UNIVERSIDADE FEDERAL DO PARANÁ

ANGÉLICA MARIA TORTOLA RIBEIRO

A KRONECKER-BASED MODEL FOR MULTIVARIATE SPATIAL DATA

CURITIBA

2022

ANGÉLICA MARIA TORTOLA RIBEIRO

A KRONECKER-BASED MODEL FOR MULTIVARIATE SPATIAL DATA

Tese apresentada ao curso de Pós-Graduação em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática: Métodos Estatísticos, Setor de Tecnologia, Universidade Federal do Paraná, como requisito parcial para a obtenção do título de Doutora em Ciências.

Orientador: Prof. Dr. Paulo Justiniano Ribeiro Jr

CURITIBA

2022

Catalogação na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

R484k Ribeiro, Angélica Maria Tortola

A Kronecker-Based Model for Multivariate Spatial Data [recurso eletrônico] / Angélica Maria Tortola Ribeiro - Curitiba, 2022.

Tese (doutorado) - Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Setor de Ciências Exatas, Universidade Federal do Paraná.

Orientador: Prof. Dr. Paulo Justiniano Ribeiro Junior

1. Análise de covariância. 2. Geologia – Métodos estatísticos. 3. Estatística matemática. I. Universidade Federal do Paraná. II. Ribeiro Junior, Paulo Justiniano. III. Título.

CDD 526.640285

Bibliotecário: Nilson Carlos Vieira Junior CRB9/1797

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação MÉTODOS NUMÉRICOS EM ENGENHARIA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **ANGÉLICA MARIA TORTOLA RIBEIRO** intitulada: **A KRONECKER-BASED MODEL FOR MULTIVARIATE SPATIAL DATA**, sob orientação do Prof. Dr. PAULO JUSTINIANO RIBEIRO JUNIOR, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutora está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 21 de Fevereiro de 2022.

Assinatura Eletrônica
22/02/2022 17:09:56.0
PAULO JUSTINIANO RIBEIRO JUNIOR
Presidente da Banca Examinadora

Assinatura Eletrônica
22/02/2022 10:27:05.0
FERNANDA DE BASTIANI
Avaliador Externo (UNIVERSIDADE FEDERAL DE PERNAMBUCO)

Assinatura Eletrônica
21/02/2022 17:07:18.0
CESAR AUGUSTO TACONELI
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica
21/02/2022 17:11:18.0
MARCOS OLIVEIRA PRATES
Avaliador Externo (55002642)

RESUMO

Neste trabalho apresentamos uma proposta de especificação de função de covariância para dados multivariados espacialmente contínuos. Este modelo é baseado no produto de Kronecker de matrizes de covariância para campos aleatórios gaussianos. A estrutura é válida para diferentes funções de covariância marginais, possibilitando que diferentes variáveis tenham diferentes estruturas de dependência espacial, o que faz dele mais flexível. O modelo permite que seus parâmetros variem em seus domínios usuais, o que torna a estimação menos restrita quando comparada com outras abordagens clássicas. O tempo computacional para estimação e a fácil generalização para dimensões maiores decorrem de estrutura do modelo. São apresentados resultados teóricos para a função de verossimilhança e as derivadas da matriz de covariância. São relatados resultados de um estudo de simulação, considerando diferentes cenários paramétricos para avaliar as propriedades assintóticas dos estimadores de máxima verossimilhança. Ilustramos o uso do modelo proposto em dois conjuntos de dados da literatura. Para os dados *soil250*, medidas de adequação, erros de previsão e tempos de estimação são comparados com os obtidos com base em modelos clássicos. O conjunto de dados *meuse* é utilizado para ilustrar a flexibilidade do modelo em uma análise para quatro variáveis. A estrutura simples do modelo, aliada à interpretabilidade de seus parâmetros e tempos computacionais para inferência fazem deste modelo um candidato promissor para a análise de dados espaciais multivariados.

Palavras-chaves: Campos aleatórios gaussianos. Dados espaciais multivariados. Funções de covariância. Geoestatística. Modelo de correlação de Matérn.

ABSTRACT

In this work we present a proposal for a covariance function specification for spatially continuous multivariate data. This model is based on the Kronecker product of covariance matrices for Gaussian random fields. The structure is valid for different marginal covariance functions, allowing different variables to have different spatial dependence structures, which makes it more flexible. Our model allows its parameters to vary in its usual domains, which makes the estimation less constrained when compared to other classical approaches. The reduced computational times and easy generalization to larger dimensions follows from the model definition. Theoretical results for the likelihood function and the derivatives of the covariance matrix are presented. A simulation study considering different parametric scenarios to evaluate the asymptotic properties of the maximum likelihood estimators, is conducted. We illustrate the proposed model in the literature datasets: *soil250*, for which adequacy measures, forecast errors and estimation times are compared with those obtained based on classical models, and *meuse*, for illustrate the flexibility of the model in a four-variate analysis. The simple structure of the model, combined with the interpretability of the parameters and computational time for inference make this model a promising candidate for modeling spatially continuous multivariate data.

Key-words: Gaussian random fields. Multivariate spatial data. Covariance functions. Geostatistics. Matérn correlation model.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

The analysis of multivariate random fields have been of interest from the very early days of the geostatistical literature, with an increasing number of proposed approaches as data sets became richer and the ever-increasing computational power. The specification of the covariance structure is central in the estimation and prediction process. Recent contributions include asymmetric models (QADIR et al., 2021; ALEGRÍA; PORCU; FURRER, 2018), modeling on spheres (BEVILACQUA; DIGGLE, et al., 2020; EMERY; PORCU; BISSIRI, 2019; ALEGRÍA; PORCU; FURRER; MATEU, 2019; EMERY; PORCU, 2019), on mapping disease (MARTINEZ-BENEITO, 2020; MACNAB, 2018; MACNAB, 2016), on multinary problems (TEICHMANN et al., 2020), to name a few. However, the definition of a covariance structure that jointly models different variables over a region of the space still presents challenges, especially when the number of variables increases, since obtaining a valid covariance structure becomes more restrictive.

