Supplementary Materials for "Comparing two spatial variables with the probability of agreement"

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

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

Additional contents for this article are available online. In Web Appendix A we provide the proofs that were omitted in the main manuscript. In Web Appendix B we introduced some examples that illustrates the theoretical results given in the paper. In Web Appendix C we described simulations of a bivariate Gaussian process, simulations of spatiotemporal processes and additional material for the application discussed in the paper. Some additional references are given at the end this manuscript.

Web Appendix A Proofs

Proof of Theorem 2 Note that $\sigma'_D(h) = -\rho_{XY}\sigma_X\sigma_Y M'(h, \nu_{XY}, a_{XY})/\sigma_D(h)$. Also note that $\rho_{XY} \geq 0$, $\sigma_X > 0$, $\sigma_Y > 0$ and $\sigma_D(h) > 0$ for all $h > 0$. Thus $\sigma'_D(h) > 0$ if and only if $M'(h, \nu_{XY}, a_{XY}) < 0$. Without loss of generality, we assume that $a = 1$ and $\nu = \nu_{XY}$ in (7) Noticing that the terms $M'(h, \nu, 1)$ and $g'_\nu(h)$ have the same sign, where $g_\nu(h) = h^\nu K_\nu(h)$, and using the properties of the modified Bessel functions of the second kind (Lebedev, 1965, p. 110), we have that $g'_\nu(h) = -h^\nu K_{\nu-1}(h)$. Since $K_\alpha(x) = K_{-\alpha}(x)$ (Lebedev, 1965, p. 110), $K_\alpha(x) > 0$, for all $x > 0$ and $\alpha \in \mathbb{R}$ (Lebedev, 1965, p. 136), it follows that $g'_\nu(h) < 0$, and the proof is complete. □

Lemma 1. *The function*

$$\mathcal{GW}(h; \kappa, \mu) = \begin{cases} \dfrac{1}{B(2\kappa, \mu+1)} \displaystyle\int_h^1 u(u^2 - h^2)^{\kappa-1}(1-u)^\mu du, & \text{if } 0 \leq h < 1, \\ 0, & \text{if } h \geq 1, \end{cases}$$

is decreasing in h for all $\kappa \geq 0$.

Proof of Lemma 1 Let $h_1 > 0$ and $h_2 > 0$ be such that $0 \leq h_1 \leq h_2$. If $h_1 \geq 1$, then $\mathcal{GW}(h_1;\kappa,\mu) = \mathcal{GW}(h_2;\kappa,\mu) = 0$ and $\mathcal{GW}(\cdot;\kappa,\mu)$ is a monotone function . If $h_1 < 1$ and $h_2 \geq 1$, then $\mathcal{GW}(h_1;\kappa,\mu) \geq 0$, and $\mathcal{GW}(h_2;\kappa,\mu) = 0$, then $\mathcal{GW}(\cdot;\kappa,\mu)$ is a decreasing monotone function. If $h_2 < 1$, we distinguish the following two cases:

- For $\kappa = 0$, note that $0 \leq 1 - h_2^2 \leq 1 - h_1^2 < 1$ and $0 \leq (1 - h_2^2)^\mu \leq (1 - h_1^2)^\mu < 1$, therefore $\mathcal{GW}(h_1;\kappa,\mu) \geq \mathcal{GW}(h_2;\kappa,\mu)$.
- For $\kappa \geq 1$, we define $g(u,h) = u(u^2-h^2)^{\kappa-1}(1-u)^\mu / B(2\kappa,\mu+1)$. Clearly, $g(u,h) \geq 0$ for $0 \leq h < u < 1$, then $G(h) = \int_h^1 g(u,h)du$ corresponds to $\mathcal{GW}(h;\kappa,\mu)$. Hence, by Leibniz's formulae,

$$G'(h) = \int_h^1 \frac{\partial g(u,h)}{\partial h} du - g(h,h) = -2h(\kappa-1)\int_h^1 \frac{g(u,h)}{u^2-h^2} du.$$

  Because $h > 0$, $G'(h) < 0$ if and only if $\kappa > 1$. When $\kappa = 1$, $g(u,h) := \widetilde{g}(u) \geq 0$ and $G'(h) = -\widetilde{g}(h) \leq 0$.

Therefore $\mathcal{GW}(\cdot;\kappa,\mu)$ is a decreasing monotone function for $h_2 < 1$. □

Proof of Theorem 3 Without loss of generality, we assume $b_{12} = 1$ and note that $\sigma_D^2(h)$ is an increasing function of h if and only if $\mathcal{GW}(h;\kappa,\mu)$ is a decreasing function in h for all κ. Therefore, the result holds by Lemma 1, since $\nu + \gamma_{12} + 1 > 0$. □

Web Appendix B Examples

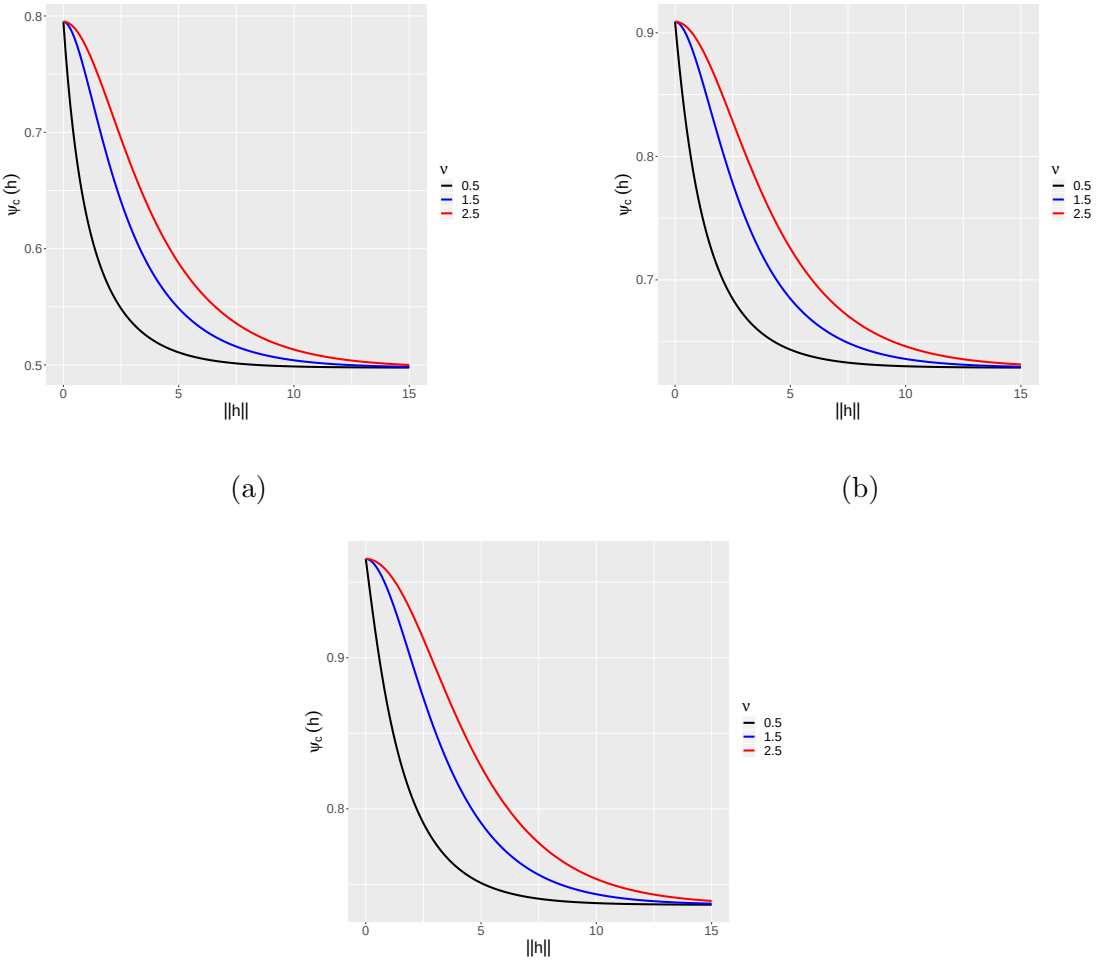
Example 1: We write $M(\cdot,\cdot,\cdot)$ in (10) as $M(h, m+\frac{1}{2}, 1) = M(h) = \exp(-h)P_m(h)$, where $P_m(x) = \sum_{n=0}^m a_n x^n$, with $a_n = (2^n/n!)\binom{m}{n}/\binom{2m}{n}$ (Acosta and Vallejos, 2018). Then, $M'(h) = \{P_m'(h) - P_m(h)\}\exp(-h)$. Consequently, solving the equation $M'(h) = 0$ is equivalent to solve $P_m'(h) = P_m(h)$. Since $P_m'(h) = a_1 + 2a_2h + \cdots + ma_m h^{m-1}$, we have that $Q(h) := P_m'(h) - P_m(h) = \sum_{n=0}^m b_n h^n$, where $b_n = (n+1)a_{n+1} - a_n, n = 0, \cdots, m-1$, and $b_m = -a_m$. Consider

$$a_n = \frac{2^n m!(2m-n)!}{n!(m-n)!}, \quad n = 0, 1, \ldots, m-1; \qquad a_m = \frac{2^m m!}{(2m)!} > 0.$$

Note that

$$b_n = \frac{-2^n m!(2m-n-1)!)}{n!(m-n-1)!(2m)!}\left(\frac{2m-n}{m-n} - 2\right) = \frac{-2^n m!(2m-n-1)!)n}{n!(m-n-1)!(2m)!(m-n)} < 0,$$

for $n = 0, \cdots, m-1$, and $m \geq 1$. Thus, $b_0 = 0$ and $b_n < 0$, for $n = 1, \cdots, m$, and $m \geq 1$. It then follows that $M'(h) = 0 \iff h = 0$. Moreover, $M'(h) < 0$ for $h > 0$, because $M'(h) = Q(h)e^{-h}$, and $Q(\cdot)$ is a polynomial with negative coefficients. Therefore, we have explicitly shown that $\psi_c(h)$ is a decreasing function of h. □



(a)

(b)

(c)

Figure S1: $\psi_c(\|\boldsymbol{h}\|)$ versus $\|\boldsymbol{h}\|$. (a) $c = 1.5$; (b) $c = 2$; (c) $c = 2.5$.

