Statistical Methods & Applications (2026) 35:181–205
https://doi.org/10.1007/s10260-026-00841-4

ORIGINAL PAPER

On the use of difference-based covariance models

Sandra De Iaco[1,2] · **Monica Palma**[1] · **Donato Posa**[1]

Received: 23 July 2025 / Accepted: 26 March 2026 / Published online: 6 May 2026
© The Author(s) 2026

Abstract
There exist applications for which the phenomenon under study can have not always positive correlation structures. The traditional Whittle-Matern family of covariance functions is distinguished by being positive, hence this class is not able to model correlation structures with negative values. In the very recent years, theoretical results regarding families of covariance functions developed through the difference of covariance functions and characterized by negative values in a subset of the corresponding domain, have been obtained in the literature, both in the complex and in the real domain. In this paper, some models belonging to these families are reviewed and their flexibility, with respect to the traditional covariance models, is underlined from a practical point of view. Various case studies referred to distinct temporal and spatial datasets whose correlation structures are featured by positive and negative values, are discussed in the paper with the aim of highlighting the pro of fitting difference-based covariance models with respect to the traditional ones. The advantages of the difference-based covariance models are assessed in terms of predictive performances. Finally, an application on a 2-dimensional simulated dataset is also furnished.

Keywords Negative correlation · Hole effect model · Bochner's theorem · Predictive performances

✉ Monica Palma
monica.palma@unisalento.it

Sandra De Iaco
sandra.deiaco@unisalento.it

Donato Posa
donato.posa@unisalento.it

[1] Department of Economic Sciences, University of Salento, Via per Monteroni, Lecce 73100, Italy

[2] National Biodiversity Future Center - NBFC, Palermo, Italy

1 Introduction

Over the years, the concept of spatial correlation has grown significantly in mining, geography, ecology, epidemiology, sociology, environmental studies and many other disciplines based on geo-referenced data. The notion of spatial correlation is a fundamental basis in the field of spatial statistics and plays a central role in Geostatistics and also in spatial econometrics. Among the crucial geostatistical contributions, it is worth mentioning the ones of the geologist-statistician Matheron (1963), who developed the theory of the regionalized random variables, with his concepts of intrinsic stationarity and intrinsic variogram, as well as the fundamental textbooks of Geostatistics by Journel and Huijbregts (1978) and Cressie (1993). Getis (2008) traced the evolution of spatial correlation, stating that scientists have been addressing the idea of distance effects for a very long time. Some references to distance-related concepts can be traced back to von Thünen with his land use theory (Von Thünen 1826) and the to the German economic geographers Losch (1940) and Christaller (1966), who emphasized spatial interaction models where distance-decay and cost-distance were fundamental ideas. Getis recalled also the Moran Index (Moran 1948), with his measure of spatial autocorrelation, which was successively formalized and generalized by Cliff and Ord in their monograph (Cliff and Ord 1968), followed by Anselin (1988), in the field of spatial econometrics, who provided a comprehensive approach to understand spatial effects in econometric models. In some particular circumstances, the phenomenon under study can be characterized by not always positive correlation structure. In spatial econometrics, several geo-referenced variables with positive and negative spatial autocorrelations can be recognized. Jacob et al. (2008) were more likely among the first authors to suggest spatial correlation with negative values and proposed this model in the context of spatially structured effects associated with mosquito habitats. Some examples of negative autocorrelation in a spatial context were discussed by Griffith and Arbia (2010), who investigated empirical situations characterized by negative spatial autocorrelation masked sometimes by positive spatial autocorrelation, by Hu et al. (2018, 2020) who employed Moran eigenvector spatial filtering methodology to account for the positive and negative spatial autocorrelation mixture for the investigated female breast cancer incidences in Florida, as well as by Griffith (2019) where several real cases characterized by negative spatial autocorrelation structure were mentioned. In a previous work, Griffith (2016) affirmed that the negative spatial autocorrelation occurs in spatial patterns of competition and mentioned the case where a school attempts to enlarge its number of students enrollment by reducing the number of students enrolled to the schools in the neighborhood. However, negative spatial autocorrelation has long been ignored in the literature and often neglected in empirical studies.

In spatial or spatio-temporal geostatistical analysis, covariance functions play a crucial role for analyzing the spatial or the spatio-temporal behavior of the phenomenon under study. The Whittle-Matérn family (Matérn 1980) and the various families of covariance functions constructed by applying some classical properties (Yaglom 1987; Cressie 1993; De Iaco et al. 2002, 2019, 2020; Hristopulos 2020)

have been largely utilized in the applications for modeling and prediction purposes. However, Whittle-Matérn class, as well as the generalized Cauchy family (Gneiting and Schlather 2004), cannot ever describe negative spatial autocorrelation (Posa 2021; Alegria et al. 2024).

Additionally, a special class of covariance functions (the generalized sum of product models), characterized by negative weights, was analyzed by Gregori et al. (2008) and further discussed in De Iaco et al. (2011). It is worth to mention some linear combinations containing negative coefficients that do not describe negative spatial autocorrelation function, namely the circular model (it is conditionally valid only in one and two dimensions), the cubic, the spherical, the tetra-spherical and the penta-spherical model. Note that none of the aforementioned class of covariance functions can be used in case of negative spatial dependencies. Moreover, some other linear combinations of continuous covariance functions with negative weights were proposed by Vecchia (1988), Vargas-Guzmán et al. (2002) and Ma (2005).

For decades negative spatial correlation structures have been modelled through the hole effect covariance functions (Journel and Froidevaux 1982; Ma and Jones 2001; Asghari 2015), which belong to the Bessel family (Hristopulos 2020; Posa 2025). The Bessel function of the first kind enables valid specifications that can capture negative spatial autocorrelation, due to its oscillating form, which can align with periodic behavior or alternating patterns in geo-referenced data. However the Bessel family is featured by a parabolic behavior near the origin and a countable infinity of zeros, neglecting the cases with a different behaviour near the origin and only one zero. Note that, as discussed in Griffith (2019), negative spatial autocorrelation differentiates between discrete (polygons) and continuous (points separated by distance) geography cases.

Recently, Posa (2021, 2025) analyzed the general characteristics concerning the difference between two covariance functions, providing the conditions to obtain valid covariance models from the difference of two complex as well as real-valued covariance functions. In particular, the properties of new classes of isotropic covariance models were explored by highlighting their flexibility in describing positive covariances and also covariance functions which are negative in a subset of their domain, according to the respective parameters values. Moreover, the new classes of covariance models proposed by Posa (2021) can be characterized by linear or parabolic behavior near the origin.

In this paper, after recalling the theoretical basis concerning the class of covariance functions (Section 2) and the traditional models adopted in presence of negative correlation (Section 3), some examples for differences of two covariance functions are reported. Thus, temporal and spatial datasets which exhibit negative correlation, where the sample correlograms can be fitted by using the new class of models are provided (Section 4) and an effective simulation study is discussed (Section 5). The above applications demonstrate the importance to get flexible models capable to describe and analyze several and various correlation structures, as further underlined in the conclusions of the paper.

2 A brief theoretical review

An early study addressing the validity of linear combinations of two covariances was proposed by Ma (2005), without exploring their analytical features. However, the properties concerning the difference between two covariance families were investigated by Posa (2021), as direct consequence of Bochner Theorem (Bochner 1959)[1].

In his very recent papers, Posa has formalized this innovative result and analyzed the conditions required for the difference of two covariance functions to be a valid covariance function in a complex domain, as well in a real domain. Moreover, in Posa (2023a, 2025) the main features of the new families of covariance functions have been detailed in the Euclidean spaces $\mathbb{R}$, $\mathbb{R}^2$ and $\mathbb{R}^3$.

In this Section, the main analytic features of the difference-based covariance models are summarized to point out their flexibility in describing correlation structures that the family of classical covariance functions is unable to model.

The results discussed hereafter represent theoretical and practical advances particularly useful in the modelling stage, since they allow the analysts to identify the most appropriate covariance model when the investigated data present correlation structure with positive and negative values.