We are interested in multivariate random fields analysis, in the specific context of spatially continuous data. Possible applications cover a wide range of disciplines, such as climatology, meteorology, geophysics, among others, where spatially referenced data is usually of interest. We consider two illustrative examples, one on chemical soil properties relevant for agriculture and the other on heavy metal concentrations within a study area.

Our main goal is to propose a valid covariance specification in the case of spatially continuous multivariate data. The model, presented in Chapter 4, is based on the Kronecker products and it is quite flexible to handle with two or more variables. Furthermore, we present the conditions for the positive definiteness and detail the inferential process of the proposed model.

Different covariance structures for multivariate random fields have been studied by several authors, for example we can mention the linear model of corregionalization (LMC) (GOULARD; VOLTZ, 1992; BOURGAULT; MARCOTTE, 1991; WACKERNAGEL, 2003), which is built as a linear combination of random functions, the class of Matérn cross-covariance functions (GNEITING; KLEIBER, et al., 2010; APANASOVICH et al., 2012), the class of separable models, which considers that the components of the multivariate random field share the same correlation structure (BEVILACQUA; ALEGRÍA, et al., 2016; VALLEJOS et al., 2020) and appears as a parsimonious modeling alternative because it allows a simplification of the more complex models. Genton and Kleiber (2015) and Salvaña and Genton (2020) present a review of the main works in the area. Other recent works on covariance functions for multivariate spatial data can be

found at Ip and Li (2017), Kleiber (2017), Prause, Steland, et al. (2018), Alegría, Porcu, and Furrer (2018) and Bevilacqua, Diggle, et al. (2020), to name a few.

The aforementioned models are widely assessed in the geostatistical literature. Cressie (1993), Gneiting, Kleiber, et al. (2010), Goovaerts et al. (1997), Porcu et al. (2013) and Bevilacqua, Fassò, et al. (2016) noticed some difficulties to handling with the LMC due, for example, to its lack of flexibility and difficulty in recovering the smoothness of latent processes. Separable models are not capable to capture the different scales and smoothness for the variables under study (BEVILACQUA; ALEGRÍA, et al., 2016; BEVILACQUA; VALLEJOS, et al., 2015). The covariance Matérn model presents some restrictions in the parametric space. Vallejos et al. (2020) note that variation of the colocated correlation parameter is constrained by the values of the scale and smoothness parameters, resulting in difficulties for the estimation process and parameter interpretation. Thus, more robust covariance models for multivariate random fields are still a challenge.

The covariance specification presented here emerges as an additional modeling alternative for multivariate random fields that can be more flexible to deal with two or more variables, as it allows the model parameters to vary freely in their usual parametric domains. With a simple construction, the computational implementation has no major difficulties for a number of variables and sampling location points, with a parsimonious estimation computational time, especially for small sample sizes. Furthermore, unlike the separable models, our proposal allows different marginal correlation structures, making it able to capture the structure of each variable.

This work is organized as follows. In Chapter 2 we present a bibliometric study, highlighting relevant information in order to provide an overview of the study area. In Chapter 3 we present a brief literature review of the some basic concepts and we revisit the classical covariance specifications for univariate and multivariate spatial data. In Chapter 4 we present our proposal of specification for multivariate random fields and some theorical results. Through a simulation study, we evaluate the asymptotic properties of the proposed model estimators and the computational times results. The data analyses are presented in Chapter 5. Finally, the main conclusions are summarized in Chapter 6. The model implementation and reported analysis are performed using the computational statistical software R (R CORE TEAM, 2021).

2 BIBLIOMETRIC STUDY

In this Chapter, we detail the steps and the methodological profile of the research conducted in the area of covariance functions for multivariate geostatistical data and present some important results of the bibliometric analysis carried out in order to provide an overview of the study area.

2.1 RESEARCH METHODOLOGY

The research methodology was defined in 3 steps. Initially, we conducted bibliographic searches in the *Web of Science* and *Scopus* bases in August 2021. We consider article type and english language publications, resulting in 108 articles from *Web of Science* and 120 from *Scopus*. Subsequently, 99 duplicated articles were removed, resulting in 129 articles referring to the research performed. Finally, we proceeded to the analysis and main conclusions obtained from the final base of articles. A summary of the research steps is shown in Figure 1.