Web Appendix C Numerical Experiments

Web Appendix C.1 Bivariate Simulations

Consider a Gaussian bivariate random field with mean vector $(\mu_1, \mu_2)^\top = (0, 0)^\top$ and Matérn Cross-Covariance function (Gneiting, 2010) with the following two set of parameters (following Gneiting, 2010): (a) separable case: $\sigma_X^2 = \sigma_Y^2 = 1$, $\rho_{XY} = 0.5$, $\nu_X = \nu_Y = \nu_{XY} = \nu = 0.5$, and $a_X = a_Y = a_{XY} = a = 0.2$; (b) non-separable case $\sigma_X^2 = \sigma_Y^2 = 1$, $\rho_{XY} = 0.5$, $\nu_X = 1.5$, $\nu_Y = 0.5$, $\nu_{XY} = 1.0$, $a_X = 1$, $a_Y = 0.2$, and $a_{XY} = 0.6$. The case (a) is know as the parsimonious bivariate Matérn model, and (b) as

the full bivariate Matérn model. As in Gnieiting (2010) we simulate in a grid of size 20×20, for which each of 1000 simulations were generated from both a correct and a misspecified model, i.e., the parameter in the separable case was estimated under a separable and a non-separable scheme for the covariance function, and similarly for the non-separable case. To highlight the effect of covariance estimates, the means of the process were not estimated, and the smoothness parameter, ν's, were also fixed.

Regarding the range parameters, a's, we estimated $\phi = 1/a$ because this is the parameterization programmed in the GeoFit function of the R package GeoModels (Bevilacqua et al., 2023), which we used in this simulation study. Table S1 shows the average, standard deviation, and root mean square error (RMSE) of the estimated parameters for the separable model when the data are simulated from a separable and a nonseparable model, respectively. In addition, Table S2 shows the same information as in Table S1 but for the estimated parameters for a nonseparable model when the data are simulated from a nonseparable and a separable model, respectively. In both cases, the estimates for the well-specified model perform better in terms of bias and variance. However, when the data come from a separable model and we estimate assuming a non-separable model, this misspecification causes greater uncertainty in the estimations. In all cases, only the simulations that include the estimation of the variance of the parameters were left, since it is essential for the hypothesis testing described below.

	Well-specified				Misspecified			
	ρ_{XY}	ϕ	σ^2_X	σ^2_Y	ρ_{XY}	ϕ	σ^2_X	σ^2_Y
mean	0.4982	5.0452	1.0208	1.0211	0.4866	4.8462	1.1632	0.9911
sd	0.0188	0.7377	0.1503	0.1484	0.0313	0.8473	0.2671	0.1695
RMSE	0.0189	0.7387	0.1516	0.1498	0.0341	0.8605	0.3128	0.1696

Table S1: Simulation mean, standard deviation, and root mean square error (RMSE) of estimations for separable model when the data are simulated from a separable and a nonseparable model, respectively.

	Well-specified						Misspecified					
	ρ_{XY}	ϕ_X	ϕ_Y	ϕ_{XY}	σ^2_X	σ^2_Y	ρ_{XY}	ϕ_X	ϕ_Y	ϕ_{XY}	σ^2_X	σ^2_Y
mean	0.497	1.001	1.657	5.167	1.010	1.031	0.414	0.836	1.224	4.641	0.690	0.957
sd	0.077	0.069	0.243	1.700	0.152	0.320	0.073	0.097	0.196	1.534	0.169	0.305
RMSE	0.077	0.069	0.243	1.708	0.152	0.321	0.112	0.190	0.484	1.575	0.352	0.308

Table S2: Simulation mean, standard deviation, and root mean square error (RMSE) of estimations for a nonseparable model when the data are simulated from a nonseparable and a separable model, respectively.

For the hypothesis testing problem $\mathrm{H}_0 : \ \psi_c(\|\boldsymbol{h}\|) \geq 0.95$ versus $\mathrm{H}_1 : \ \psi_c(\|\boldsymbol{h}\|) < 0.95$

the critical region of size α is

$$\mathcal{R}_\alpha = \left\{ \widehat{\psi}_c(\boldsymbol{h}) : \ \widehat{\psi}_c(\boldsymbol{h}) < 0.95 - z_\alpha \sqrt{\widehat{\text{var}(\widehat{\psi}_c(\boldsymbol{h}))}} \right\},$$

where z_α is the upper quantile of the standard normal distribution, and

$$\begin{aligned} \widehat{\text{var}(\widehat{\psi}_c(\boldsymbol{h}))} &= \frac{2}{\pi} \exp\left\{ -\frac{c^2}{\sigma_D^2(\boldsymbol{h}; \widehat{\boldsymbol{\theta}})} \right\} \left\{ \frac{c^2}{\sigma_D^2(\boldsymbol{h}; \widehat{\boldsymbol{\theta}})} \widehat{V_{\sigma_D}(\boldsymbol{h})} \right\}, \\ \widehat{V_{\sigma_D}(\boldsymbol{h})} &= \frac{1}{4\sigma_D^2(\boldsymbol{h}; \widehat{\boldsymbol{\theta}})} \nabla \sigma_D^2(\boldsymbol{h}; \widehat{\boldsymbol{\theta}})^\top \boldsymbol{V}_{\widehat{\boldsymbol{\theta}}} \nabla \sigma_D^2(\boldsymbol{h}; \widehat{\boldsymbol{\theta}}). \end{aligned}$$

Because $\mu_D = 0$ was not estimated, $V_{\mu_D} = 0$, and $\boldsymbol{V}_{\widehat{\boldsymbol{\theta}}}$ was obtained from the inverse of the observed Fisher information matrix. Denote by $\boldsymbol{\theta}_0$ the true parameter vector of the cross-covariance function, then in both parametrizations when $c = z_{0.025} \approx 1.96$ and $\boldsymbol{h} = \boldsymbol{0}$, we have that $\psi_c(\boldsymbol{0}, \boldsymbol{\theta}_0) = 0.95$.

To evaluate the rate of rejection of the hypothesis test defined above, we vary $\|\boldsymbol{h}\|$ and c. Indeed, $\|\boldsymbol{h}\|$ varies from 0 to 30 with step 0.1, and c from 1 to 2 with step 0.01, leaving the set of parameters fixed. For each value of $\|\boldsymbol{h}\|$ and c we quantify the number of times that $\widehat{\psi}$ belongs to the rejection region for a test of size $\alpha = 0.05$. Figure S2 shows the proportion of times (with respect to the number of simulations) the test rejected the null hypothesis in function of $\|\boldsymbol{h}\|$, each panel shows the model that was used to estimate the parameters, whereas the curves correspond to two fixed c values (1.645 and 1.960, respectively), and to whether the estimation method is correct or incorrect. In the nonseparable model there is a greater discrepancy in rejection rates between a correctly specified model and a misspecified when the value of c is high, but this discrepancy decreases when the value of c decreases, possibly influenced by the fact that the PA decreases considerably as the value of c varies, so that for relatively small values of $\|\boldsymbol{h}\|$, $\widehat{\psi}$ belongs to the rejection region. The same pattern is observed for the separable model, but the increasing ratio is slower. Similar information is displayed in Figure S3 but in this case it is shown a plot of the rejection rate versus c. All curves are decreasing and the differences for the well specified and misspecified models can be noticed looking at the rate of decaying of each curve.