I. By recalling Posa Theorem (Posa 2021) and the theoretical findings shown in Posa (2023a), the family of functions

$$C(x; A, B, \alpha, \beta) = Ae^{-\alpha x} - Be^{-\beta x}, \tag{1}$$

with $x = ||\mathbf{x}||, A > 0, B > 0, \alpha > 0, \beta > 0$, is a family of admissible isotropic covariance functions which, according to the relationships among its parameters, can assume in its domain positive or negative values, as well as a linear or parabolic behavior near the origin. In particular, as demonstrated in Posa (2023a), the family of functions in (1) is an admissible class of isotropic covariance functions which assumes always positive values in any dimensional space $\mathbb{R}^m$, if and only if $1 < \frac{\beta}{\alpha} \leq \frac{A}{B}$. Moreover, if $1 < \frac{\beta}{\alpha} < \frac{A}{B}$, the covariance function (1) shows a linear behavior near the origin; instead, if $1 < \frac{\beta}{\alpha} = \frac{A}{B}$ then the covariance function (1) assumes a parabolic behavior near the origin. The same family of functions defined in (1) is a family of admissible covariance functions with negative values in a subset of their domain and linear behavior near the origin, if $1 < \alpha/\beta < \left(A/B\right)^{1/m}$ for any dimensional space $\mathbb{R}^m$ (Posa 2025). In this case, for $x > 0$, the family of covariance functions (1) has a unique zero $x_0 = (\alpha - \beta)^{-1} \cdot ln\left(A/B\right)$ and a minimum value

[1] A continuous function $C : \mathbb{R}^m \to \mathbb{C}$, where $\mathbb{C}$ is the set of the complex numbers and $\mathbb{R}^m$ is the Euclidean m-dimensional space, is a covariance function if and only if it is the Fourier transform of a finite, non negative and non decreasing measure F, i.e., $C(\mathrm{s}) = \int_{\mathbb{R}^m} \exp(i\boldsymbol{\omega}^T \mathrm{s}) dF(\boldsymbol{\omega})$, where i is the imaginary unit.

$$C(x_m) = A\left(\frac{\beta B}{\alpha A}\right)^{\frac{1}{1-\beta/\alpha}} - B\left(\frac{\beta B}{\alpha A}\right)^{\frac{\beta/\alpha}{1-\beta/\alpha}} \quad \text{for} \quad x_m = (\alpha-\beta)^{-1} ln\left(\frac{\alpha A}{\beta B}\right)$$

. From a practical point of view, it is particularly interesting to compute an approximate value for $C(x_m)$ in $\mathbb{R}, \mathbb{R}^2$ and $\mathbb{R}^3$, under the respective admissible conditions. More in detail, by fixing $C(0) = 1$, i.e. $A - B = 1$, the value $C(x_m)$ becomes more and more negative

- as $\frac{\alpha}{\beta} \to \frac{A}{B}$ in $\mathbb{R}$; in this case the minimum value for $C(x_m)$ is approximately equal to -0.1355;
- as $\frac{\alpha}{\beta} \to \left(\frac{A}{B}\right)^{1/2}$ in $\mathbb{R}^2$; in this case the minimum value for $C(x_m)$ is approximately equal to -0.0249;
- as $\frac{\alpha}{\beta} \to \left(\frac{A}{B}\right)^{1/3}$ in $\mathbb{R}^3$; in this case the minimum value for $C(x_m)$ is approximately equal to -0.0061.

Indeed, the value of the negative correlation vanishes as the dimension *m* of the Euclidean space $\mathbb{R}$ increases (Stein 1999; Alegria et al. 2024; De Iaco and Posa 2025).

II. On the basis of Posa Theorem (Posa 2021), given two Gaussian covariance functions, the family of functions

$$C(x; A, B, \alpha, \beta) = Ae^{-\alpha x^2} - Be^{-\beta x^2}, \tag{2}$$

with $x = ||\mathbf{x}||, A > 0, B > 0, \alpha > 0, \beta > 0$, is a family of admissible isotropic covariance functions in $\mathbb{R}^m$ if and only if $1 < \alpha/\beta < (A/B)^{2/m}$. This family of covariance functions has always a parabolic behavior near the origin; moreover, it has negative values in a subset of its domain, a unique zero $x_0 = \sqrt{\frac{1}{\alpha-\beta} ln\left(\frac{A}{B}\right)}$, and a minimum value $C(x_m) = A\left(\frac{\beta B}{\alpha A}\right)^{\frac{\alpha}{\alpha-\beta}} - B\left(\frac{\beta B}{\alpha A}\right)^{\frac{\beta}{\alpha-\beta}}$ at point $x_m = \sqrt{\frac{1}{\alpha-\beta} ln\left(\frac{\alpha A}{\beta B}\right)}$. For the present family of covariance functions, it is straightforward to show that, by fixing $C(0) = 1$, namely $A - B = 1$, the value $C(x_m)$ becomes more and more negative

- as $\frac{\alpha}{\beta} \to \left(\frac{A}{B}\right)^2$ in $\mathbb{R}$; in this case the minimum value for $C(x_m)$ is approximately equal to -0.4453;

– as $\frac{\alpha}{\beta} \to \frac{A}{B}$ in $\mathbb{R}^2$; in this case the minimum value for $C(x_m)$ is approximately equal to -0.1355 (the same minimum value reached by the covariance family defined as difference of two exponential models in $\mathbb{R}$);
– as $\frac{\alpha}{\beta} \to \left(\frac{A}{B}\right)^{2/3}$ in $\mathbb{R}^3$; in this case the minimum value for $C(x_m)$ is approximately equal to -0.0547.

III. According with the Posa Theorem (Posa 2021), the following families of isotropic covariance functions

$$C(x; A, B, \alpha, \beta) = \frac{A}{(x^2+\alpha^2)} - \frac{B}{(x^2+\beta^2)}, \tag{3}$$

$$C(x; A, B, \alpha, \beta) = \frac{A}{(x^2+\alpha^2)^{3/2}} - \frac{B}{(x^2+\beta^2)^{3/2}}, \tag{4}$$

and

$$C(x; A, B, \alpha, \beta) = \frac{A}{(x^2+\alpha^2)^2} - \frac{B}{(x^2+\beta^2)^2}, \tag{5}$$

defined, respectively, in $\mathbb{R}$, $\mathbb{R}^2$ and $\mathbb{R}^3$, with $x = ||\mathbf{x}||$, $A > 0, B > 0, \alpha > 0, \beta > 0$, are admissible covariances with a behavior near the origin which is always parabolic, if and only if $\beta > \alpha$ and $B/A < \beta/\alpha$. Moreover, depending on the relationships among the parameters, the above families of covariance can assume, in their respective domains, positive or negative values. More in detail, the covariance functions (3), (4) and (5) are positive for all $x \in \mathbb{R}$ if it results that $B/A < 1 < \beta/\alpha$; instead, they assume negative values if $1 < B/A < \beta/\alpha$ (Posa 2023a). In this last situation, it is particularly interesting and useful from a practical point of view, to determine the value $x_0 > 0$ at which each of the above covariance classes vanishes, as well as the minimum value $C(x_m)$ for each of the above families. In the following, these features are briefly reviewed from Posa (2023a, 2025).

– In the Euclidean space $\mathbb{R}$, the family of covariances in (3) assumes only one zero, for $x > 0$, at $x_0 = \left(\frac{\alpha^2 B - \beta^2 A}{A - B}\right)^{1/2}$ and then it assumes negative values for all $x > x_0$. Moreover, for $x > 0$, the family of covariances (3) presents a minimum value $C(x_m) = \frac{(k-1)(Ak-B)}{k(\beta^2-\alpha^2)}$ at point $x_m = \left(\frac{\beta^2 - k\alpha^2}{k-1}\right)^{1/2}$, where $k = (B/A)^{1/2}$. For the present family of

covariance functions, it is straightforward to show that, by fixing $C(0) = 1$, namely $A/\alpha^2 - B/\beta^2 = 1$, the above $C(x_m)$ becomes more and more negative as $\beta/\alpha \to 1$ and $B/A \to \beta/\alpha$; in this case the minimum value $C(x_m) \approx -0.1248$.

