On the reliability of some tests on type of non-separability and type of class of covariance models

C. Cappello[1], S. De Iaco[1,*], M. Palma[1] and D. Posa[1]

[1] *University of Salento, Via per Monteroni, Complesso Ecotekne, Lecce, Italy; claudia.cappello@unisalento.it, sandra.deiaco@unisalento.it, monica.palma@unisalento.it, donato.posa@unisalento.it*
[*]*Corresponding author*

Abstract. *In the literature, various tests for evaluating some characteristics of space-time covariance functions, such as symmetry and separability, are widely used. Recently, in case of rejection of the separability hypothesis, innovative tests have been proposed for evaluating the type of non-separability of space-time covariance functions and testing some well known classes of non-separable positive or negative covariance function models.*
In this paper a study on simulated data is proposed in order to assess the performance of the tests on the type of non-separability and on the classes of covariance functions.

Keywords. *Space-time covariance; Non-separability; Type of non-separability test; Test on class of covariance function models.*

1 Introduction

Apart from various tests for checking some second order properties such as symmetry and separability (Mitchell et al., 2005, 2006; Li et al., 2007, 2008), a test for the type of non-separability, as well as a statistical test for some classes of space-time covariance models were proposed in Cappello et al. (2018) and implemented in the R package `covatest`, which is available on CRAN (De Iaco et al., 2017). These tests help researchers in choosing the appropriate class of spatio-temporal covariance function model, for the spatio-temporal data analyzed.
In this paper a study on simulated data is proposed in order to assess the performance of the test on the type of non-separability and on some well known classes of covariance functions (i.e., the Gneiting and product-sum class of covariance functions).

2 Simulation study

In the following simulation study, the reliability of the test statistics defined in Cappello et al. (2018) and implemented in the R package `covatest` (De Iaco et al., 2017) is discussed.
Zero-mean simulated space-time realizations have been used to test the null hypotheses formulated on different types of non-separability and types of class of models. In particular the product-sum model

$C(\mathbf{h},u) = k_1 C_s(\mathbf{h}) C_t(u) + k_2 C_s(\mathbf{h}) + k_3 C_t(u)$, $k_1 > 0, k_2 \geq 0, k_3 \geq 0$, and the Gneiting model $C(\mathbf{h},u) = \sigma^2 \left(\frac{1}{(b|u|^{2\alpha}+1)^{\tau}}\right) \cdot exp\left(-\frac{a||\mathbf{h}||^{2\gamma}}{(b|u|^{2\alpha}+1)^{\gamma\beta}}\right)$, $\beta \in [0,1], \tau \geq \beta d/2$, have been used to generate simulated space-time data regularly distributed over a range of grid sizes (spatial grids of dimensions 9×9 and 15×15), with temporal lengths $|T_n|$=600 and $|T_n|$=1000. The product-sum model has exponential marginals with spatial and temporal effective ranges equal, respectively, to 3 and 20 and parameters $(k_1, k_2, k_3) = (0.5, 0.3, 0.2)$. On the other hand the Gneiting model has marginals with linear behaviour near the origin (with smoothness parameters γ and α equal to 0.5) and $(a, b, \beta, \tau, \sigma^2) = (1, 0.75, 1, 1, 1)$ (which correspond to spatial and temporal marginals that decay approximately at 3 and 20, respectively).
These two classes of covariance function models have been considered to produce alternative simulations since they present two different types of non-separability, i.e., the product-sum class is negative non-separable and the Gneiting class is positive non-separable.
The goodness of the tests have been evaluated through the study of 900 simulations, obtained through a Gaussian-related program, that is the sequential simulation algorithm, based on the above mentioned classes of covariance function models.
The simulation study focused on the analysis of the empirical size and power of the tests for different grid sizes, temporal lengths and classes of models. In particular, for the test on the type of non-separability

- data sets simulated through the product-sum model, which is uniform negative non-separable, have been considered to compute (a) the empirical size through the frequency of rejecting the uniform negative non-separability ($Fr\{R_{H_0^{(-)}}|H_0^{(-)}\}$), and (b) the empirical power through the frequency of rejecting the uniform positive non-separability ($Fr\{R_{H_0^{(+)}}|H_1^{(-)}\}$);
- data sets simulated through the Gneiting model, which is uniform positive non-separable, have been used to compute (a) the empirical size through the frequency of rejecting the uniform positive non-separability ($Fr\{R_{H_0^{(+)}}|H_0^{(+)}\}$), and (b) the empirical power through the frequency of rejecting the uniform negative non-separability ($Fr\{R_{H_0^{(-)}}|H_1^{(+)}\}$).