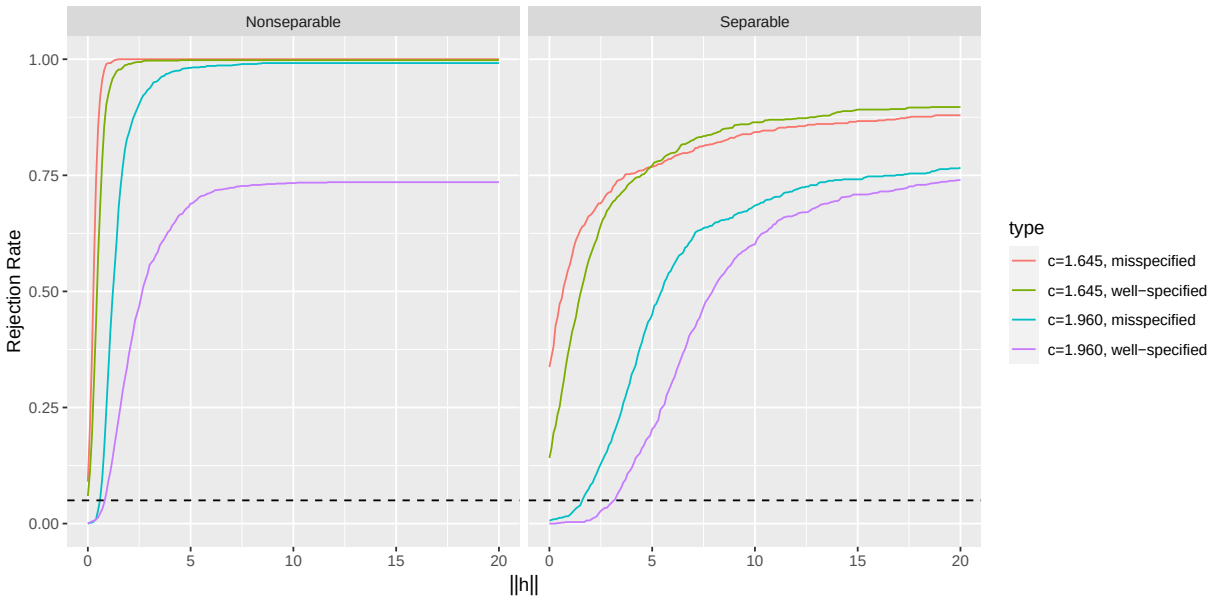


Figure S2: Rejection rate versus $\|\boldsymbol{h}\|$ for the test $\mathrm{H}_0 : \psi_c(\|\boldsymbol{h}\|) \geq 0.95$ versus $\mathrm{H}_1 : \psi_c(\|\boldsymbol{h}\|) < 0.95$, for $c = 1.960$ and $c = 1.645$, considering the both well-specified misspecified models.

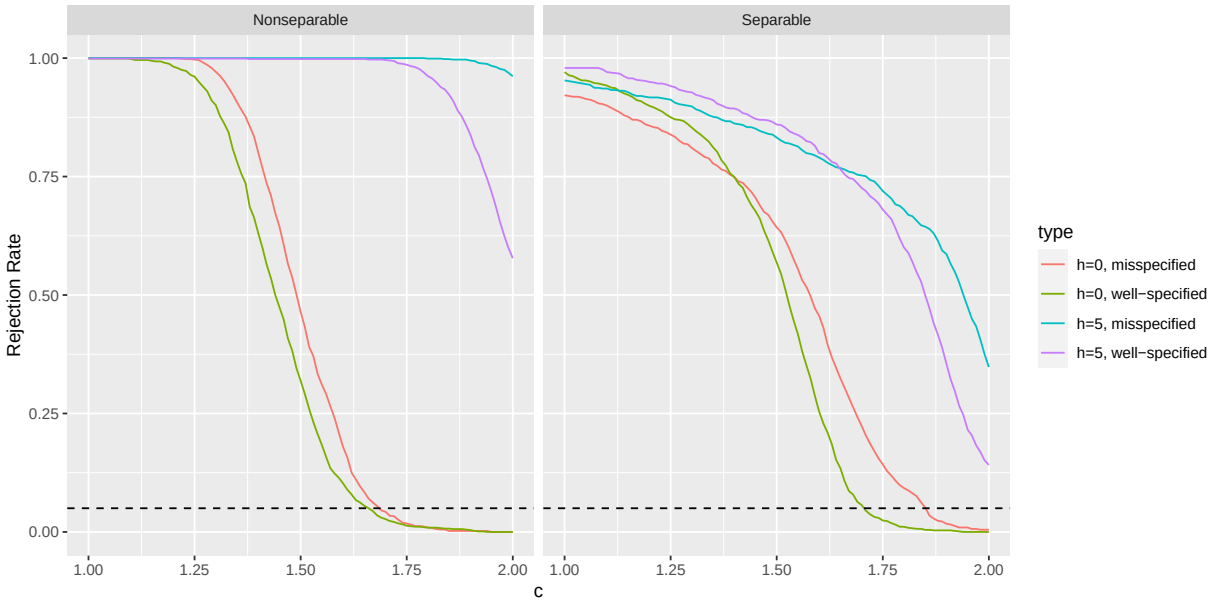


Figure S3: Rejection rate versus c for the test $\mathrm{H}_0 : \psi_c(\|\boldsymbol{h}\|) \geq 0.95$ versus $\mathrm{H}_1 : \psi_c(\|\boldsymbol{h}\|) < 0.95$ for $\boldsymbol{h} = 0$ and $\|\boldsymbol{h}\| = 5$, considering the both well-specified misspecified models.

Web Appendix C.2 Spatio-temporal Simulations

We used the GeoModels package version 1.0.8 (Bevilacqua et al., 2023) in R for the simulations, as it allowed us to simulate spatiotemporal processes with linear mean and our defined correlation structures. A regular grid of size $N_S \times N_S$ was considered for the spatial coordinates, N_T points in time from 1 to N_T with step 1. Five hundred runs were considered in this study. Examples are shown in Figures S4 and S5.

In Figures S6–S12 we display boxplots with the estimates of the probability of agreement as a function of $\|\boldsymbol{h}\|$, u and c for a spatiotemporal Gaussian process with negative and positive linear trends, with separable exponential covariance functions and also a non-separable Iacocesare covariance structure with fixed parameters given in Table 1 in the manuscript. The figures are generated for $N_S = 20$ and $N_S = 50$. The boxplots show the true values in red and the estimates in blue in each case. To observe the evolution of the estimated PA, each panel includes five boxplots, one for each different value of $\|\boldsymbol{h}\|$ for s set of fixed c and u values.

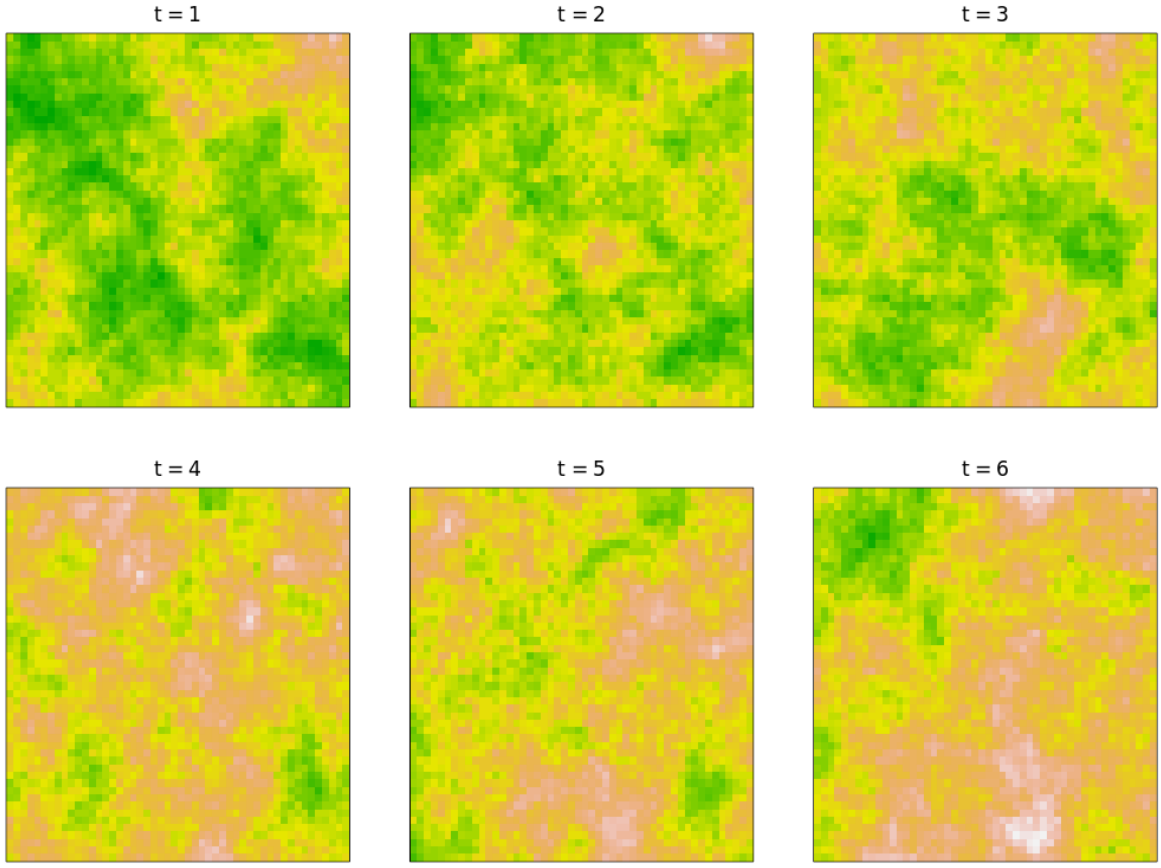


Figure S4: Simulated realizations of a spatiotemporal process defined by a Gaussian random field with an exponential separable covariance function for $N_S = 50$ and $N_T = 6$.