– In the Euclidean space $\mathbb{R}^2$, the family of covariances in (4) presents only one zero, for $x > 0$, at $x_0 = \left(\frac{\alpha^2 B^{2/3} - \beta^2 A^{2/3}}{A^{2/3} - B^{2/3}}\right)^{1/2}$, then it assumes negative values for all $x > x_0$. For $x > 0$, the family of covariance functions in (4) reaches the minimum value $C(x_m) = \frac{(k-1)^{3/2}(A\,k^{3/2} - B)}{k^{3/2}(\beta^2 - \alpha^2)^{3/2}}$ at point $x_m = \left(\frac{\beta^2 - k\alpha^2}{k-1}\right)^{1/2}$, where $k = (B/A)^{2/5}$. Note that, by fixing $C(0) = 1$, i.e. $A/\alpha^3 - B/\beta^3 = 1$, the minimum value becomes more and more negative as $\beta/\alpha \to 1$ and $B/A \to \beta/\alpha$; in this case the minimum value $C(x_m) \approx -0.0149$.
– In the Euclidean space $\mathbb{R}^3$, the family of covariances in (5) assumes a unique zero, for $x > 0$, at $x_0 = \left(\frac{\alpha^2\sqrt{B} - \beta^2\sqrt{A}}{\sqrt{A} - \sqrt{B}}\right)^{1/2}$ and then it assumes negative values for all $x > x_0$. Moreover, for $x > 0$, the family of covariances (5) presents a minimum value $C(x_m) = \frac{(k-1)^2(A\,k^2 - B)}{k^2(\beta^2 - \alpha^2)^2}$ at point $x_m = \left(\frac{\beta^2 - k\alpha^2}{k-1}\right)^{1/2}$, where $k = (B/A)^{1/3}$. For this family of covariance functions, fixing $C(0) = 1$, namely $A/\alpha^4 - B/\beta^4 = 1$, the minimum value $C(x_m)$ becomes more and more negative as $\beta/\alpha \to 1$ and $B/A \to \beta/\alpha$; in this case the minimum value $C(x_m) \approx -0.0031$.

Some of the aforementioned classes of models with their characteristics will be recalled in the case studies proposed in the following sections, highlighting their flexibility in describing correlation structures with positive and negative values and a unique zero. This last feature represents one of their peculiarity together with the possibility of showing a parabolic or a linear behavior at the origin, as well as positive or negative values according to the parameters values, unlike the traditional wave-hole models which always present a parabolic behavior near the origin and an infinite number of zeros. While the traditional oscillating models can be appropriate to describe spatial stratigraphy, the difference-based models can be used for describing phenomena whose correlation is positive for some lags, then becomes negative and decays, as it will be illustrated in the following case studies. It is worth pointing out that the above mentioned classes of models are smoother than the traditional hole effect models, thus in some examples the minimum value is not close to the theoretical lower bound reached by the J-Bessel models, as discussed in the following section. Evidently, the absolute minimum of a class of covariance models is not the

unique aspect considered by the researcher in choosing the best model for the data at hand: among the different families of covariances with negative values, the selection depends on a set of characteristics, such as the absolute minimum, but also the behavior near the origin, as well the number of zero values.

Note that, a generalization of the above models has been discussed in Posa (2025). In this last work, the author has also provided the analytic features of the models obtained as the product of a polynomial with an exponential.

3 An overview on negative correlation models

As already underlined, in the last decades correlation structures characterized by negative values have been modelled through hole effects covariance functions (Journel and Huijbregts 1978; Journel and Froidevaux 1982; Ma and Jones 2001; Asghari 2015). Although the hole effect model (also called wave-hole model) is not mainstream in Geostatistics, it holds value when a true periodic spatial structure exists (i.e., negative spatial autocorrelation) and the same structure is substantively or physically justified. In such contexts, the above mentioned authors support its use, often in attenuated or composite forms, to maintain realism and stability in modeling. As known, the hole effect models are defined as non-monotonic structures that show cyclic patterns. The presence of hole effect structures, often disregarded, provides valuable information regarding spatial variability, typically indicating a form of cyclicity or periodicity, which is a common and legitimate characteristic in geology (Pyrcz and Deutsch 2003).

The traditional wave-hole models belong to the Bessel family and they are all characterized by a) a parabolic behavior near the origin and b) a countable infinity of zeros (Hristopulos 2020; Posa 2025).

The negative covariance models introduced by Vecchia (1988) are derived from a general class of two-dimensional rational spectral density functions. As a consequence, they are valid only in two-dimensional spaces, differently from the new covariance models obtained by Posa (2023a, 2023b). In addition, the covariance models by Vecchia are obtained evaluating the derivatives of the order zero modified Bessel function of the second kind: these derivatives are not so trivial and they can be computed through some special routines, as underlined by the same author. It is worth highlighting that recently general classes of covariance models have been obtained considering the difference between covariance functions belonging to the Matèrn class, moreover the results are valid for any dimension m of the Euclidean space $\mathbb{R}^m$ (Posa 2025).

Compactly supported correlation functions with negative values, obtained through the product of a compactly supported non-negative correlation function with a Bessel correlation function, have been proposed by Gneiting (2002): as the Bessel family, all these models are characterized by a parabolic behavior near the origin.

In geological and mining applications, as well as in signal processing, the hole effect models have been often defined by combining cosine covariance model with the exponential covariance model (Ma and Jones 2001). In the analysis of brain imaging data (Ye et al. 2015), the use of the hole effect structure in the variogram model

has represented a good approach in terms of prediction results; moreover, hole effect models have also been employed to model concentrations of airborne particulate matter (Alegria et al. 2024).

Other correlation models with negative values can be found in various papers (Shkarofsky 1968; Levinson et al. 1984; Yakhot et al. 1989; Pomeroy et al. 2003; Xu et al. 2003a, b), as well as in Gregori et al. (2008) which provided an example involving the Gaussian and the Matérn family. Ma (2005) analyzed a linear combination of two real isotropic correlation and variogram functions, whereas Hristopulos (2015) derived a peculiar family of infinitely differentiable correlation models which are also characterized by an hole effect.

Further up-to-date contributions have been provided by Hristopulos (2024), who introduced a hybrid spectral technique to obtain covariance kernels and by Alegria et al. (2024) who introduced a hybrid model, which takes into account both long memory and mean square differentiability of the random function. Faouzi et al. (2022) studied Zastavnyi operators acting on rescaled weighted difference between two positive definite radial functions. Alegria and Emery (2026) discussed covariance models that are suitable to achieve negative values of varying intensities depending on the spatial orientation. In the multivariate setting, Alegria and Emery (2024) proposed new parametric families of isotropic matrix-valued functions exhibiting hole effects and cross-dimples associated with local extreme values on the covariance function.

In various applications concerning financial data and agriculture, a negative correlation can be detected in the corresponding time series. The same negative correlation can be found in spatial data concerning medical, biological and physical problems (Griffith 2019; Hu et al. 2018).

It is evident that these traditional hole effect models are not able to describe correlation structures that present positive and decreasing values at small distances (in time or space), only one zero value and, as the distance increases, negative values. On the contrary, the difference-based covariance models are able to catch this type of empirical structure. Note that, similarly to the interpretation of the nested models, when the difference between two covariances is admissible, the adoption of this class of models can be justified by the need to remove the effect of some factors from the correlation structure.

4 Some applications on real datasets

The aim of the present Section is to discuss the advantages of fitting covariance models characterized by linear or parabolic behavior near the origin, negative values and just one zero. Therefore, various case studies are presented and the pro of fitting difference-based covariance models with respect to the traditional models, is highlighted. The discussed case studies are referred to demographic and socio-economic variables defined in $\mathbb{R}$ (one-dimensional space) and in $\mathbb{R}^2$ (two-dimensional space).

A twofold check of the performances of the fitted models has been carried out, namely

(a) a leave-one-out cross-validation procedure, by which in turn each datum has been temporally removed from the dataset and estimated by using the fitted model and the available neighboring data; then the estimates have been compared with the true values,
(b) a jackknife-type estimation procedure developed by predicting data at some time points of a selected sub-set of the observed temporal span (in one-dimensional space case study) or at some locations over a sub-area inside the analyzed spatial domain (in two-dimensional space case studies) where the original data have been previously removed; then the estimates have been compared with the true data.

The performance of the fitted models has been assessed through the computation of some error metrics, such as the mean absolute error (MAE), the root mean squared error (RMSE) and the corresponding relative versions (r-MAE and r-RMSE), obtained by comparing the true values and the estimated ones.

Moreover, as a comparative analysis, alternative traditional covariance models, available in the literature, have been also fitted and used for estimation. Thus, a cross-check with respect to the covariance models generated through the difference has been provided within the above (a) and (b) procedures and the supremacy of these last models has been highlighted.

The proposed applications have confirmed the importance to get flexible models capable to describe and analyze various correlation structures characterized by jointly positive and negative values.

4.1 Times series of emigrants

The investigated time series refers to the yearly data recorded from 1931 to 2004, at the Municipality of Lecce (Italy) for the number of emigrants, i.e. citizens who permanently move to another municipality (source: http://dati.comune.lecce.it).

