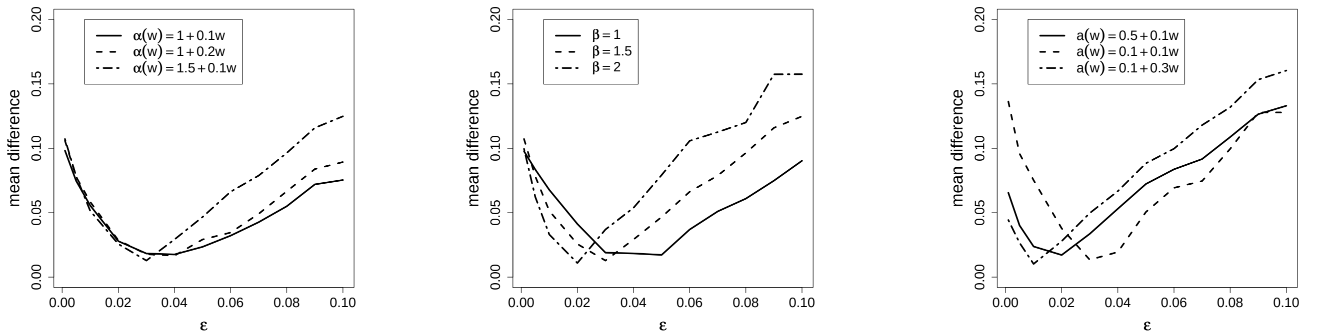


Figure 11: Proposed measure of separability (v_0) against c for different regions of integration when $a(w) = 0.5 + 0.2w$ in Example 5 with $\sigma^2(w) = 0.5$ and $\alpha(w) = 1.5$. The straight line represents $c = v_0$. Values of ϵ refer to the stationary model of Figure 10(a).



(a) $\beta = 1.5$, $a(w) = 0.5$. (b) $\alpha(w) = 1.5 + 0.1w$, $a(w) = 0.5$. (c) $\alpha(w) = 1.5 + 0.1w$, $\beta = 1.5$.

Figure 12: ϵ against mean difference $|c - v_0|$ for several parametrisations of the covariance model (15). Values of ϵ were obtained by letting $C(s_1, 0, 0) < \epsilon$, $C(0, s_2, 0) < \epsilon$ and $C(0, 0, t) < \epsilon$.