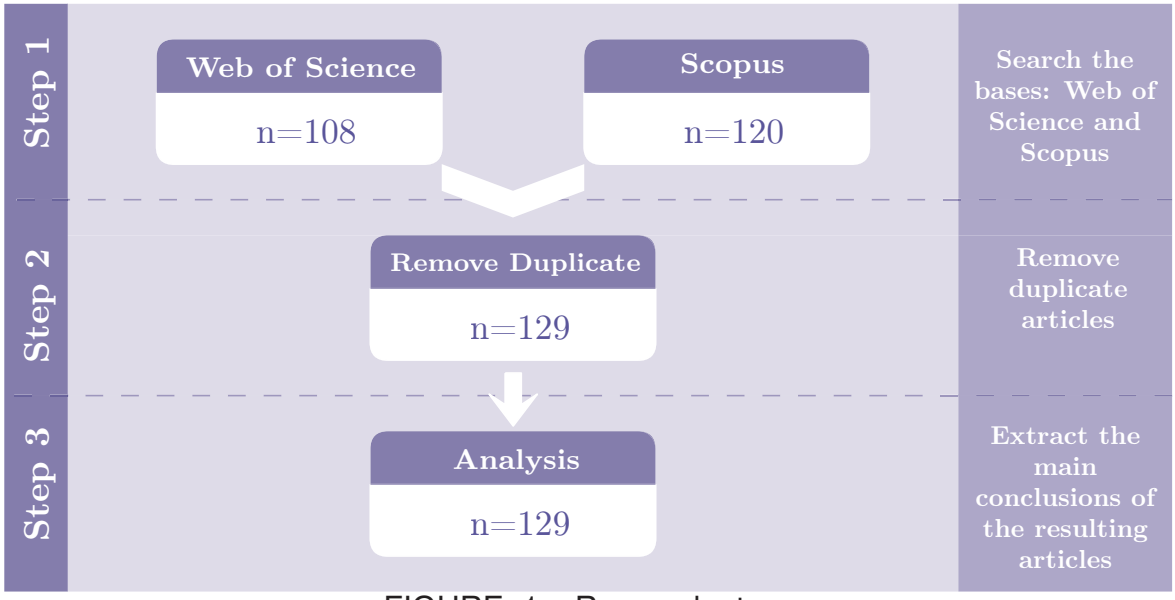

FIGURE 1 – Research steps

In step 1, we considered search *strings* in the fields "title" and "abstract" according to the specific search nomenclatures of each base and using the Boolean operators "AND" and "OR". The entire publication period was considered, which was 1987 to 2021. The table 1 details the search criteria. Bibliometric analysis are presented in subsection 2.2.

TABLE 1 – Search criteria

Item	Especification
Databases	Web of Science and Scopus
Search fields	Title and Abstract
Search string	(Title = ("Multivariate" OR "Bivariate") AND Abstract = ("Geostatistics" OR "Geostatistical" OR "Random Fields" OR "Randomfields") AND Abstract = ("Covariance" OR "Covariances" OR "Cross-Covariance" OR "Cross-Covariances")) OR (Title = ("Geostatistics" OR "Geostatistical" OR "Random Fields" OR "Randomfields") AND Abstract = ("Multivariate" OR "Bivariate") AND Abstract = ("Covariance" OR "Covariances" OR "Cross-Covariance" OR "Cross-Covariances")) OR (Title = ("Covariance" OR "Covariances" OR "Cross-Covariance" OR "Cross-Covariances") AND Abstract = ("Geostatistics" OR "Geostatistical" OR "Random Fields" OR "Randomfields") AND Abstract = ("Multivariate" OR "Bivariate")) OR (Title = ("Covariance" OR "Covariances" OR "Cross-Covariance" OR "Cross-Covariances") AND Title = ("Geostatistics" OR "Geostatistical" OR "Random Fields" OR "Randomfields") AND Title = ("Multivariate" OR "Bivariate"))
Document type	Article
Language	English
Period	1987 to 2021

2.2 BIBLIOMETRIC ANALYSIS

The bibliometric analysis considered the 129 resulting articles from the research steps defined in 1, in order to better understand and describe the researched area. Graphical analyzes were performed using the `Bibliometrix` package (ARIA; CUCCURULLO, 2017) of the R software (R CORE TEAM, 2021).

The world production of articles in the area of covariance functions for multivariate geostatistical data can be seen in Figure 2, which shows the number of articles by country of the corresponding author. Based on this, the country with the highest production in this field are the USA (28 articles), followed by Chile (17 articles), Canada (10 articles) and Saudi Arabia (10 articles) (Table 2).

Each article was also classified as a single country publication (SCP), if all authors are from the same country, or as a multiple country publication (MCP), if there are authors from different countries. Table 2 presents these results for the 10 countries with the highest production of articles in the research area, in which we observe that Chile is the country with the highest international production (8 articles), followed by Australia and Saudi Arabia, both with 4 international articles each. Brazil is in 7th position, with 5 national and 2 international productions.

Performing an analysis of scientific production by journals, the journals with the highest number of articles in the research area are the *Stochastic Environmental*

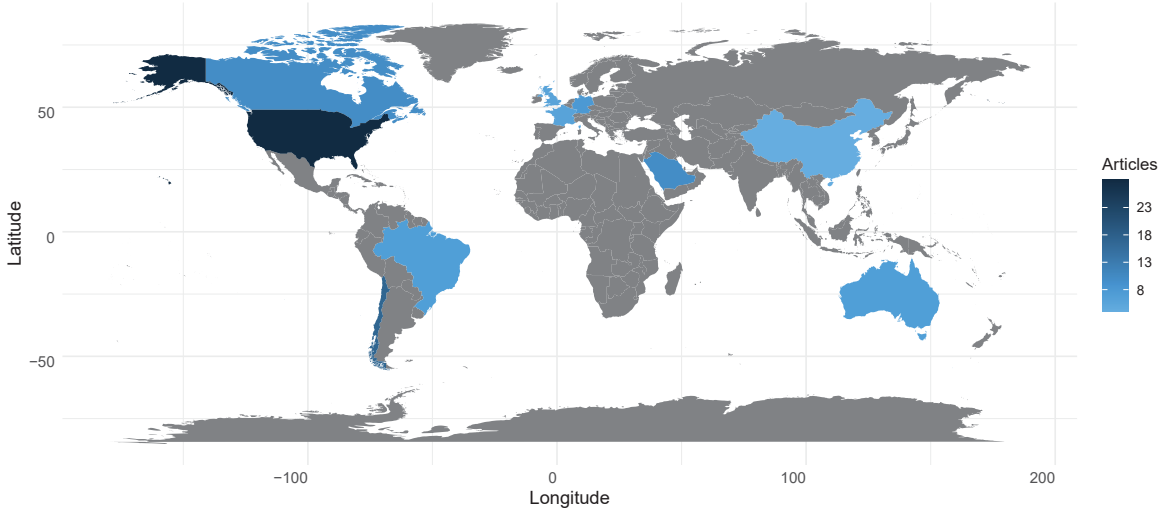

FIGURE 2 – Number of articles by country of the corresponding author

TABLE 2 – The most productive countries in the searched area and its classification into SCP (single country publication) and MCP (multiple country publication)