4.1.1 Structural analysis for emigrants

For the present case study, the sample correlogram has been computed for 17 temporal lags and the obtained values are shown in Fig. 1a (empty black circles). Then a model has been chosen to be fitted to the experimental correlogram, accordingly to its main features. In particular, the estimated correlogram shows positive and decreasing values till the temporal lag approximately equal to 13.5 and then negative values for temporal lags greater than 13.5. In this case, a model based on the difference between covariance functions can be adequate to describe such behavior. Through a visual inspection, the following linear combination of the difference between two exponential covariance functions and the difference between two Gaussian covariance functions has been chosen to fit the sample correlogram:

$$\rho(x) = \begin{cases} 1, & x = 0 \\ 0.90\,\{0.45[A_1 exp(-\alpha_1 x) - B_1 exp(-\beta_1 x)] + \\ +\,0.55[A_2 exp(-\alpha_2 x^2) - B_2 exp(-\beta_2 x^2)]\}, & x \neq 0, \end{cases} \tag{6}$$

with $x = |x|, A_1 = 4, B_1 = 3, \alpha_1 = 1.6/20, \beta_1 = 1.21/20, A_2 = 9, B_2 = 8, \alpha_2 = 0.1/20$ and $\beta_2 = 0.0792/20$.

This model (Fig. 1a) is characterized by a nugget effect equals to 0.10, a linear behavior near the origin, and a unique zero for $x_0 = 11.139$. The model in (6) is able to describe the behavior of the sample correlogram which, for time lags greater than 13.5, assumes negative values. Moreover, this model reaches the minimum value equal to -0.266 at a distance approximately equal to 19 years.

Evidently none of the traditional correlation models developed in the literature is suitable to fit the correlogram estimated for the analyzed time series. Nevertheless, in order to assess the predictive performances of the model (6), a counterpart traditional model has been fitted to the sample correlogram and the comparison has been carried out on the basis of the findings from both the validation procedures (a) and (b), previously detailed.

At this purpose, the following nested model defined by combining an exponential model with temporal range equal to 4 years and a Gaussian model with temporal range equal to 27 years, has been considered:

$$\rho'(x) = \begin{cases} 1, & x = 0 \\ 0.90\left[0.30\,exp\left(\dfrac{-x}{4}\right) + 0.70\,exp\left(\dfrac{-x^2}{27^2}\right)\right], & x \neq 0, \end{cases} \tag{7}$$

with $x = |x|$.

This last model shows a linear behavior near the origin and decreasing values as the temporal distances increase; however, it fails to match the negatives values of the sample correlogram for large time lags. As already pointed out, any other correlation model developed in the literature (Chilès and Delfiner 1999), neither the exponential

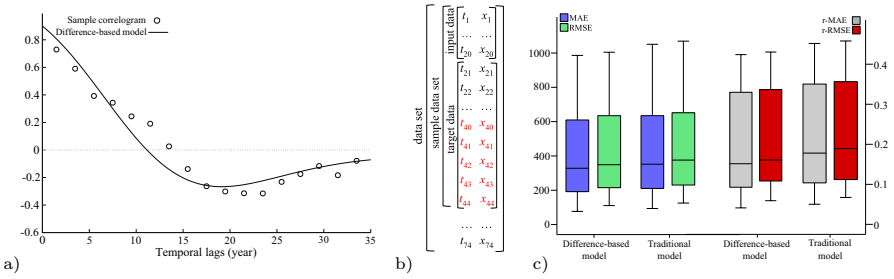

Fig. 1 **a** Sample correlogram with the fitted difference-based model (the horizontal gray dotted line represents the *x* axis); **b** an example of the sampling scheme used for the jackknife kriging predictions (in red the time points whose closest points are associated with the negative correlation values); **c** box plots of the error metrics related to the jackknife predictions obtained with the two fitted models

model which is always positive in the respective domain, nor the traditional hole effect models which present a countable infinity of zeros, are able to fit the sample correlogram computed for the analyzed time series. On the contrary, the difference-based model expressed in (6), which is always negative in a subset of its domain, presents a unique zero and a linear behavior near the origin, adequately describes the temporal behavior shown by the data.

In line with the aim of the present paper, the usefulness of adopting a difference-based model has been underlined from a practical point of view by analyzing the predictive performances of the model, in comparison with those of a traditional model. Therefore, in the following section the two models (6) and (7) have been used for prediction purposes and the findings have been evaluated by the corresponding absolute and relative error metrics.

4.1.2 Predictive performances of the models for emigrants

As previously detailed, a comparative analysis of model in (6) generated by the difference of two covariance functions with respect to the traditional model in (7) has been carried out through a) a leave-one-out cross-validation procedure and b) a jackknife-type estimation. To carry out the above procedures, several sequences of 44 time points have been selected and temporal kriging estimations have been obtained by using the GSLib routines (Deutsch and Journel 1997) which have been properly modified to make predictions with the new classes of models based on differences.

Procedure a) is well known and has been applied by considering alternatively model (6) and model (7), while the implementation of the procedure b) requires some explanation. In particular, for each selected sequence of 44 time points, the data measured at the first 20 time points have been retained, while the data at the 24 subsequent time points have been temporarily removed and considered as target data to be estimated through a jackknife temporal kriging based, alternatively, on model (6) and on model (7). Note that in this jackknife prediction procedure, the negative correlation occurs mainly when estimating the values at the last 4–5 years of the target time span, since their closest data are available at temporal distances at which negative correlation values arise. In Fig. 1b an example of the adopted sampling scheme has been illustrated: the target time points whose closest neighborhoods are associated with the negative correlation values, are depicted in red. The predictions obtained at these last time points by using the two fitted models (6) and (7) have been compared to assess the advantages of using the new class of models in the presence of positive and negative correlation.

Finally, the absolute (MAE and RMSE) and relative (r-MAE and r-RMSE) error metrics constructed with respect to the true values, have been calculated for both models. A summary of these prediction errors are reported in Table 1.

It is worth noting that absolute errors are largely influenced by the order of magnitude of the analyzed data; thus relative errors metrics are generally better suited to compare very different distributions.

The findings have highlighted the superior predictive performance of the difference-based model (6) with respect to the traditional correlation model (7). In Table 1, the percentage relative variations (Δ) between the error metrics referred to the model

(7) and the ones related to the model (6) have been also listed in order to immediately catch the improvement/worsening in the prediction results. Indeed, a positive value of Δ indicates an improvement in predicting by using the difference-based model (6); on the contrary, a negative value of Δ indicates a worsening in prediction results when model (6) is adopted. In the present case study, Δ values have expressed an overall enhancement when the correlation model constructed as difference of covariance functions has been used in the kriging system. The improvement seems slight for the cross-validation, where the beahavior of the correlation function near the origin is prevalent; instead, for the jackknife the percentage variation clearly supports the use of the difference-based model compared to the more conventional one. In particular, the jackknife prediction errors have improved of 5.4% on average when the kriging procedure has been performed with model (6), instead of model (7). This is because in this case in the neighborhood fall points for which the global feature of the correlation function becomes more influential.

A final comparative assessment of the predictive performances of the two different fitted models has been also performed by testing, with the underlying distribution assumptions asymptotically satisfied, the null hypothesis that the difference between the error metrics of the jackknife predictions is on average equal to zero versus the alternative hypothesis that the difference is negative, i.e., the errors are on average lower when the predictions are made by using the difference-based model (6) with respect to the ones obtained with the traditional Gaussian model (7). In particular, the *T*-tests have been performed to compare the mean values of MAE and r-MAE of the jackknife predictions, and the *p*-values of the computed statistical tests, equal respectively to 0.087 and 0.083, have indicated the rejection of the null hypotheses at a significance level of 10%.

All the above findings have confirmed, from a practical point of view, the goodness of the new classes of isotropic covariance models defined as difference between covariance functions (Posa 2021, 2023a).

4.2 Spatial data of population density in Abruzzo

The present case study refers to the population density values (number of inhabitants per square kilometer) registered at the municipalities of an Italian region in 2019 (data collected by the Italian Institute of Statistics - ISTAT - https://demo.istat.it/). In particular, the analyzed data concerns the population density values recorded in 2019 at all 305 municipalities of Abruzzo Region, one of the region located in the central-southern Italy (Fig. 2a).