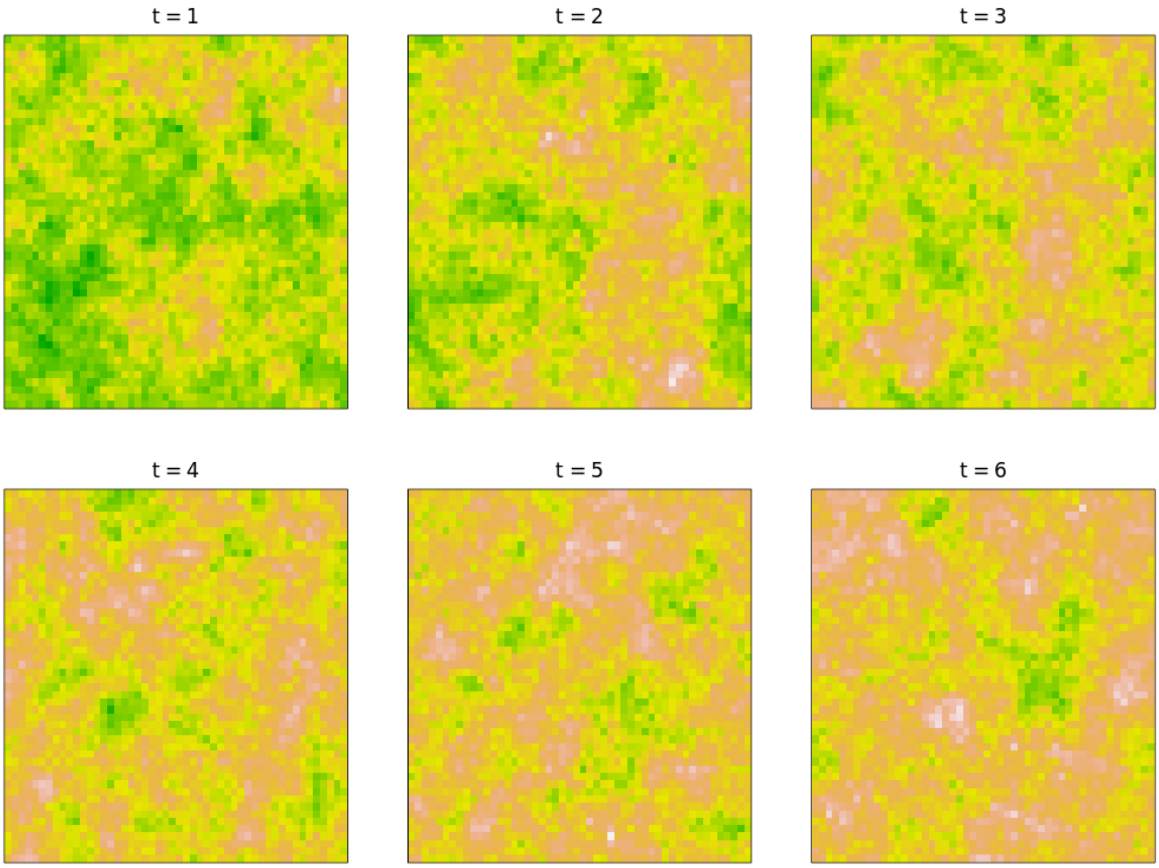


Figure S5: Simulated realization of a spatiotemporal process defined by a Gaussian random field with an Iacosecare non-separable covariance function for $N_S = 50$ and $N_T = 6$.

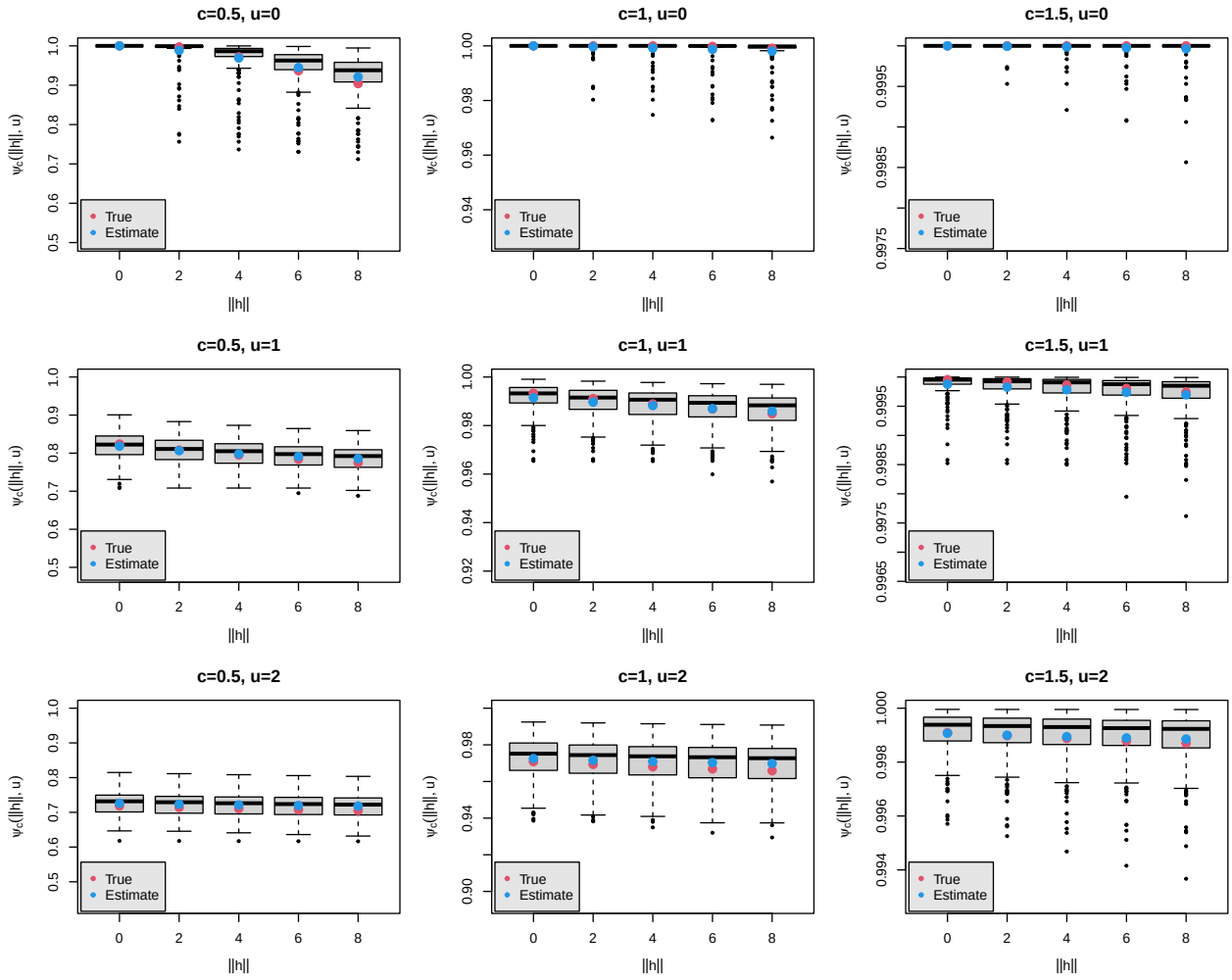


Figure S6: Estimates of the probability of agreement as a function of $\|\boldsymbol{h}\|$, u and c for a spatiotemporal Gaussian process with a negative linear trend and an exponential separable covariance structure with fixed parameters given in Table 1 (in the manuscript), and for $N_S = 20$. Note differences in range limits of the y-axis among the nine panels.