Country	Articles	Frequency	SCP	MCP
USA	28	0.2258	25	3
CHILE	17	0.1371	9	8
CANADA	10	0.0806	9	1
SAUDI ARABIA	10	0.0806	6	4
GERMANY	8	0.0645	5	3
AUSTRALIA	7	0.0565	3	4
BRAZIL	7	0.0565	5	2
FRANCE	6	0.0484	4	2
UNITED KINGDOM	6	0.0484	5	1
CHINA	4	0.0323	4	0

Research and Risk Assessment (13 articles), followed by *Mathematical Geology* (10 articles) and *Spatial Statistics* (9 articles). Table 3 presents these results for the 10 with the largest production, while Figure 3 shows the smoothed behavior of articles production per journal over the years. We can see that the journals *Spatial Statistics*, *Stochastic Environmental Research and Risk Assessment* and *Mathematical Geosciences*, showed, respectively, the largest growth in article production over the years. In particular the *Spatial Statistics* journal had a more accentuated growth in the production of articles in the research area in recent years.

Figure 4 presents the authors with the highest production in the area. Figure 5 shows the evolution of scientific production in the area over the years. The scientific production of the most productive authors over the years is shown in Figure 6, according to the number of articles and total citations per year (TC per year), in which we observe that Emilio Porcu, Marc G. Genton and Moreno Bevilacqua are the authors with the highest productions (as main authors or co-authors) in the researched area. R codes

TABLE 3 – The journals with the highest production in the researched area

Journals	Articles
STOCHASTIC ENVIRONMENTAL RESEARCH AND RISK ASSESSMENT	13
MATHEMATICAL GEOLOGY	10
SPATIAL STATISTICS	9
MATHEMATICAL GEOSCIENCES	8
BIOMETRIKA	6
JOURNAL OF MULTIVARIATE ANALYSIS	4
ENVIRONMETRICS	3
JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION	3
STAT	3
STATISTICA SINICA	3

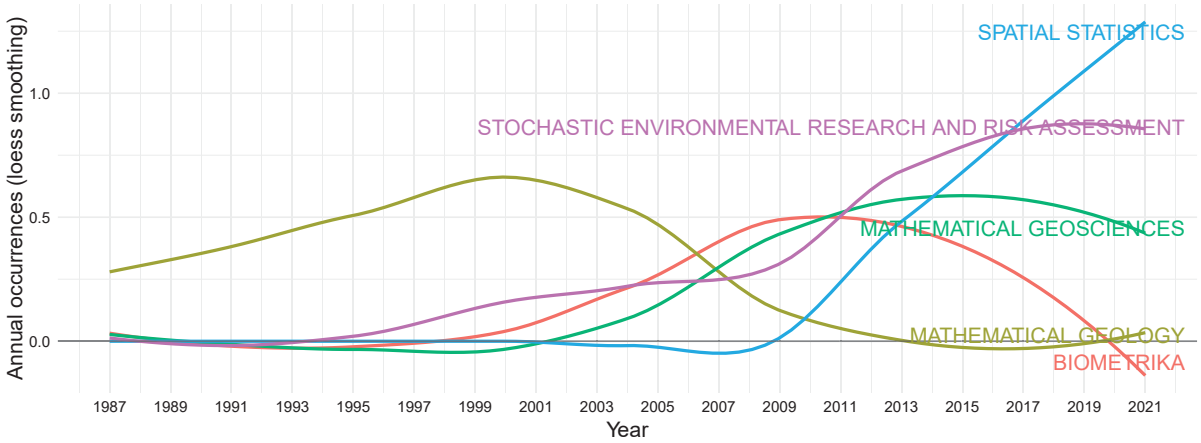

FIGURE 3 – Smoothed behavior of articles production by journal over the years

are available in the supplementary materials.

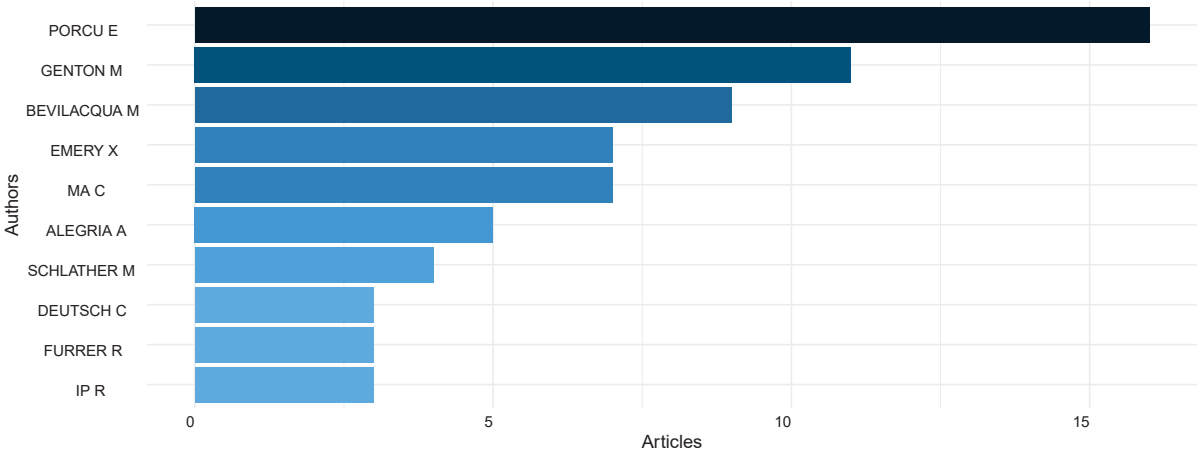

FIGURE 4 – The most productive authors by number of articles

In Chapter 3 we present a literature review of the some basic concepts on geostatistical modeling.