Table 1 Statistics on the prediction errors of models (6) and (7) for emigrants time series

	Cross-valid predictions	Jackknife predictions	Cross-valid predictions	Jackknife predictions	Cross-valid predictions	Jackknife predictions
	Difference-based model (6)		Traditional model (7)		$\Delta(\%)$	
MAE	163.261	424.896	162.928	449.048	−0.20	5.38
RMSE	216.611	446.738	218.401	471.491	0.82	5.25
r-MAE	0.085	0.203	0.086	0.214	1.18	5.52
r-RMSE	0.112	0.213	0.113	0.225	0.88	5.37

4.2.1 Structural analysis of population density in Abruzzo

For the analyzed dataset, the correlogram has been computed for 7 spatial lags, which have been properly chosen on the basis of the geometry of the spatial domain. From Fig. 2b it is evident that the sample correlogram shows a parabolic behavior near the origin and decreasing values which get negative after the fourth spatial lag and remain negative till the last fixed spatial lag corresponding approximately to 88 km.

In this case, the following model, based on the difference of two Gaussian covariance functions, has been used:

$$\rho(x) = \begin{cases} 1, & x = 0 \\ 0.79\,[A\exp(-\alpha x^2) - B\exp(-\beta x^2)], & x \neq 0, \end{cases} \tag{8}$$

where $x = ||\mathrm{x}||$ with $\mathrm{x} = (x_1, x_2)$ and $A = 9, B = 8, \alpha = 0.128, \beta = 0.114$ and a nugget effect equals to 0.21.

This model is an admissible model in $\mathbb{R}^2$, since it results that $1 < \alpha/\beta < A/B$, moreover it is characterized by a parabolic behavior near the origin, a unique zero $x_0 \approx 2.898$ (28.98 km) and negative values for spatial distances greater than x_0. Furthermore, the above model assumes the minimum value, which is equal to -0.103, at a spatial distance approximately equal to 4.082 (i.e. 40.82 km).

Similarly to the previous case study, none of the traditional correlation models valid in $\mathbb{R}^2$ is suitable to fit the correlogram estimated for the analyzed data. Nevertheless, in order to assess the predictive performances of the model (8), a counterpart traditional model has been considered and the comparison has been carried out on the basis of the findings of both estimation procedures (a) and (b), detailed in Sect. 4.

Thus the following Gaussian correlation model with a spatial range equal to 5 (namely 50 km) and a nugget effect equal to 0.21, has been fitted to the sample correlogram:

$$\rho'(x) = \begin{cases} 1, & x = 0 \\ 0.79\,exp\left(\dfrac{-x^2}{5^2}\right), & x \neq 0, \end{cases} \tag{9}$$

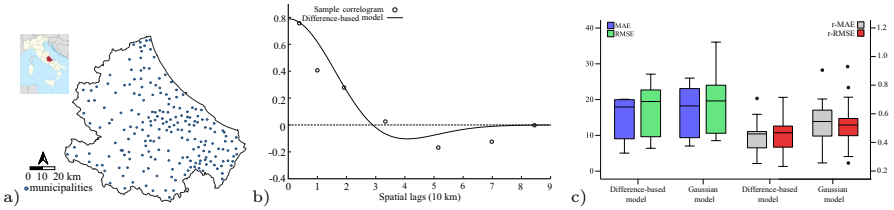

Fig. 2 **a** Location map of the municipalities belonging to Abruzzo Region (central-southern Italy); **b** sample correlogram and fitted model (the horizontal gray dotted line represents the *x* axis); **c** box plots of the error metrics related to the jackknife predictions obtained with the two fitted models

where $x = ||\mathrm{x}||$ with $\mathrm{x} = (x_1, x_2)$.

As for the case study proposed in one-dimensional space, also in the present application the appropriateness of using a difference-based model, rather than a traditional one has been underlined through the predictive performances of the two fitted models, evaluated in terms of the corresponding absolute and relative error metrics.

4.2.2 Predictive performances of the models for population density in Abruzzo

The fitted models (8) and (9) have been evaluated in a comparative way through the results obtained by a) the leave-one-out cross-validation procedure and b) the jackknife-type estimation procedure. As first step, various small subsets of sample points have been randomly selected from the original dataset; then the well known leave-one-out cross-validation procedure has been performed for each selected subset of data by using alternatively model (8) and model (9), successively the absolute and relative error metrics (i.e. MAE, RMSE, r-MAE and r-RMSE) have been computed comparing the estimated and the true values.

As regards the model validation procedure based on jackknife predictions, it can be performed in different ways on the basis of the dataset under study. By taking into account that the usefulness of the difference-based correlation model with respect to a traditional model can be highlighted in predicting data at spatial distances where the negative correlation occurs, in the present case study, the jackknife estimation procedure has been developed as follows. By using each selected subset of data, jackknife predictions of the population density have been obtained at the remaining locations whose true values have been temporarily removed. Then, the comparison between trues values and predicted ones has been made for those isolated points which have been estimated by using data at distances where the negative correlation occurred. Finally, prediction error metrics (absolute, MAE and RMSE, and relative, r-MAE and r-RMSE) have been computed and the results have been listed in Tab. 2.

The findings from both cross-validation and jackknife procedures have shown a better performance of the difference-based model (8) with respect to the Gaussian model (9), with an improvement in the prediction results which is evident especially in the case of the jackknife procedure; indeed, predicting with the traditional correlation model (9) has increased the estimation errors so that the percentage relative variations (Δ index in Table 2) with respect to the prediction errors related to the model (8) are very large ranging from 11.28% to 20.22%. This is due to the fact that in the neighborhood constructed for the jackknife predictions fall points for which the global feature of the correlation function are more predominant.

The distributions of the absolute and relative error indexes related to the jackknife predictions have been also summarized through the box plots shown in Fig. 2c. From these graphs it is evident that the difference-based model (8) performs better with respect to the Gaussian model (9) in terms of median value of the errors distributions, as well as in terms of their respective variability. Moreover, the errors distributions related to the jackknife estimations obtained with the Gaussian model have shown some extreme values (black dots in the Fig. 2c), underlying once again the poor suitability of this model for prediction purposes.

Table 2 Statistics on the prediction errors of models (8) and (9) for population density in Abruzzo

	Cross-valid predictions	Jackknife predictions	Cross-valid predictions	Jackknife predictions	Cross-valid predictions	Jackknife predictions
	Difference-based model (8)		Gaussian model (9)		$\Delta(\%)$	
MAE	78.815	20.567	79.003	23.155	0.24	12.58
RMSE	205.102	24.770	208.702	27.563	1.76	11.28
r-MAE	0.589	0.445	0.591	0.535	0.34	20.22
r-RMSE	0.631	0.460	0.642	0.542	1.74	17.83

Finally, statistical *T*-tests have been carried out to verify, with the underlying distribution assumptions asymptotically satisfied, the null hypothesis that the difference between the error metrics of the jackknife predictions is on average equal to zero versus the alternative hypothesis that the difference is negative (i.e., the errors are on average lower when the predictions are made by using the difference-based model (8) with respect to the ones obtained when the traditional Gaussian model is adopted). Hence, the *T*-tests, performed to compare the mean values of MAE and r-MAE of the jackknife predictions, have indicated the rejection of the null hypotheses at a significance level equal at least 10% (p-values 0.005 and 0.066, respectively).

The above discussed results have confirmed the effectiveness of the correlation models defined as differences between covariance functions: these new families of models are able to describe a correlation structure with jointly positive and negative values and a unique zero, which the traditional correlation models do not catch. As pointed out in the discussed case studies, this limitation of the traditional models is especially noticeable in predicting data at (temporal or spatial) distances where the negative correlation occurs. Hence, the analysts are encouraged to adopt the difference-based correlation models which are characterized by more flexible properties than the ones of the traditional models.

4.3 Spatial data of population density in Veneto

In this case study, the population density values (number of inhabitants per square kilometer) registered at the 571 municipalities belonging to Veneto Region (north-eastern Italy) in 2019 (Fig. 3a), have been analyzed.

4.3.1 Structural analysis of population density in Veneto

Unlike the previous case study related to Abruzzo Region, in Veneto the population density has shown a non-geometric anisotropy which has emerged through the estimation of the directional correlograms. Indeed, the sample correlogram in the east–west direction (an azimuth of 90 degrees) and the one in the north–south direction (an azimuth of 0 degrees) have presented different behaviors and range values, as illustrated in Fig. 3b, c. Both the directional sample correlograms present a linear behavior near the origin, a nugget effect equal to 0.10 and decreasing values which get negatives around 70 km in the east–west direction and 30–35 km in the north–south direction. Moreover, the estimated correlograms are characterized by only one zero. These correlation structures can be adequately modeled by the difference-based

covariance models, since none of the traditional models developed in the literature is capable of catching the above features shown by the sample correlograms.