Moreover, an indirect way of approximating the power of the test has been also proposed. It has been evaluated how large is the p-value for the decision of non-rejection (when the null hypothesis is true), therefore the frequencies of non-rejecting the null hypotheses with large p-values (greater than 0.9), denoted with $Fr\{\bar{R}_{H_0^{(+)}}|H_0^{(+)}; \text{p-values} > 0.9\}$ and $Fr\{\bar{R}_{H_0^{(-)}}|H_0^{(-)}; \text{p-values} > 0.9\}$, have been computed. For the tests on the type of class of models the size and power have been also determined. In particular

- Gneiting model-based data have been used to compute (a) the empirical size, through the frequency of rejecting the same Gneiting model ($Fr\{R_{H_0^{\text{Gn}}}|H_0^{\text{Gn}}\}$) and (b) the empirical power through the frequency of rejecting the null hypotheses formulated on two different classes, such as the product-sum model ($Fr\{R_{H_0^{\text{PS}}}|H_1^{\text{Gn}}\}$) and the integrated product model ($Fr\{R_{H_0^{\text{IP}}}|H_1^{\text{Gn}}\}$). Note that the power of the test on the Gneiting class has been analyzed with respect to the product-sum model, which is negative non-separable and the integrated product model, which is positive non-separable;
- product-sum model-based data have been used to determine (a) the empirical size through the frequency of rejecting the product-sum model, ($Fr\{R_{H_0^{\text{PS}}}|H_0^{\text{PS}}\}$) and (b) the empirical powers, the frequency of rejecting the null hypotheses formulated on the Gneiting class ($Fr\{R_{H_0^{\text{Gn}}}|H_1^{\text{PS}}\}$) and the integrated product class ($Fr\{R_{H_0^{\text{IP}}}|H_1^{\text{PS}}\}$), which are positive non-separable.

In addition, the frequencies of non-rejecting the null hypotheses (when it is true) with large p-values (greater than 0.9) have been computed as an indirect way to approximate the power of the test. These

frequencies are denoted with $Fr\{\bar{R}_{H_0^{\mathrm{Gn}}}|H_0^{\mathrm{Gn}};\text{p-values} > 0.9\}$ and $Fr\{\bar{R}_{H_0^{\mathrm{PS}}}|H_0^{\mathrm{PS}};\text{p-values} > 0.9\}$.
As stated above, the testing procedure has been applied to the zero-mean simulated data sets, obtained for different alternatives in terms of grid size, temporal length and class of models; spatial couples and temporal lags at distances 1 and 2 have been considered for the tests. The results of the test on the type of non-separability, i.e., the empirical size with respect to the nominal level 0.05 and power are given in Tab. 1. Looking at the results, it is clear that the size of the test (p_1 and p'_1) is close to the nominal level and the power (p_3 and p'_3) approaches 1 as the grid size and temporal length increase; similarly for the approximated powers (p_2 and p'_2), measured in terms of frequencies of non-rejecting the null hypotheses (when it is true) with large p-values (greater than 0.90). These results confirm the reliability of the test and that there is strong confidence in rejecting the null hypothesis of negative/positive non-separability when the alternative hypothesis is valid, as well as in failing to reject the null hypothesis when the null hypothesis is valid.

		Negative non-separable model-based simulations			Positive non-separable model-based simulations		
		p_1	p_2	p_3	p'_1	p'_2	p'_3
9×9	$\|T_n\| = 600$	0.080	0.093	0.747	0.080	0.093	0.693
	$\|T_n\| = 1000$	0.053	0.107	0.933	0.040	0.107	0.920
15×15	$\|T_n\| = 600$	0.067	0.107	0.893	0.067	0.093	0.813
	$\|T_n\| = 1000$	0.040	0.120	0.987	0.053	0.120	0.973

Table 1: Values of the empirical size and power for the tests on type of non-separability for data simulated through a uniform negative non-separable model ($p_1 = Fr\{R_{H_0^{(-)}}|H_0^{(-)}\}$, $p_2 = Fr\{\bar{R}_{H_0^{(-)}}|H_0^{(-)};\text{p-values} > 0.9\}$ and $p_3 = Fr\{R_{H_0^{(+)}}|H_1^{(-)}\}$) and through a uniform positive non-separable model ($p'_1 = Fr\{R_{H_0^{(+)}}|H_0^{(+)}\}$, $p'_2 = Fr\{\bar{R}_{H_0^{(+)}}|H_0^{(+)};\text{p-values} > 0.9\}$ and $p'_3 = Fr\{R_{H_0^{(-)}}|H_1^{(+)}\}$).

The results for the test on the type of class of models are show in Tab. 2. The size (p_1 and p'_1) is close to the nominal level for each option, while the empirical power (p_3, p_4 and p'_3, p'_4) supports the rejection decision of the null hypothesis when it is false. The approximated powers (p_2 and p'_2) are consistent with respect to the nominal frequency of the non-rejection decision of the null hypothesis (when it is valid) with p-value greater than 0.9. Note also that the powers and the approximated powers of all alternatives are nearly equivalent when the temporal length is equal to 1000. Moreover, the tests have greater power when the underlining data are generated by a covariance model characterized by a different type of non-separability with respect to the class of model under the null hypothesis (i.e., p_3 is greater than p_4).