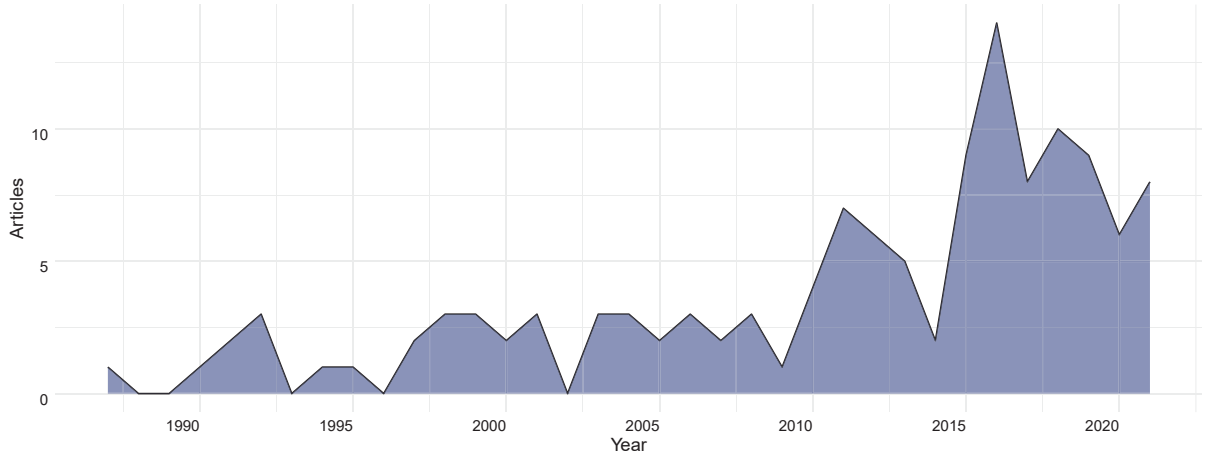

FIGURE 5 – Annual Scientific Production

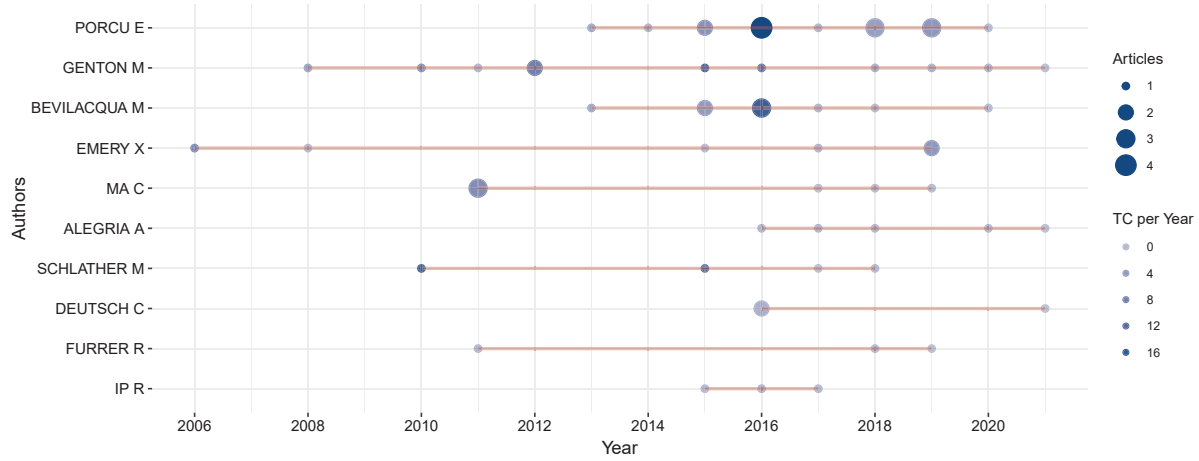

FIGURE 6 – The most productive authors over the years

3 LITERATURE REVIEW

In the spatial data analysis, it is of interest to study how some regionalized variables behave over a region of space. Some situations involve the study of climatology data, meteorology, environmental monitoring, among others. This Chapter presents some basic concepts in order to provide a background on spatial data modeling.

3.1 RANDOM FUNCTIONS

We assume that we have n sample locations $\mathbf{s}_k, k = 1, \ldots, n$, such that $\mathbf{s}_k \in \mathbb{R}^d$, with d being the domain dimension in which the $\mathbf{s}_k$ locations are contained. In each location $\mathbf{s}_k$ we have measures of some amount y, whose values are represented by $y(\mathbf{s}_k)$. The value $y(\mathbf{s}_k)$ represents one of the infinite values that could be observed for the amount y in the location $\mathbf{s}_k$. Therefore, it is usual to represent the infinite possible values of the quantity y in each location $\mathbf{s}_k$, by a random variable $Y(\mathbf{s}_k)$. In addition, these random variables can be seen as a subset of an infinite family of random variables (WACKERNAGEL, 2003) which is called the Random Function (RF) and represented by $\mathbf{Y}(\mathbf{s})$, for any $\mathbf{s} \in \mathbb{R}^d$.

Chilès and Delfiner (2012) presents a RF as a *stochastic process* when the quantity of interest varies in a one-dimensional space, which can be interpreted as time, for example, and is called a *random field* when the quantity of interest varies in a space of a dimension greater than one, for example, in geographical coordinates, which consist of two dimensions.

The notation $Y(\mathbf{s})$ is, in general, used to denote the random function, while $y(\mathbf{s})$ is used to represent its realization.