In particular, in this case study the following models have been adopted:

(a) in the east–west (or horizontal) direction,

$$\rho_H(x) = \begin{cases} 1, & x = 0 \\ 0.90[A_1\, exp(-\alpha_1 x) - B_1\, exp(-\beta_1 x)], & x \neq 0, \end{cases} \tag{10}$$

where $x = |x_1|$ and $A_1 = 2, B_1 = 1, \alpha_1 = 0.325, \beta_1 = 0.23$;

(b) in the north–south (or vertical) direction,

$$\rho_V(x) = \begin{cases} 1, & x = 0 \\ 0.90\,\{0.54\,[A_2\, exp(-\alpha_2 x^2) - B_2\, exp(-\beta_2 x^2)]\} + \\ +\,0.46\,[(1 - 1.1x)\, exp(-1.11x)], & x \neq 0, \end{cases} \tag{11}$$

where $x = |x_2|$ and $A_2 = 7.12, B_2 = 6.12, \alpha_2 = 0.04737, \beta_2 = 0.035$.

Note that the above models are valid in $\mathbb{R}$, since they satisfy the admissible conditions specified in Posa (2023a, 2025) and reviewed in Section 2. Moreover, it is worth noting that model (11) has been constructed by combining the difference of two Gaussian models with a model discussed in Posa (2025) and given by the product of a polynomial with an exponential.

The models in (10) and (11) well catch the behavior of the estimated directional correlograms, as it is evident from their graphical representations in Fig. 3b, c. In particular: a) in the east–west direction, the fitted model presents only one zero for a distance equal approximately to 73 km and a minimum value of −0.023 for a distance corresponding to approximately 109.357 km, b) in the north–south direction, the model used has one zero and a minimum value of −0.218 at distances that are roughly 33.393 km and 60.072 km, respectively. Both models are characterized by a linear behavior near the origin.

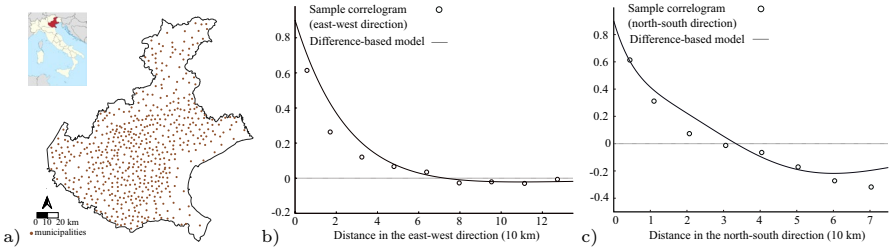

Fig. 3 **a** Location map of the municipalities belonging to Veneto Region (north-eastern Italy); **b** sample correlogram in the east–west direction (an azimuth of 90 degrees) with fitted model; **c** sample correlogram in the north–south direction (an azimuth of 0 degrees) with fitted model (the horizontal gray dotted line represents the *x* axis)

It is worth pointing out that the correlation structures shown by the data along the two directions above discussed, are such that none of the traditional models developed in the literature can adequately describe the estimated correlograms which exhibit a linear behavior near the origin, with one zero, positive and decreasing values at small distances, and negative values as distances between points increase. However, in order to underline the effectiveness of the difference-based models in terms of predictive results, two traditional correlation models have been chosen as the counterparts and referred to as "competing" models of the ones in (10) and (11), respectively:

$$\rho'_H(x) = \begin{cases} 1, & x = 0 \\ 0.90\, exp\left(\frac{-x}{6}\right), & x \neq 0, \end{cases} \qquad \rho'_V(x) = \begin{cases} 1, & x = 0 \\ 0.90\, exp\left(\frac{-x}{4}\right), & x \neq 0. \end{cases} \tag{12}$$

They are both exponential models with nugget effect equal to 0.10, linear behavior near the origin and spatial range equal to 60 km in the east–west direction and 40 km in the north–south direction. Evidently, these models always present positive values as the distance between points increases, thus failing to capture the negative correlation detected for the variable under study, especially in the north–south direction.

Finally, on the basis of the fitted directional models, the non-geometric anisotropy detected for the population density in Veneto has been described by a model defined as the product of the two directional models. Hence, two different product models have been constructed, one by using the two difference-based models in (10) and (11), and the other one by multiplying the two traditional exponential models in (12). The predictive performances of these latter models have been analyzed and compared through the absolute and relative prediction errors and the improvement of one model over the other, in terms of predictions, has been underlined by the percentage relative variations between the error metrics.

The findings from the above mentioned comparative analysis have been reported and discussed in the following section.

4.3.2 Predictive performances of the models for population density in Veneto

In the present case study, the two product models defined, as above clarified, have been used in the spatial kriging procedure, which has been properly implemented to obtain predictions of the population density over the study area.

In particular, the results from a) the leave-one-out cross-validation and b) the jackknife estimation procedure have been assessed. Firstly, some samples of locations have been selected from the available data points. Each sample has been composed of a few spatial points located far away from all the other sample points in the direction characterized by prevailing negative correlation (i.e. about 45–55 km in the north–south direction). These data have been considered as the test data in the jackknife procedure, since their closest neighbors are located at distances where the negative correlation occurs, mainly in the north–south direction, so that the jackknife predic-

tions have been computed only for the test data. Through the leave-one-out cross-validation procedures, all the points of the selected samples have been estimated.

Finally, the absolute and relative error indicators, namely MAE, RMSE, r-MAE and r-RMSE, have been calculated for the predictions obtained by using the two fitted product models and the improvement/worsening of the prediction results has been measured through the percentage relative variation (Δ), as already done in the other case studies discussed in this paper.

From the error values listed in Table 3, it is evident that the product model constructed with the difference-based models in (10) and (11) has better performed with respect to the model defined as the product of the traditional exponential models in (12). In particular, the improvements in prediction results have been higher with the jackkinfe procedure compared to those associated to cross-validation estimation: in the jackknife procedure the Δ index has, in fact, shown values varying between approximately 6.7% and 9.7% (note that positive values of Δ indicate that the predictions computed with the difference-based models have been more reliable than those obtained by using the traditional models). The percentage of variation clearly supports the use of the difference-based model, especially for jackknife predictions where the neighborhood considers points for which the global characteristic of the correlation function is predominant.

The distributions of the absolute and relative error indexes related to the jackknife predictions computed for the various samples selected from the dataset, have been also summarized through the box plots shown in Fig. 4. From these graphs it is evident that the product model computed by using the difference-based models (10) and (11) performs better with respect to the product model based on the exponential models (12), also in terms of median values, especially for the distributions of the relative errors.

An ultimate comparative evaluation of the predictive capabilities of the two distinct product models has been conducted by testing, under the distribution assumptions asymptotically satisfied, the null hypothesis that the difference between the error metrics (MAE and r-MAE) of the jackknife predictions is on average equal to zero, versus the alternative hypothesis that the difference is lower when the predictions are made by using the product of the two difference-based models (10) and (11) with respect to the ones obtained with the product of the traditional models (12). The *T*-tests findings obtained for the comparison of MAE and r-MAE and the *p*-values equal respectively to 0.089 and 0.019, have indicated the rejection of the null hypoth-

Table 3 Statistics on the prediction errors of the two product models constructed for the population density in Veneto

	Cross-valid predictions	Jackknife predictions	Cross-valid predictions	Jackknife predictions	Cross-valid predictions	Jackknife predictions
	Model based on (10) and (11)		Model based on (12)		$\Delta(\%)$	
MAE	92.627	74.630	93.690	81.454	1.15	9.14
RMSE	124.839	87.745	125.897	96.255	0.85	9.70
r-MAE	0.398	0.301	0.399	0.322	0.29	6.98
r-RMSE	0.469	0.343	0.470	0.366	0.39	6.66

eses at a significance level equal at least 10%. This last results have once more confirmed the usefulness of the new models in terms of predictive performances.