		Gneiting model -based simulations				Product-sum -based simulations			
		p_1	p_2	p_3	p_4	p'_1	p'_2	p'_3	p'_4
9×9	$\|T_n\| = 600$	0.080	0.080	0.827	0.707	0.067	0.080	0.893	0.853
	$\|T_n\| = 1000$	0.040	0.107	0.973	0.773	0.040	0.107	0.987	0.973
15×15	$\|T_n\| = 600$	0.067	0.093	0.947	0.720	0.053	0.107	0.907	0.880
	$\|T_n\| = 1000$	0.053	0.133	0.987	0.813	0.040	0.120	1.000	0.987

Table 2: Values of the empirical size and power for the tests on type of class of covariance function models for data simulated through the Gneiting model ($p_1 = Fr\{R_{H_0^{\mathrm{Gn}}}|H_0^{\mathrm{Gn}}\}$, $p_2 = Fr\{\bar{R}_{H_0^{\mathrm{Gn}}}|H_0^{\mathrm{Gn}};\text{p-values} > 0.9\}$, $p_3 = Fr\{R_{H_0^{\mathrm{PS}}}|H_1^{\mathrm{Gn}}\}$ and $p_4 = Fr\{R_{H_0^{\mathrm{IP}}}|H_1^{\mathrm{Gn}}\}$) and through the product-sum model ($p'_1 = Fr\{R_{H_0^{\mathrm{PS}}}|H_0^{\mathrm{PS}}\}$, $p'_2 = Fr\{\bar{R}_{H_0^{\mathrm{PS}}}|H_0^{\mathrm{PS}};\text{p-values} > 0.9\}$, $p'_3 = Fr\{R_{H_0^{\mathrm{Gn}}}|H_1^{\mathrm{PS}}\}$ and $p'_4 = Fr\{R_{H_0^{\mathrm{IP}}}|H_1^{\mathrm{PS}}\}$).

From the results in Tab. 1 and 2 it is evident that (a) the grid size does not significantly affect the size of the test, which is around the nominal level even if the series length is equal to 600 and (b) the power

increases as temporal length increases.
Finally, the product-sum and the Gneiting model-based simulations have been also used to evaluate how rapidly the empirical distribution function of the test statistic on the type of non-separability and on the type of class of models converges in distribution to a normal and a Chi-square, respectively, according to the results of the multivariate delta theorem of Mardia et al. (1979) and Li et al. (2007). In particular, the temporal length of simulated data increases from 400 up to 1000, with increments of 200 time points for each step and the Kolomogorov-Smirnov tests have been applied for comparing the observed cumulative distribution functions of the test statistics with the corresponding theoretical distributions. The empirical distribution function of the test statistic on the type of non-separability rapidly converges to a normal distribution, even when the temporal length is greater than 400. The p-values for the Kolomogorov-Smirnov tests support the non-rejection of the null hypothesis for all options. The same goes for the empirical distribution function of the test statistic on the class of models. In this case, the p-values, which support the non-rejection of the null hypothesis for all options, are greater than 0.8 when the temporal length is greater than 800 and approach 1 when the temporal length is equal to 1000.

3 Conclusions

In this paper the reliability of the statistical tests for checking different forms of non-separability and some classes of space-time covariance function models was analyzed. The empirical results obtained through the simulated data confirm the goodness of these tests and can stimulate their use in the applications.

References

Cappello, C., De Iaco, S., Posa, D. (2018). Testing the type of non-separability and some classes of space-time covariance function models. *Stoch. Environ. Res. and Risk Assess.* **32**, 17–35.

De Iaco, S. and Posa, D. (2013). Positive and negative non-separability for space-time covariance models. *J. Stat. Plan. Infer.* **143**(2), 378–391.

De Iaco, S., Cappello, C., Posa, D., Maggio, S. (2017). Covatest: Tests on properties of space-time covariance functions. The Comprehensive R Archive Network, https://CRAN.R-project.org/package=covatest, 1–18.

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**(478), 736–744.

Li, B., Genton, M. G. and Sherman, M. (2008). On the asymptotic joint distribution of sample space-time covariance estimators. *Bernoulli*. **14**(1), 228–248.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979). Multivariate Analysis, New York: Academic Press. 521p.

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

Mitchell, M. W., Genton, M. G., and Gumpertz, M. L. (2006). A Likelihood Ratio Test for Separability of Covariances. *Journal of Multivariate Analysis*. **97**, 1025–1043.

Rodrigues, A., Diggle, P. J. (2010). A Class of Convolution-Based Models for Spatio-Temporal Processes with Non-Separable Covariance Structure. *Scand. J. Stat.* **37**(4), 553–567.