3.1.1 Probability distribution of an RF

The probability function of a random variable Y at a point $\mathbf{s}_0$, will be given by:

$$F_{\mathbf{s}_0}(y) = P(Y(\mathbf{s}_0) \leq y).$$

The expected value of an RF at a $\mathbf{s}$ point is defined as $E(Y(\mathbf{s}))$, while the covariance function between two different locations $\mathbf{s}_1$ and $\mathbf{s}_2$, is given by:

$$Cov(\mathbf{s}_1, \mathbf{s}_2) = E[Y(\mathbf{s}_1) - E(Y(\mathbf{s}_1))][Y(\mathbf{s}_2) - E(Y(\mathbf{s}_2))].$$

The RF is said to be stationary of second order if for any vector $\mathbf{h} \in \mathbb{R}^d$ separating two points, the first two moments of the random function are stationary, that

is,

$$E(Y(\mathbf{s})) = E(Y(\mathbf{s}+\mathbf{h}))$$
$$Cov(Y(\mathbf{s}), Y(\mathbf{s}+\mathbf{h})) = \boldsymbol{\Sigma}(\mathbf{h}).$$

Thus, the function is second order stationary (also known as weakly stationary) if the mean remains constant and the covariance depends only on the vector $\mathbf{h}$. Some authors refer to this as an invariant translation (CHILÈS; DELFINER, 2012; WACKER-NAGEL, 2003). It is said to be isotropic, if its covariance function does not depend on the orientation of the $\mathbf{h}$ vector, but only on its length $|\mathbf{h}|$.

When the random function is not isotropic it is called anisotropic. According to Diggle and Ribeiro Jr (2007) this anisotropy occurs when there is a differential stretching and rotation of the coordinate axes of the covariance structure.

3.1.2 Gaussian random functions

A random function is called Gaussian, if for any finite collection of locations $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_k$, with $\mathbf{s}_k \in \mathbb{R}^d$, the probability distribution of $Y(\mathbf{s}_1), Y(\mathbf{s}_2), \ldots, Y(\mathbf{s}_k)$ is a multivariate Gaussian. Considering that the first two moments are sufficient to characterize the Gaussian distribution, a Gaussian random function it is also completely characterized by its mean and covariance function. Sometimes, depending on the complexity of the problem, a weak Gaussianity may be necessary. Chilès and Delfiner (2012) cites two ways of building such weak Gaussianity: the first one is to consider bivariate distributions with Gaussian distribution; the second one is when only the marginal distribution of the random function is Gaussian.

3.1.3 Covariance function

As previously specified, the covariance function of a random function is defined by:

$$\begin{aligned}\boldsymbol{\Sigma}(\mathbf{h}) &= Cov(Y(\mathbf{s}), Y(\mathbf{s}+\mathbf{h})) \\ &= E(Y(\mathbf{s}) - E(Y(\mathbf{s})))(Y(\mathbf{s}+\mathbf{h}) - E(Y(\mathbf{s}+\mathbf{h}))).\end{aligned} \tag{3.1}$$

Considering a stationary process, we have $E(Y(\mathbf{s}+\mathbf{h})) = E(Y(\mathbf{s})) = \boldsymbol{\mu}$, and (3.1) can be simplified by:

$$\begin{aligned}\boldsymbol{\Sigma}(\mathbf{h}) &= E(Y(\mathbf{s}) - \boldsymbol{\mu})(Y(\mathbf{s}+\mathbf{h}) - \boldsymbol{\mu}) \\ &= E[Y(\mathbf{s})Y(\mathbf{s}+\mathbf{h})] - \boldsymbol{\mu}E(Y(\mathbf{s})) - \boldsymbol{\mu}E(Y(\mathbf{s}+\mathbf{h})) + \boldsymbol{\mu}^2 \\ &= E[Y(\mathbf{s})Y(\mathbf{s}+\mathbf{h})] - 2\boldsymbol{\mu}^2 + \boldsymbol{\mu}^2 \\ &= E[Y(\mathbf{s})Y(\mathbf{s}+\mathbf{h})] - \boldsymbol{\mu}^2.\end{aligned}$$

In addition, if $\mathbf{h} = \mathbf{0}$, we have:

$$\begin{aligned}\boldsymbol{\Sigma}(\mathbf{0}) &= E[Y(\mathbf{s})Y(\mathbf{s})] - \boldsymbol{\mu}^2 \\ &= E[Y(\mathbf{s})^2] - \boldsymbol{\mu}^2 \\ &= Var(Y(\mathbf{s})).\end{aligned}$$

The covariance function is bounded by the variance, that is,

$$|\boldsymbol{\Sigma}(\mathbf{h})| \leq \boldsymbol{\Sigma}(\mathbf{0}) = Var(Y(\mathbf{s})).$$

The correlation function can be obtained by dividing the covariance function by the variance (WACKERNAGEL, 2003):

$$\boldsymbol{\rho}(\mathbf{h}) = \frac{\boldsymbol{\Sigma}(\mathbf{h})}{\boldsymbol{\Sigma}(\mathbf{0})},$$

where $\boldsymbol{\rho}(\mathbf{h}) \in [-1, 1]$.

An important condition for the covariance function is that it must meet the positive definite condition, that is, $\mathbf{a}^\top \boldsymbol{\Sigma} \mathbf{a} > 0$, for any vector $\mathbf{a} \neq \mathbf{0}$. In this sense, many covariance models have been proposed to represent random functions.

Next, we will present some of the classic covariance models for univariate and multivariate problems and some of their characteristics.