5 Simulation study

Through the following simulation study, the estimation capability of the difference-based Gaussian covariance model has been examined and a comparison with the standard Gaussian covariance model has been provided. At this purpose, a total of 3 hundred datasets have been generated in the two-dimensional Euclidean space, over a grid $[0, n_2]/5 \times [0, n_2]/5$, with $n_2 = 19$ (that is, 20×20, with horizontal/vertical step equal to 0.2) (Fig. 5a).

The following difference-based isotropic model

$$C(x) = 20.1\, exp(-7.44x^2) - 19.1\, exp(-7.19x^2) \tag{13}$$

has been defined through the weighted difference of two Gaussian models in $\mathbb{R}^2$ and employed for simulation (Fig. 5b).

The robustness of the model has been assessed using jackknife analyses. Specifically, each simulated dataset has been partitioned into a kriging input dataset (black points in Fig. 5a) and a test dataset (red points in Fig. 5a). The latter represents the jackknife dataset, which includes the points where the kriging estimates have to be computed and the simulated data to be used for comparison. The enhancement in estimation reliability, relative to the standard Gaussian model, has also been assessed for those points whose nearest neighbors lie at distances that correspond to negative values of the newly proposed covariance models (red points enclosed by the dashed line in Fig. 5a). Thus, the differences in jackknife estimation performance, based on the model specified in (13), have been calculated.

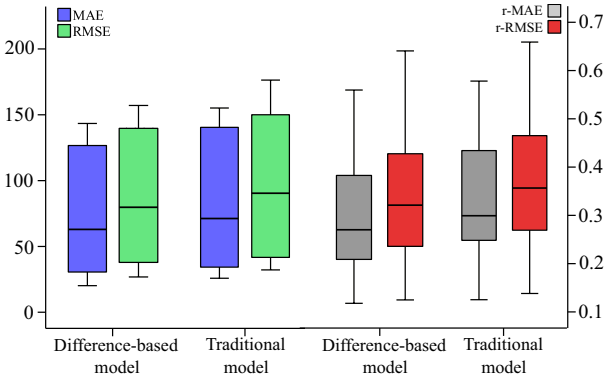

Fig. 4 box plots of the error metrics related to the jackknife predictions obtained with the fitted product models constructed with the difference-based models in (10) and (11), or the traditional exponential models in (12)

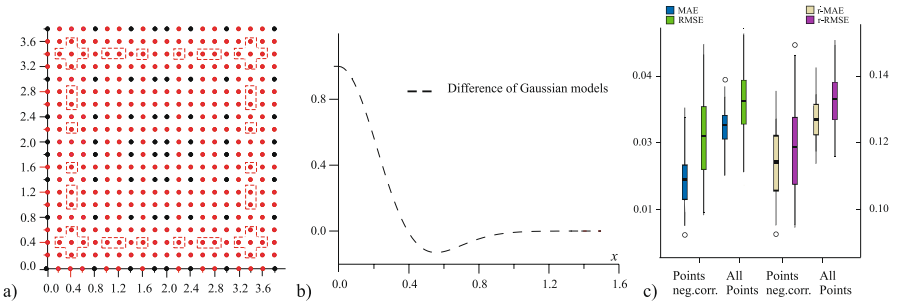

Fig. 5 **a** Geometric structure of the grid used for simulation. The points in black in the simulated dataset are used in the kriging input dataset, while the ones in red belong to the test dataset, and are outlined with a dashed line when their nearest neighbors lie at distances that correspond to negative values of the difference-based covariance models; **b** isotropic covariance model obtained through the weighted difference of Gaussian models; **c** box-plots of the error indexes computed for all the estimated points and for the points delimited by the dashed line in (**a**)

As in the previous case studies, the evaluation of the estimation performance for the models involving negative values has been conducted using the error measures MAE, RMSE, r-MAE and r-RMSE, which have been calculated between the simulated values and the jackknife estimated ones (Table 4). Moreover, for each error indicator, a percentage relative variation (Δ) between the error metrics associated with the estimates – computed either over all test points or only over the subset of points along the dashed line – have been evaluated, in order to emphasize the differences in predictive performance. In particular, negative values of Δ indicate a performance improvement achieved by models that account for negative covariance values. A discussion of the benefits of employing a difference-based covariance model has also drawn on the basis of the statistical *T*-test of the null hypothesis that the error indices are on average equal to zero, as well as on the associated p-values.

Table 4 reports the values of the error indicators for the estimates obtained with model (13) and the corresponding statistics, computed with respect to the variation over the two sets of test points. The findings highlight that the kriging estimates obtained by using the model (13) are accurate and the improvement, measured in terms of Δ, remains negative for all the indices (in the range -31.250, -7.143%). It is worth noting that the deviations of predictive performances are more negative for the points delimited by the dashed line in Fig. 5a (whose closest neighbors are at distances associated to negative values of the adopted difference-based covariance model).

Moreover, the test statistics and p-values offer an additional quantitative evaluation of the improvement achieved in the estimation procedure when the difference-based covariance model in (13) is used, as evidenced by the reduced errors. The p-values (ranging from 0.195 to 0.394) generally imply that the null hypothesis is not rejected at the 5% significance level in both scenarios. Even when considering estimations at all points, the p-values remain above the significance threshold for every error index.

Table 4 Errors related to the estimates obtained through the difference models and some statistics classified with respect to the test set used: the complete test dataset (All points) or the one which includes only the points whose closest neighbors are at distances associated to negative values of the new proposed covariance models (N.C. points)

	Estimate errors		$\Delta(\%)$	Test-Stat (p-value)	
	All points	N.C. points		All points	N.C. points
MAE	0.032	0.022	−31.250	0.213 (0.389)	0.147 (0.394)
RMSE	0.037	0.031	−16.216	0.247 (0.386)	0.207 (0.389)
r-MAE	0.126	0.117	−7.143	1.145 (0.206)	1.064 (0.225)
r-RMSE	0.131	0.119	−9.160	1.191 (0.195)	1.082 (0.221)

6 Conclusions and remarks

In this paper, some theoretical aspects regarding recent classes of difference-based covariance models were revisited and several examples based on spatial and temporal datasets were furnished, with the aim of illustrating the substantial flexibility of these models in capturing both positive and negative correlation. In particular, three different case studies on real-world data characterized by negative correlation structures in $\mathbb{R}$, $\mathbb{R}^2$ in presence of isotropy and $\mathbb{R}^2$ in presence of non-geometric anisotropy, plus a final application on simulated data, were thorough discussed and analyzed from a modeling and prediction perspective. From a practical standpoint, it was emphasized that certain phenomena exhibit negative correlation. This is typically the case, for instance, with population density measured over a geographic region or in time, where the sample covariance usually displays a linear pattern near the origin, then declines to negative values at some spatial or temporal lags, reaches the minimum value and asymptotically goes to zero. In this setting, conventional covariance models fail to capture the estimated covariance function, whereas the recently introduced classes of covariances – constructed as differences of real-valued covariance functions – enable the analysts to represent the sample covariance structure appropriately.

Additional discussions impacting the usefulness of the new models in terms of predictive performances were developed and finally the superior capability and flexibility of these models with respect to the well known covariance models exiting in the literature, were confirmed. Indeed, both the real-world and the simulated case studies well depicted that the predictions which exploit neighbors falling at distances associated to the negative values of the difference-based covariance models, were more reliable than predictions computed with the traditional covariance models. It is worth underlining that the applications discussed in this paper can be particularly useful for many practitioners in choosing the most appropriate family of covariance functions in presence of negative correlation structures.

Finally, a possible comparison of the discussed difference-based models with other traditional models not strictly related to Geostatistics, such as SARMA (Spatial Autoregressive Moving-Average) models, can be arranged as further developments of the empirical analyses shown in this paper.

Acknowledgements The authors thank the Editor and the referees for their suggestions during the reviewing process.

Author contributions The authors have jointly contributed to the conceptualization, editing and revision of the manuscript.