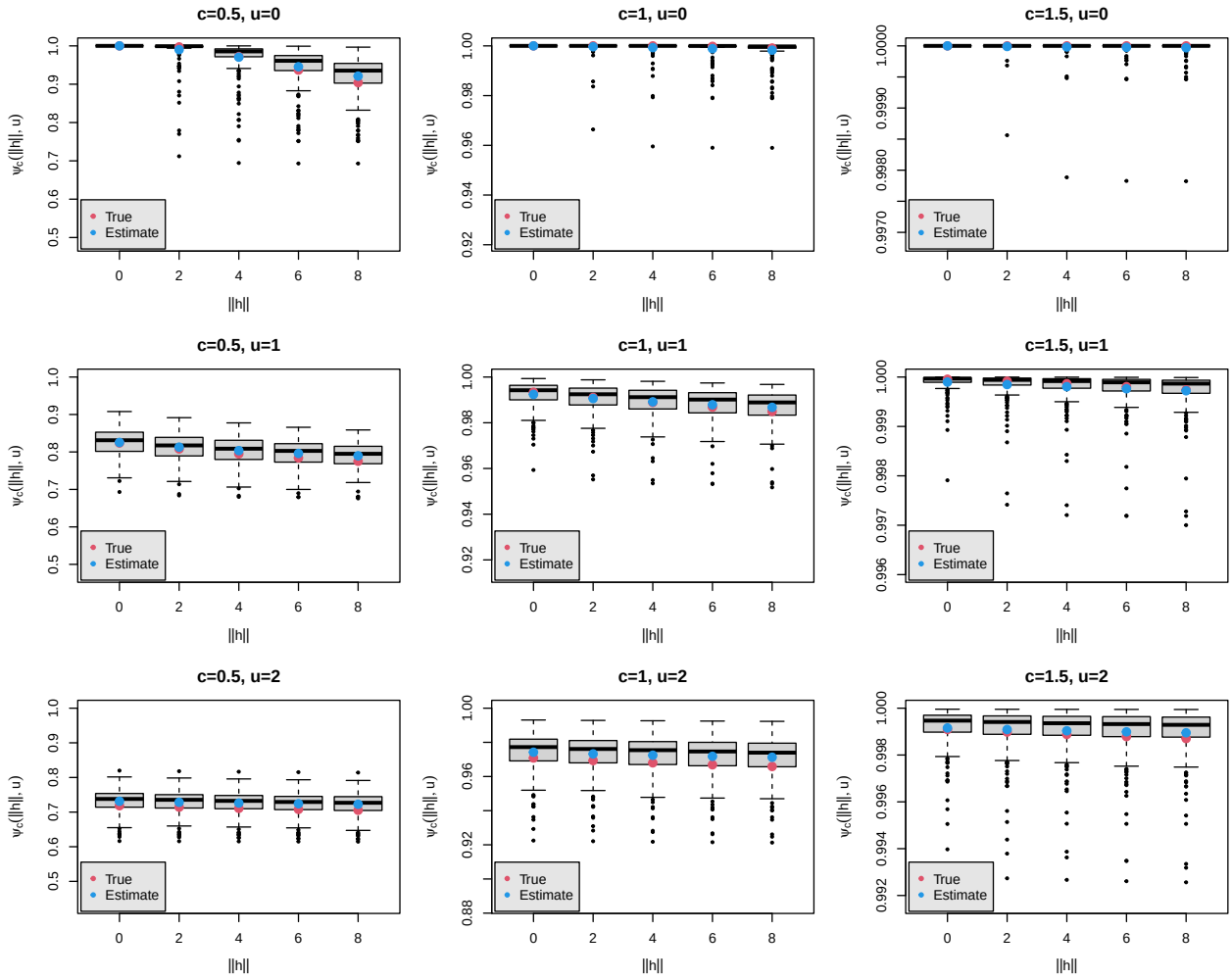


Figure S7: Estimates of the probability of agreement as a function of $\|\boldsymbol{h}\|$, u and c for a spatiotemporal Gaussian process with a positive linear trend and an exponential separable covariance structure with fixed parameters given in Table 1 (in the main text), and for $N_S = 20$. Note differences in range limits of the y-axis among the nine panels.

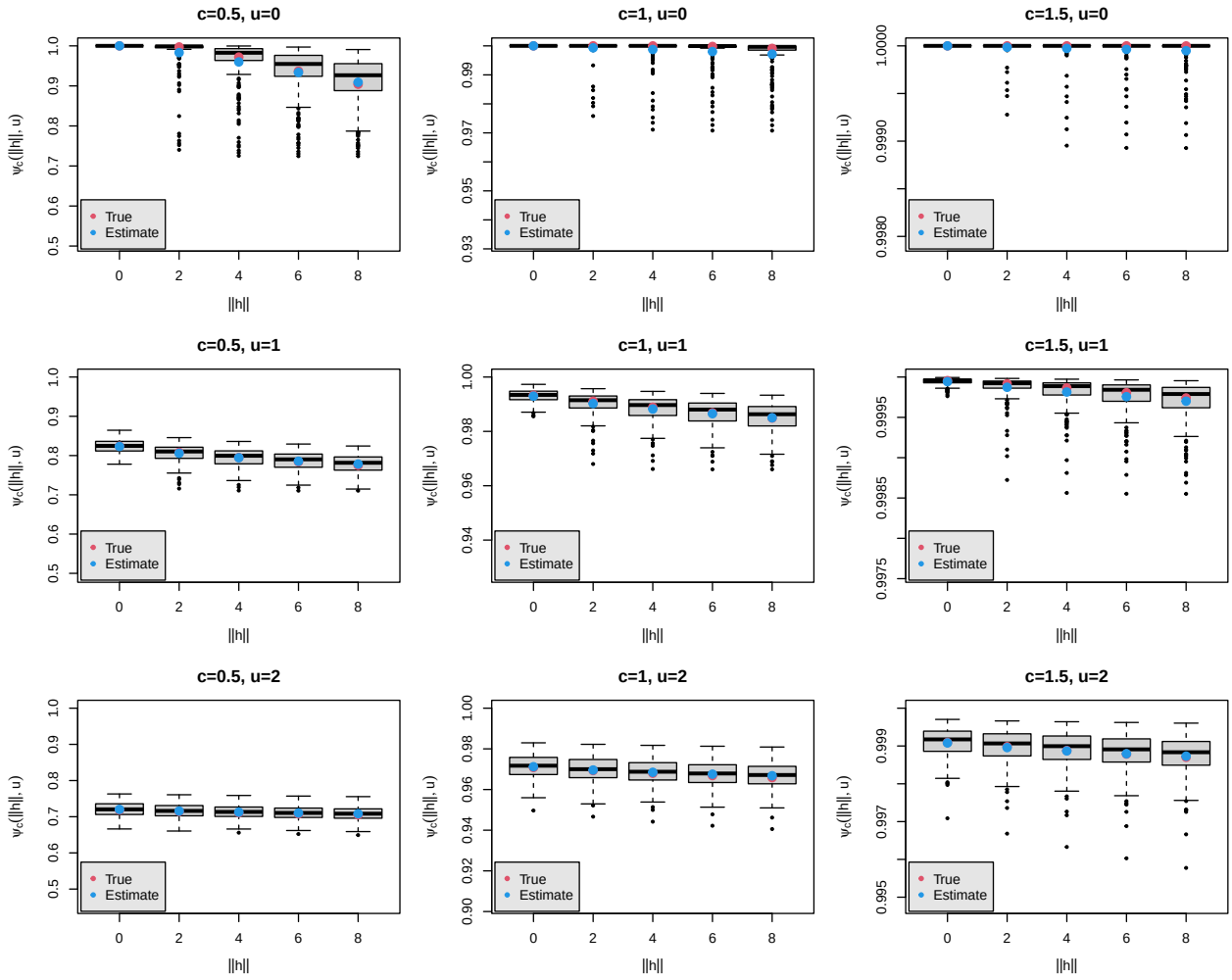


Figure S8: Estimates of the probability of agreement as a function of $\|\boldsymbol{h}\|$, u and c for a spatiotemporal Gaussian process with a negative linear trend and an exponential separable covariance structure with fixed parameters given in Table 1 (in the main text), and for $N_S = 50$. Note differences in range limits of the y-axis among the nine panels.

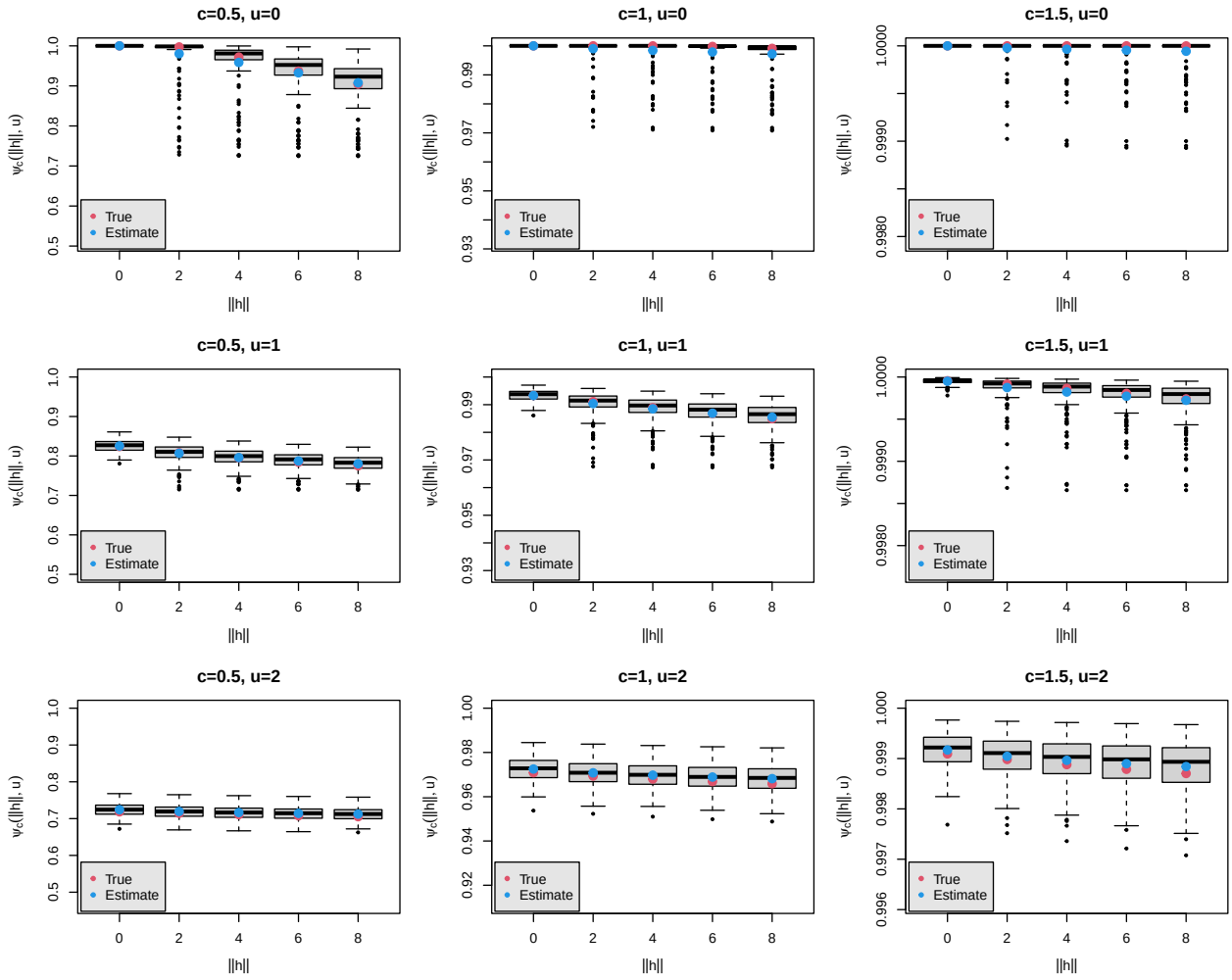


Figure S9: Estimates of the probability of agreement as a function of $\|\boldsymbol{h}\|$, u and c for a spatiotemporal Gaussian process with a positive linear trend and an exponential separable covariance structure with fixed parameters given in Table 1 (in the main text), and for $N_S = 50$. Note differences in range limits of the y-axis among the nine panels.