3.2 UNIVARIATE MODELS

3.2.1 Nugget effect covariance function

This is a covariance function that models a discontinuity at the origin and can be defined, for a value of $\tau > 0$, by:

$$\boldsymbol{\Sigma}(\mathbf{h}) = \begin{cases} \tau, & \text{if} \quad |\mathbf{h}| = 0 \\ 0 & \text{if} \quad |\mathbf{h}| > 0. \end{cases}$$

3.2.2 Exponential covariance function

The *exponential covariance function*, with parameters σ^2 and ϕ is defined by:

$$\boldsymbol{\Sigma}(\mathbf{h}) = \sigma^2 \exp\left(-\frac{|\mathbf{h}|}{\phi}\right), \quad \text{with} \;\; \sigma, \phi > 0,$$

where σ^2 represents the variance parameter of the random function, while ϕ represents the correlation length (or scale) parameter, that is, the speed of correlation reduction. The Figure 7 illustrates the exponential covariance function for different values of the ϕ parameter. The ϕ values equal to 0.3, 1.0 and 2.5 were chosen in order to illustrate

the different correlation lengths that can occur in spatial data, from a case with a lower correlation length ($\phi = 0.3$) to a case with a greater correlation length ($\phi = 2.5$).

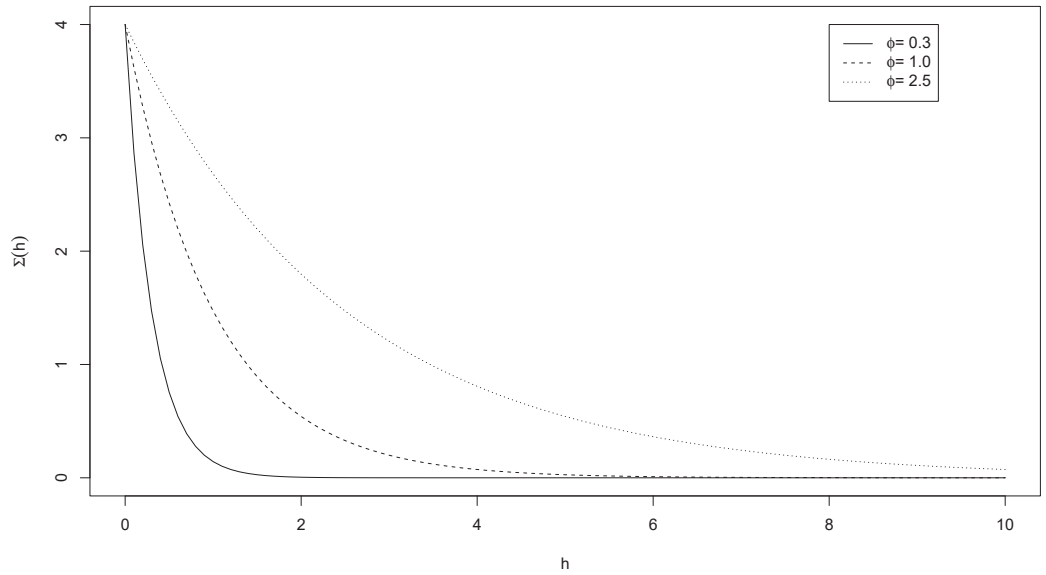

FIGURE 7 – Exponential covariance function for different ϕ values, with $\sigma = 2$.

3.2.3 Matérn covariance function

The *Matérn covariance function* (MATÉRN, 1986) represents a family of covariance functions that has been widely used in recent works and is defined by:

$$\mathbf{\Sigma}(\mathbf{h}) = \sigma^2 M(\mathbf{h}|\nu, 1/\phi),$$

where $M(\mathbf{h}|\nu, 1/\phi) = \frac{2^{1-\nu}}{\Gamma(\nu)} (|\mathbf{h}|/\phi)^{\nu} K_{\nu} (|\mathbf{h}|/\phi)$, is the Matérn spatial correlation at distance $|\mathbf{h}|$, K_{ν} is the modified Bessel function, $\sigma^2, \nu, \phi > 0$ are the variance, smoothness and scale parameters, respectively. When $\nu = 0.5$, the Matérn model reduces to the exponential covariance function and when $\nu \to \infty$, it reduces to the Gaussian covariance function.

This family of distributions is widely used because it includes two important characteristics that are often observed in empirical stationary covariance structures: the reduction of the correlation between two sample points, $Y(\mathbf{s})$ and $Y(\mathbf{s} + \mathbf{h})$, as the distance between these points increases, and the occurrence of different degrees of smoothing that the $Y(\mathbf{s})$ process can present (DIGGLE; RIBEIRO JR, 2007).

More information on this covariance function can be found at Diggle and Ribeiro Jr (2007), Chilès and Delfiner (2012), Matérn (1986) and Wackernagel (2003), to name a few. The Figure 8 illustrates the Matérn covariance function with $\sigma = 2$, $\phi = 0.2$ and considering different values of the ν parameter.

The ν parameter values equal to 0.5, 1.0 and 2.5 were chosen in order to illustrate the different cases of smoothness of the Matérn covariance function, especially

the cases in which it is reduced to exponential covariance function ($\nu = 0.5$) and the case that it has a smoothness closer to the Gaussian covariance function ($\nu = 2.5$). The values of ν and ϕ parameters allow to establish the practical range, that is, the distance from which the correlation becomes very small, typically close to 0.05. In Figure 8, the practical ranges are, respectively, 0.59, 0.79 and 1.18, illustrating different degrees of spatial dependence in the process. We observe that for a fixed ϕ parameter, the spatial dependence tends more quickly to zero for smaller values of ν.