Funding Open access funding provided by Università del Salento within the CRUI-CARE Agreement. This research was supported by National Biodiversity Future Center-NBFC, Spoke 4, Activity 4.1, sub activity 4.1.1. Funder: Project funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4 - Call for tender No. 3138 of 16 December 2021, rectified by Decree n.3175 of 18 December 2021 of Italian Ministry of University and Research funded by the European Union – Next GenerationEU. Award Number: Project code CN_00000033, Concession Decree No. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUP F87G22000290001, Project title "National Biodiversity Future Center-NBFC"; European Union-NextGenerationEU with the Cascade Open Calls published by ALMA MATER STUDIORUM - University of Bologna, inside the Project GRINS funded by PNRR - Mission 4, Component 2, Investment 1.3 "Partnership extended to Universities, Research Centers, Firms and research projects funding", D.D. 341 of 15/03/2022, "ECoST-DATA, Exploring Spatio-Temporal Environmental Conditions: Harmonized Databases and Analytical Techniques", CUP: J33C22002910001; ICSC-National Research Center in High Performance Computing, Big Data and Quantum Computing, funded by European Union-NextGenerationEU; Project name: PNRR-HPC; Project code: CN00000013; CUP: C83C22000560007

Data availibility Data are available on request from the corresponding author

Declarations

Conflict of interest The authors have no conflict of interest to declare that are relevant to the content of this article.

Ethical approval Not pertinent

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alegria A, Emery X (2024) Matrix-valued isotropic covariance functions with local extrema. J Multivar Anal 200:105250

Alegria A, Emery X (2026) Versatile parametric classes of covariance functions that interlace anisotropies and hole effects. Stat Sin 36:1–20

Alegria A, Ramirez F, Porcu E (2024) Hybrid parametric classes of isotropic covariance functions for spatial random fields. Math Geosci 56:1517–1537

Anselin L (1988) Spatial econometrics: methods and models. Kluwer Academic Publishers, Dordrecht, The Netherlands

Asghari O (2015) Geostatistical simulation of dyke systems in sungun porphyry copper deposit. Iran J Mining Environ 6(1):1–10

Bochner S (1959) Lectures on Fourier integrals. Princeton University Press, Colorado

Chilès J, Delfiner P (1999) Geostatistics. Wiley, New York

Christaller W (1966) Central places in Southern Germany. Prentice-Hall, Oxford

Cliff AD, Ord JK (1968) The problem of spatial autocorrelation. University of Bristol, Department of Economics and Department of Geography, Discussion paper
Cressie N (1993) Statistics for spatial data. Wiley, New York
De Iaco S, Posa D (2025) Characteristics of some isotropic covariance models with negative values. Spatial Stat 68:100905
De Iaco S, Myers DE, Posa D (2002) Nonseparable space-time covariance models: some parametric families. Math Geo 34(1):23–41
De Iaco S, Myers DE, Posa D (2011) On strict positive definiteness of product and product–sum covariance models. J Stat Plan Infer 141:1132–1140
De Iaco S, Posa D, Cappello C, Maggio S (2019) Isotropy, symmetry, separability and strict positive definiteness for covariance functions: a critical review. Spat Stat 29:89–108
DeIaco S, Posa D, Cappello C, Maggio S (2020) On some characteristics of Gaussian covariance functions. Int Sta Rev 89:36–53. https://doi.org/10.1111/insr.12403
Deutsch CV, Journel AG (1997) GSLib: geostatistical software library and user's guide, 2nd edn. Oxford University Press, New York
Faouzi T, Porcu E, Kondrashuk I et al (2022) Convergence arguments to bridge Cauchy and Matérn covariance functions. Stat Pap 13(2):1–16
Getis A (2008) A history of the concept of spatial autocorrelation: a geographer's perspective. Geogr Anal 40:297–309
Gneiting T (2002) Compactly supported correlation functions. J Multivar Anal 83:493–508
Gneiting T, Schlather M (2004) Stochastic models that separate fractal dimension and the hurst effect. SIAM Rev 46(2):269–282
Gregori P, Porcu E, Mateau J et al (2008) On potentially negative space time covariances obtained as sum of products of marginal ones. Ann Inst Stat Math 60:865–882
Griffith DA (2016) Spatial autocorrelation. Oxford University Press, New York
Griffith DA (2019) Negative spatial autocorrelation: one of the most neglected concepts in spatial statistics. Stats 2:388–415
Griffith DA, Arbia G (2010) Detecting negative spatial autocorrelation in georeferenced random variables. Int J Geogr Inf Sci 24(3):417–437
Hristopulos DT (2015) Covariance functions motivated by spatial random field models with local interactions. Stoch Environ Res Risk Assess 29:739–754
Hristopulos DT (2020) Random fields for spatial data modeling. Springer, Berlin
Hristopulos DT (2024) Non-separable covariance kernels for spatiotemporal gaussian processes based on a hybrid spectral method and the harmonic oscillator. IEEE Trans Inf Theory. https://doi.org/10.1109/TIT.2023.3321215
Hu L, Griffith DA, Chun Y (2018) Space-time statistical insights about geographic variation in lung cancer incidence rates: Florida, USA, 2000–2011. Environ Res Public Health 15:1–18
Hu L, Chun Y, Griffith DA (2020) Uncovering a positive and negative spatial autocorrelation mixture pattern: a spatial analysis of breast cancer incidences in Broward County, Florida, 2000–2010. J Geogr Syst 22:291–308
Jacob BG, Muturi EJ, Caamano EX (2008) Hydrological modeling of geophysical parameters of arboviral and protozoan disease vectors in internally displaced people camps in Gulu, Uganda. Int J Health Geogr 7(1):11
Journel AG, Huijbregts CJ (1978) Mining geostatistics. Academic Press, New York
Journel AJ, Froidevaux R (1982) Anisotropic hole-effect modeling. Math. Geology 14(3):217–239
Levinson SJ, Beall JM, Powers EJ et al (1984) Space-time statistics of the turbulence in a tokamak edge plasma. Nucl Fusion 24:527–540
Losch A (1954) Die Räumliche Ordnung der Wirtschaft. Gustav Fisher Verlag, 1940. English translation by Stolper, W. F.: The Economics of Location. Yale University Press, New Haven
Ma C (2005) Linear combinations of space-time covariance functions and variograms. IEEE Trans Signal Process 53(3):857–864
Ma YZ, Jones TA (2001) Modeling hole-effect variograms of lithology-indicator variables. Math Geol 33(5):631–648
Matérn B (1980) Spatial variation (2nd ed.), Lecture Notes in Statistics. Springer Verlag, New York
Matheron G (1963) Principles of geostatistics. Econ Geol 58:1246–1266
Moran PAP (1948) The interpretation of statistical maps. J Roy Stat Soc B 10:243–251
Pomeroy JW, Toth B, Granger RJ et al (2003) Variation in surface energetics during snowmelt in a subarctic mountain catchment. J Hydrometeorol 4:702–719

Posa D (2021) Models for the difference of continuous covariance functions. Stoch Environ Res Risk Assess 35:1369–1386
Posa D (2023) Special classes of isotropic covariance functions. Stoch Environ Res Risk Assess 37:1615–1633
Posa D (2023) Revised and wider classes of isotropic space-time covariance functions. Stoch Environ Res Risk Assess 37:4941–4962
Posa D (2025) Covariance models through difference between Matérn families. Stoch Environ Res Risk Assess 39:1345–1358
Pyrcz M, Deutsch C (2003) The whole story on the hole effect. Geostatistical Association of Australasia, Newsletter, p 18
Shkarofsky IP (1968) Generalized turbulence space-correlation and wave-number spectrum-function pairs. Can J Phys 46:2133–2153
Stein ML (1999) Interpolation of spatial data: some theory for kriging. Springer-Verlag, New York
von Thünen JH (1826) Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie, Hamburg, Perthes. English translation by Wartenberg C. M.: von Thunen's Isolated State, 1966. Pergammon Press, Oxford
Vargas-Guzmán JA, Warrick AW, Myers DE (2002) Coregionalization by linear combination of nonorthogonal components. Math Geo 34:405–419
Vecchia AV (1988) Estimation and model identification for continuous spatial processes. J R Stat Soc Ser B Stat Methodol 50(2):297–312
Xu ZW, Wu J, Huo WP et al (2003a) Temporal skewness of electromagnetic pulsed waves propagating through random media with embedded irregularity slab. Chin Phys Lett 20:370–373
Xu ZW, Wu J, Wu ZS (2003b) Statistical temporal behaviour of pulse wave propagation through continuous random media. Waves Random Media 13:59–73
Yaglom AM (1987) Correlation theory of stationary and related random functions, basic results. Springer, Berlin
Yakhot V, Orszag SA, She ZS (1989) Space-time correlations in turbulence - kinematical versus dynamical effects. Phys Fluids 1:184–186
Ye J, Lazard NA, Li Y (2015) Nonparametric variogram modeling with hole effect structure in analyzing the spatial characteristics of fMRI data. J Neurosci Methods 240:101–115

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.