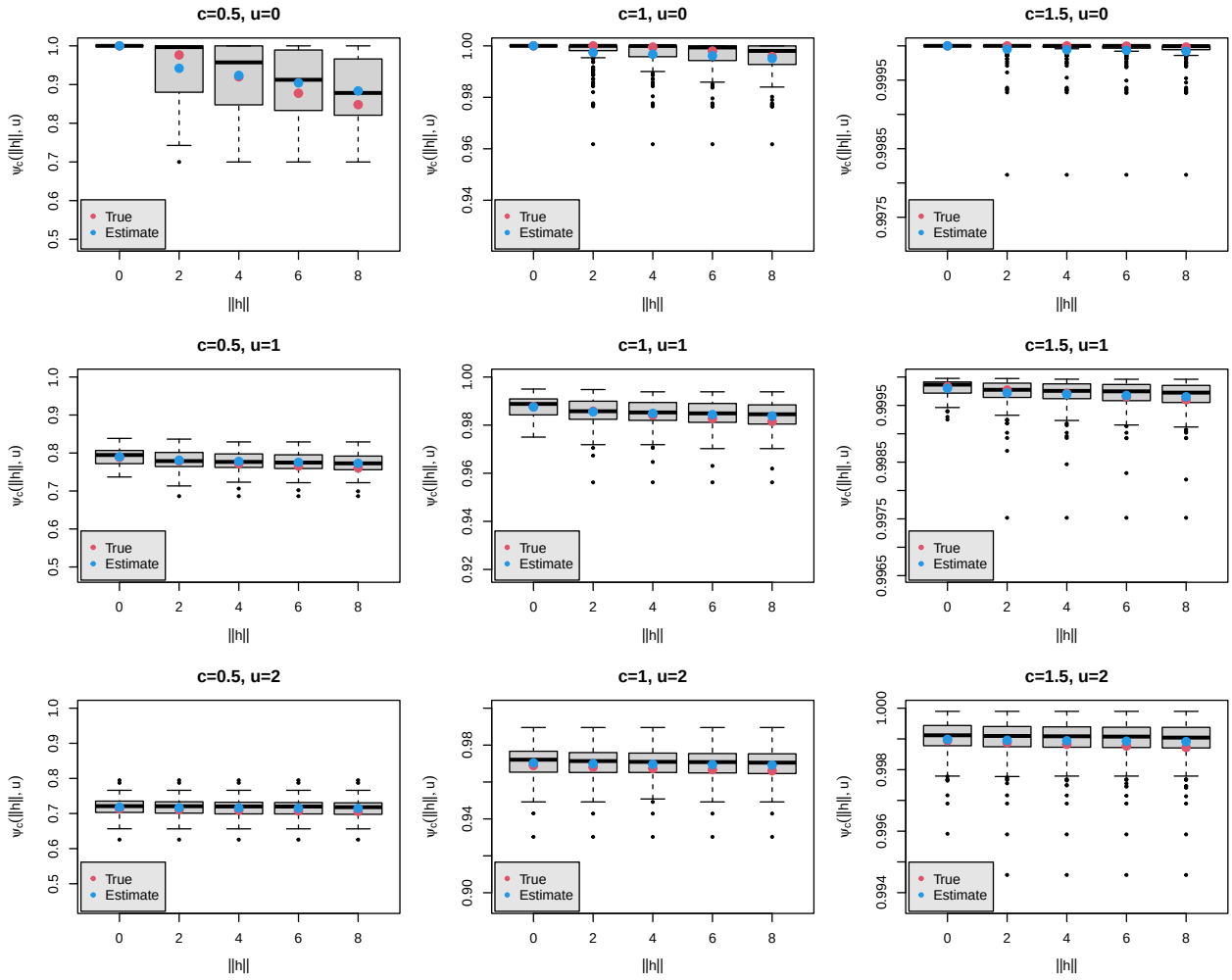


Figure S10: Estimates of the probability of agreement as a function of $\|\boldsymbol{h}\|$, u and c for a spatiotemporal Gaussian process with a negative linear trend and a non-separable Iacocesare covariance structure with fixed parameters given in Table 1 (in the main text), and for $N_S = 20$. Note differences in range limits of the y-axis among the nine panels.

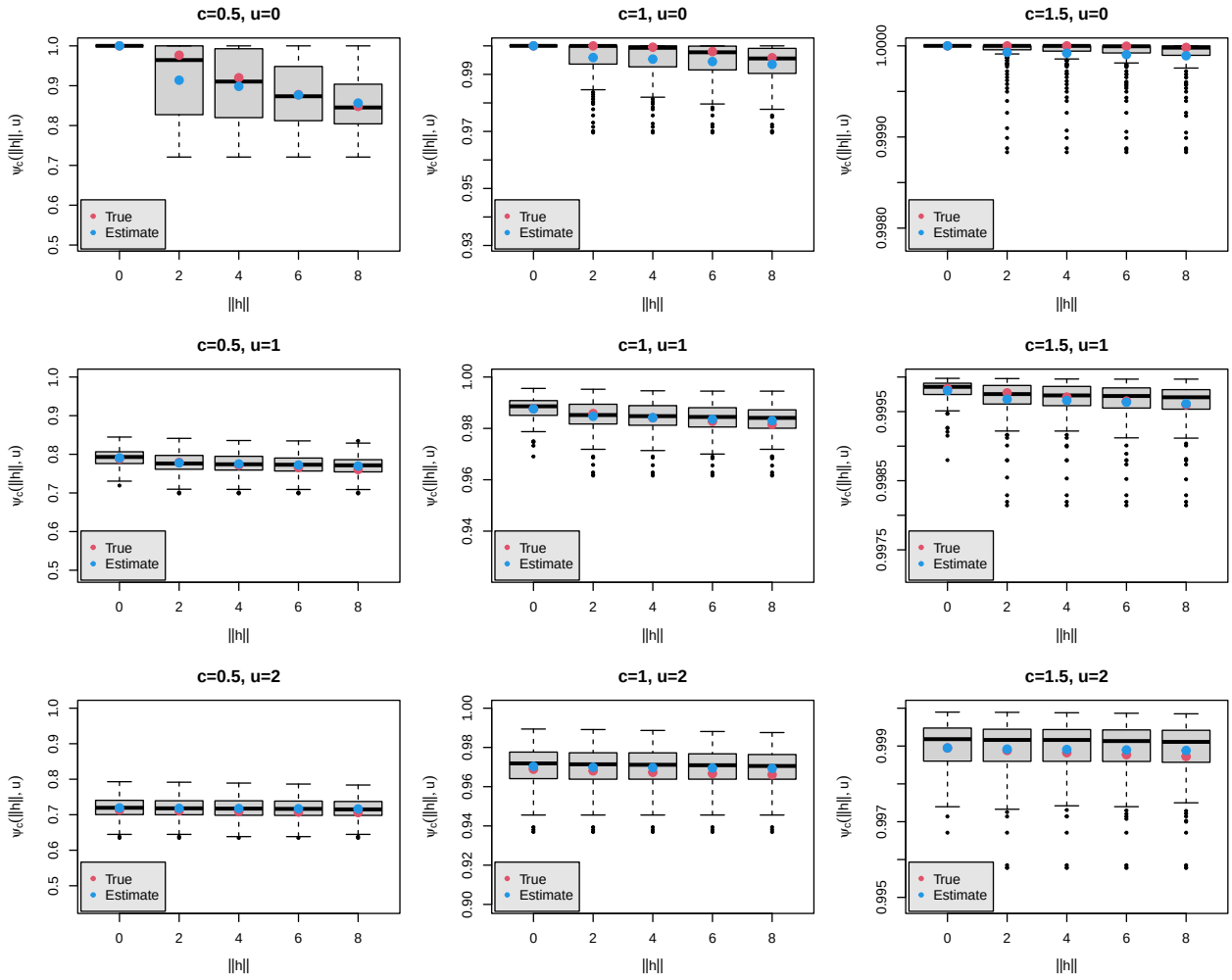


Figure S11: Estimates of the probability of agreement as a function of $\|\boldsymbol{h}\|$, u and c for a spatiotemporal Gaussian process with a negative linear trend and a non-separable Iacocesare covariance structure with fixed parameters given in Table 1 (in the main text), and for $N_S = 20$. Note differences in range limits of the y-axis among the nine panels.

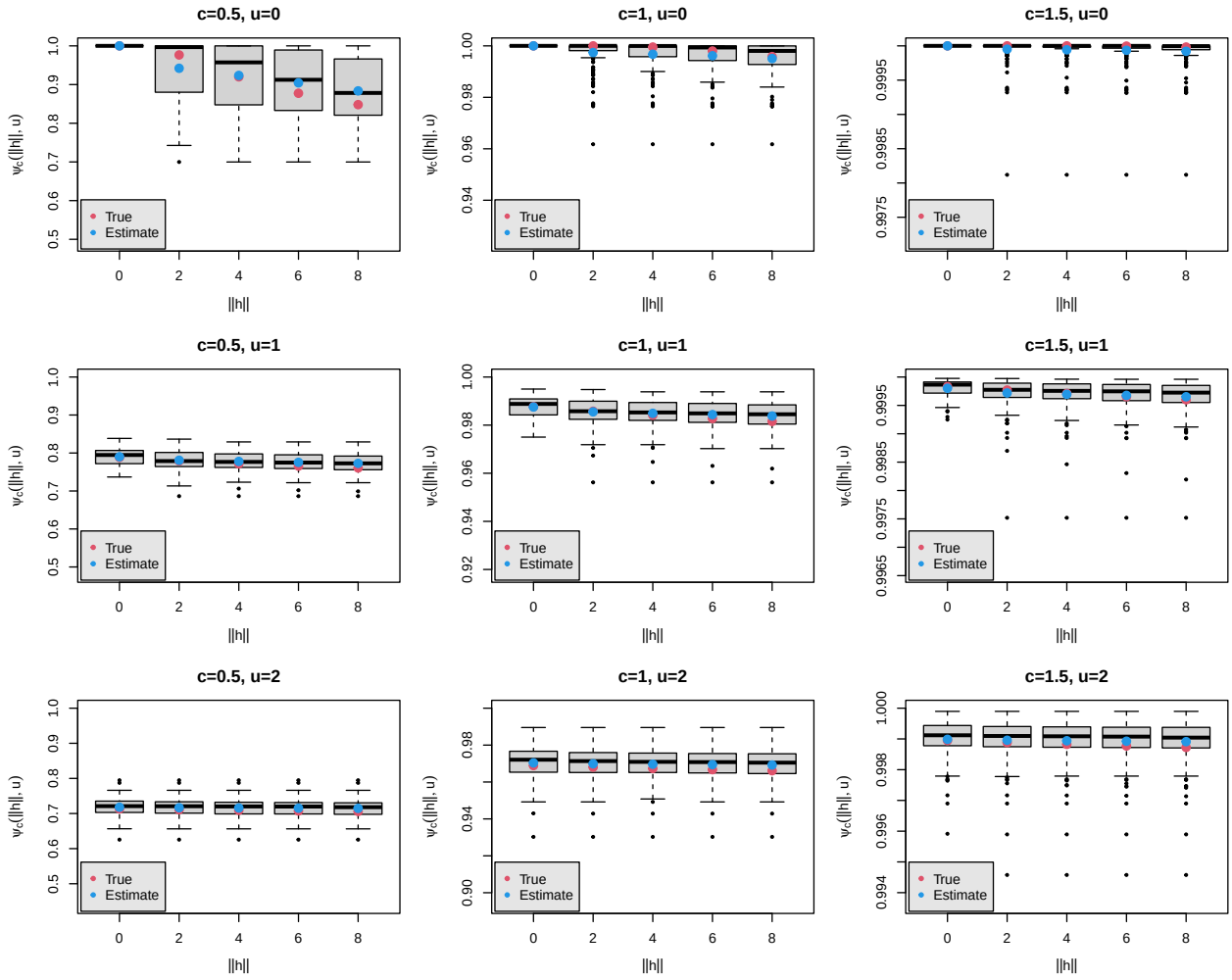


Figure S12: Estimates of the probability of agreement as a function of $\|\boldsymbol{h}\|$, u and c for a spatiotemporal Gaussian process with a negative linear trend and a non-separable Iacocesare covariance structure with fixed parameters given in Table 1 (in the main text), and for $N_S = 50$. Note differences in range limits of the y-axis among the nine panels.

The interpretation of the test when H_0 is rejected and $\boldsymbol{h}$ is variable is the same for each fixed value of u. However, when $\boldsymbol{h}$ is fixed and u is variable, the decision about H_0 does not necessarily follow the same structure unless the temporal trend is constant and the covariance model is monotonically decreasing in time. Moreover, if the covariance model is monotonically decreasing in time, then if H_0 is rejected for $u = 0$, H_0 is rejected for all u.

Web Appendix C.3 Application

In Figure S13 we present an empirical description of the dataset used in the application. Figure S13 (top left) displays an empirical semi-variogram that show the presence of spatial autocorrelation. Figure S13 (top right) shows a temporal semi-variogram which has an unstable behavior possibly due to the small sample size. Figure S13 (bottom left) shows a sequence of semi-variograms when the time is fixed. Finally, S13 (bottom left) displays a continuous space time semi-variogram as a surface summarizing the information provided in previous plots. For the hypothesis testing problem $H_0: \psi_c(\|\boldsymbol{h}\|, u) \geq 0.95$, versus $H_1: \psi_c(\|\boldsymbol{h}\|, u) < 0.95$ the test statistic and p-value were determined as follows

$$z = \frac{\widehat{\psi}_c(\boldsymbol{h}, u) - 0.95}{\sqrt{\widehat{\text{var}(\widehat{\psi}_c(\boldsymbol{h}, u))}}} \quad \Rightarrow \quad p\text{-value} = \Phi(z),$$

where $\Phi(\cdot)$ is the cumulative probability function of the standard normal distribution, and

$$\begin{aligned}
\widehat{\text{var}(\widehat{\psi}_c(\boldsymbol{h}, u))} &= \frac{2}{\pi} \exp\left\{ -\frac{(c - \widehat{\mu}_D)^2}{\sigma_D^2(\boldsymbol{h}, u; \widehat{\boldsymbol{\theta}})} \right\} \left\{ \widehat{V_{\mu_D}} + \frac{(c - \widehat{\mu}_D)^2}{\sigma_D^2(\boldsymbol{h}, u; \widehat{\boldsymbol{\theta}})} \widehat{V_{\sigma_D}(\boldsymbol{h}, u)} \right\}, \\
\widehat{V_{\sigma_D}(\boldsymbol{h}, u)} &= \frac{1}{4\sigma_D^2(\boldsymbol{h}, u; \widehat{\boldsymbol{\theta}})} \nabla \sigma_D^2(\boldsymbol{h}, u; \widehat{\boldsymbol{\theta}})^\top \boldsymbol{V}_{\widehat{\boldsymbol{\theta}}} \nabla \sigma_D^2(\boldsymbol{h}, u; \widehat{\boldsymbol{\theta}}).
\end{aligned}$$

In this case, $\widehat{\mu}_D = \widehat{\beta}_1 u$, and $\widehat{V_{\mu_D}} = u^2 V_{\widehat{\beta}_1}$. Figure S14 shows $\widehat{\text{var}(\widehat{\psi}_c(\boldsymbol{h}, u))}$ for the separable case. As expected, the variance is quite small because of the dimension of the data vector, which is equal to 25080. Accordingly the p-values are practically 0 or 1 (Figure S15). Because the dimension of the data vector is large, it is practically impossible to estimate the parameters of the variance components by maximum likelihood. Instead, we used the pairwise method provided in the R package GeoModels (Bevilacqua et al., 2023). We estimated $\boldsymbol{V}_{\boldsymbol{\theta}}$ using Fisher's information, which requires inverting matrices of order 25080×25080. This is impractical for the Iacocesare covariance. However, in the separable case, the temporal and spatial covariance matrices are of dimension 15×15 and 1672×1672, respectively. Thus, using the properties of the Kronecker product, we were able to obtain the estimates of the variances needed to perform the hypothesis testing for the dataset used in the application.

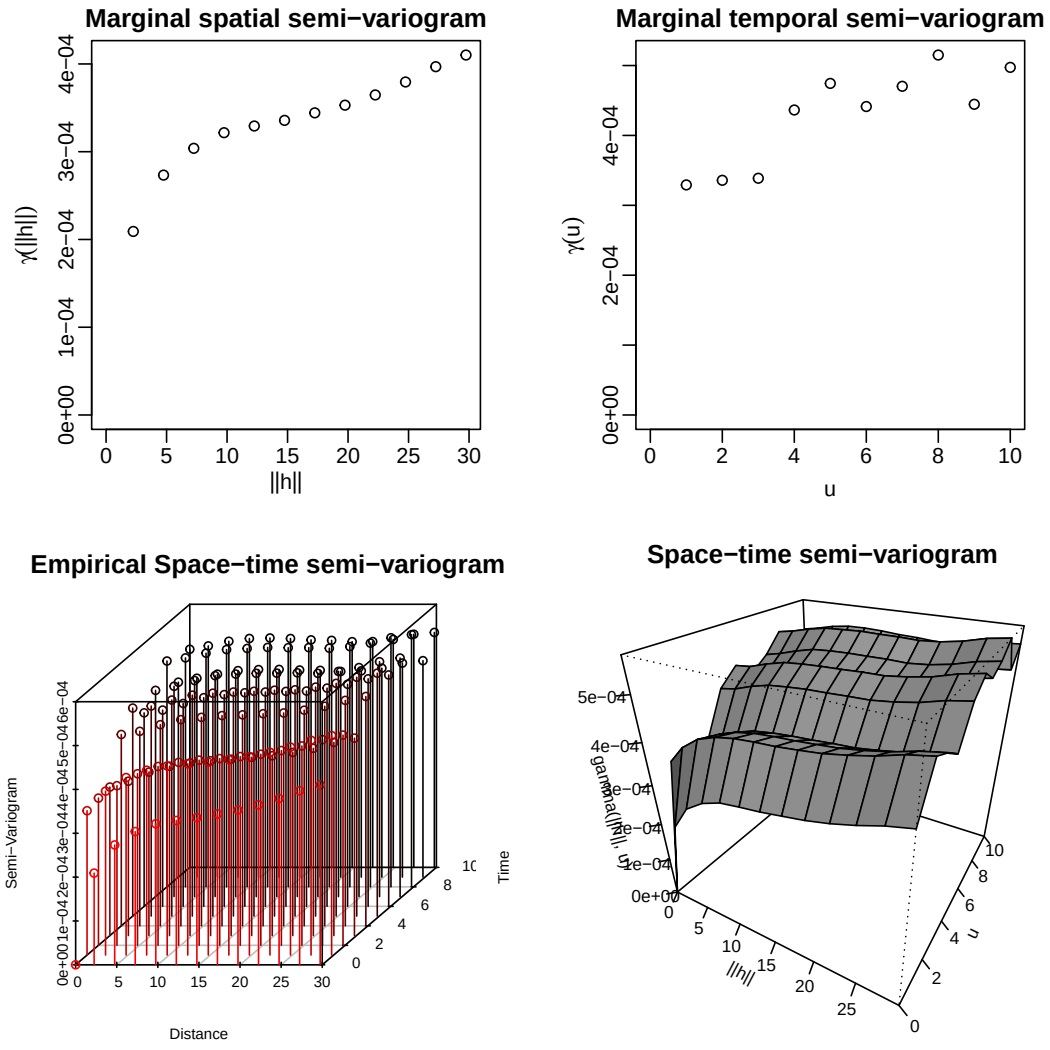


Figure S13: Empirical description of the data set.

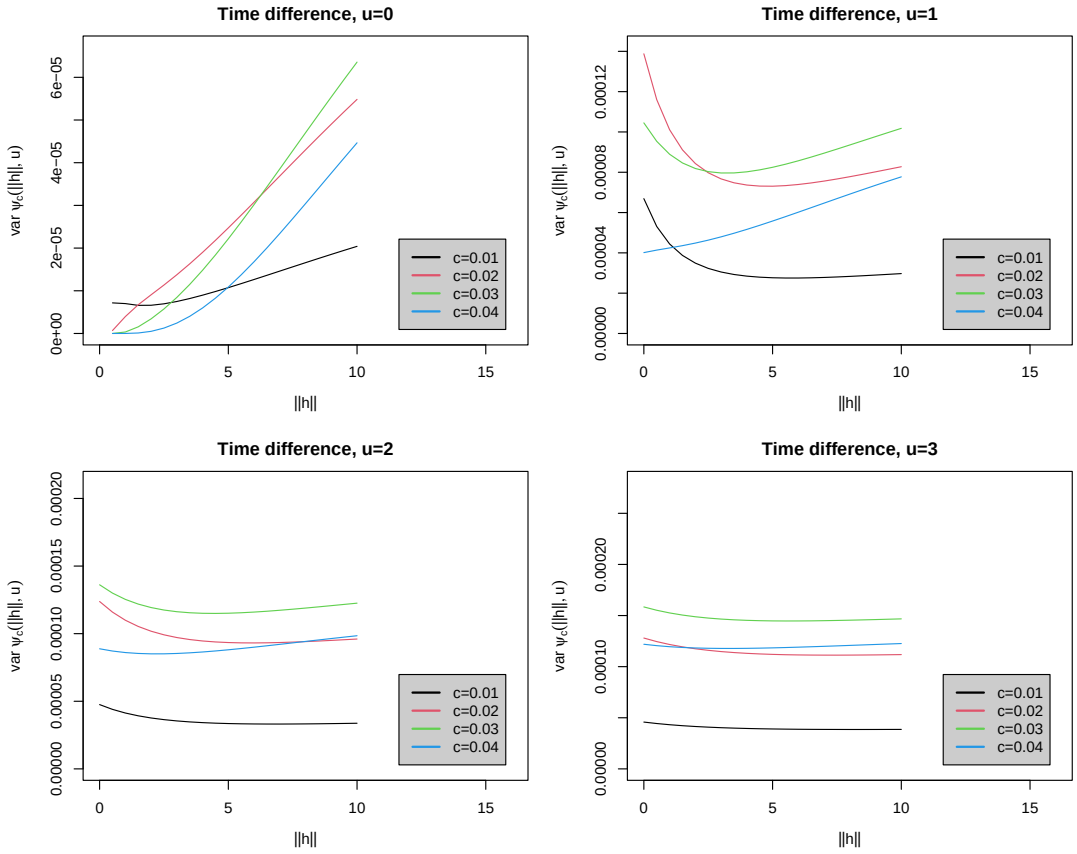


Figure S14: Variance of PA estimated in the separable case.

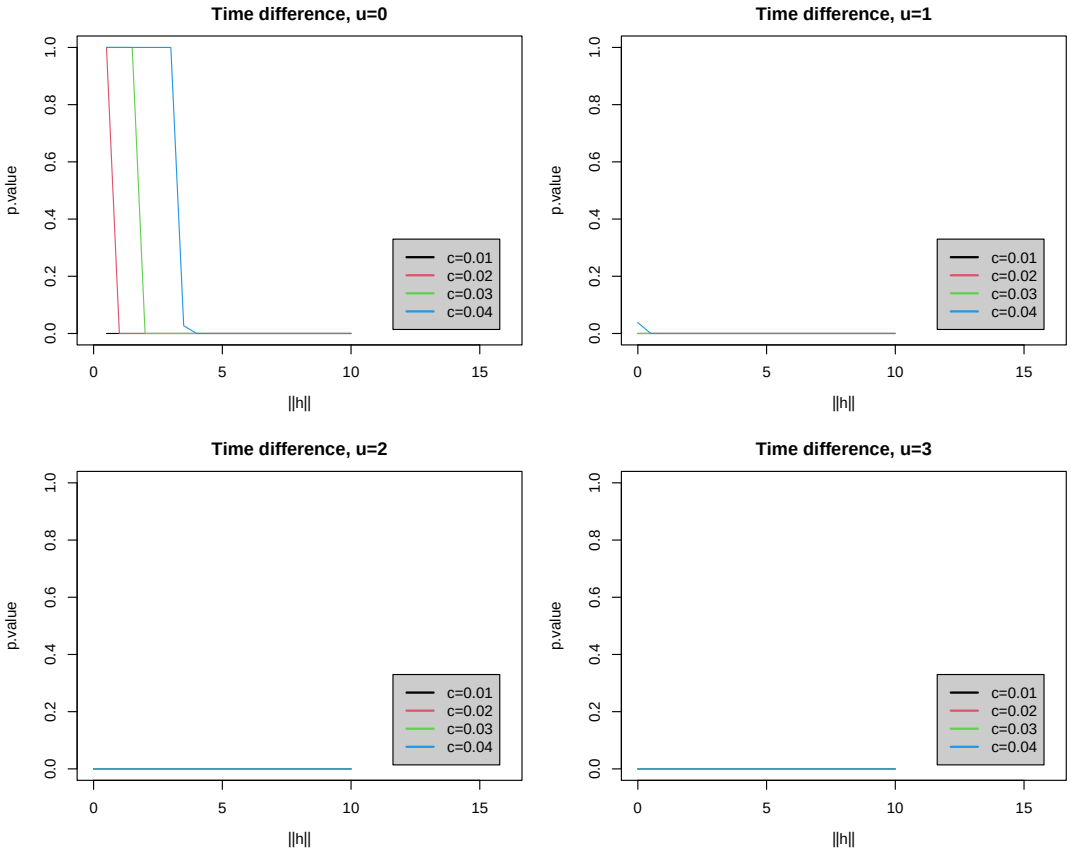


Figure S15: Variance of PA estimated in the separable case.

References

Acosta, J., and Vallejos, R. (2018). Effective sample size for spatial regression processes. Electronic Journal of Statistics 12, 3147–3180.

Gneiting, T., Kleiber, W., and Schlather, M. (2010). Matérn Cross-Covariance Functions for Multivariate Random Fields. Journal of the American Statistical Association, 105:491, 1167-1177

Bevilacqua, M., Morales-Oñate, V., and Caamaño-Carrillo, C. (2023). GeoModels: procedures for Gaussian and non-Gaussian. R package version 1.0.8, https://CRAN.R-project.org/package=GeoModels.

Lebedev, N.N. (1965). Special Functions and Their Applications, Prentice-Hall, New York.