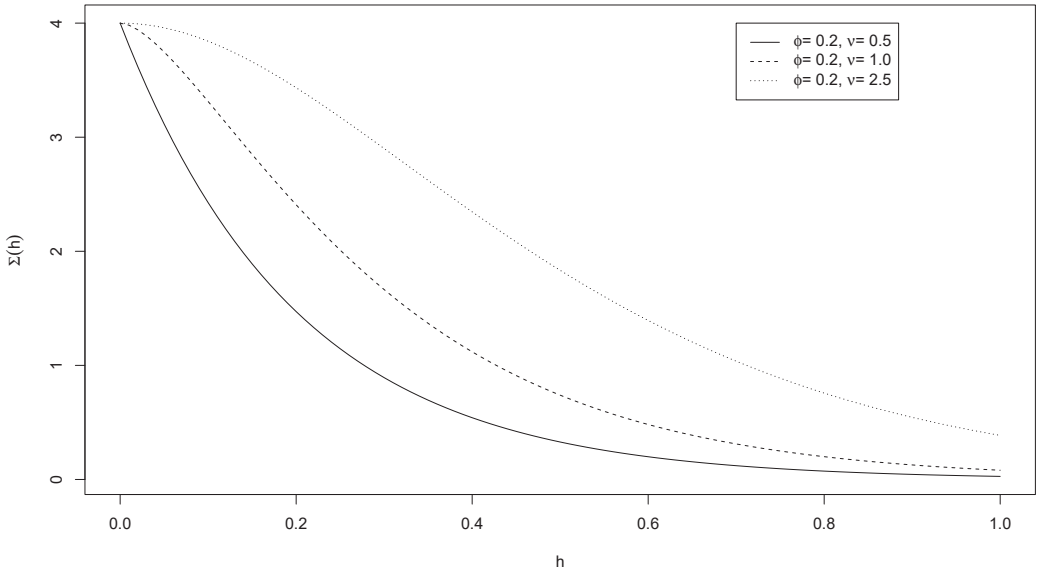

FIGURE 8 – Matérn covariance function for different ν values, with $\sigma = 2$ and $\phi = 0.2$

3.3 MULTIVARIATE MODELS

Problems involving the simultaneous analysis of several variables are increasingly necessary and, with technological advances, obtaining this information has become more efficient. As an example, we can mention the weather stations, which are able to capture information such as temperature, pressure, humidity, wind speed, among others, in the same place and time. Therefore, it is important to have a model that considers the study of all the information of interest simultaneously.

We will use the notation $\mathbf{Y}(\mathbf{s}) = (Y_1(\mathbf{s}), Y_2(\mathbf{s}), \ldots, Y_p(\mathbf{s}))'$ to denote that p RF are considered simultaneously, with mean vector $\boldsymbol{\mu}(\mathbf{s}) = (\boldsymbol{\mu}_1(\mathbf{s}), \boldsymbol{\mu}_2(\mathbf{s}), \ldots, \boldsymbol{\mu}_p(\mathbf{s}))'$, where $\boldsymbol{\mu}_i(\mathbf{s}) = E(Y_i(\mathbf{s}))$, for $i = 1, \ldots, p$, and matrix-valued covariance function:

$$\boldsymbol{\Sigma}(\mathbf{h}) = \text{cov}\{\mathbf{Y}(\mathbf{s_1}), \mathbf{Y}(\mathbf{s_2})\} = [\boldsymbol{\Sigma}_{ij}(\mathbf{h})]_{i,j=1}^{p}, \tag{3.2}$$

where $\mathbf{h} = \mathbf{s_1} - \mathbf{s_2} \in \mathbb{R}^d$ is the spatial separation vector. We consider a stationary and isotropic process (CHILÈS; DELFINER, 2012; DIGGLE; RIBEIRO JR, 2007; GNEITING, 1999). When $i = j$, the functions $\boldsymbol{\Sigma}_{ii}(\mathbf{h})$ in (3.2) describe the spatial variability of the ith process $Y_i(\mathbf{s})$, for $i = 1, \ldots, p$, and are referred as the direct- or marginal-covariance

functions (GENTON; KLEIBER, 2015) and, if $i \neq j$, the functions $\mathbf{\Sigma}_{ij}(\mathbf{h})$ in (3.2) describe the spatial variability between the process $Y_i(\mathbf{s})$ and $Y_j(\mathbf{s})$ and are called as cross-covariance functions.

From the covariance function, it is possible to obtain the correlation function, making:

$$\rho_{ij}(\mathbf{h}) = \frac{\mathbf{\Sigma}_{ij}(\mathbf{h})}{\sigma_i \sigma_j}.$$

The covariance function can be placed in terms of a matrix-valued structure, in which the diagonal matrices are the marginal-covariance functions and outside are the cross-covariance functions. Thus, considering p variables, the covariance function $p \times p$ matrix-valued will be given by:

$$\mathbf{\Sigma}(\mathbf{h}) = \begin{pmatrix} \mathbf{\Sigma}_{11}(\mathbf{h}) & \mathbf{\Sigma}_{12}(\mathbf{h}) & ... & \mathbf{\Sigma}_{1p}(\mathbf{h}) \\ \mathbf{\Sigma}_{21}(\mathbf{h}) & \mathbf{\Sigma}_{22}(\mathbf{h}) & ... & \mathbf{\Sigma}_{2p}(\mathbf{h}) \\ ... & ... & ... & ... \\ \mathbf{\Sigma}_{p1}(\mathbf{h}) & \mathbf{\Sigma}_{p2}(\mathbf{h}) & ... & \mathbf{\Sigma}_{pp}(\mathbf{h}) \end{pmatrix}, \tag{3.3}$$

where each matrix in (3.3) is obtained by (3.2).

If we consider variables centered at zero, the covariance functions in (3.3), can be defined by:

$$\mathbf{\Sigma}_{ij}(\mathbf{h}) = E[Y_i(\mathbf{s}) Y_j(\mathbf{s}+\mathbf{h})].$$

As in the univariate case, the positive definiteness condition must be maintained for (3.3). Next, we present some classical multivariate covariance specifications.

3.3.1 Proportional covariance model

In the *proportional covariance model* (CHILÈS; DELFINER, 2012) the covariance functions are given by:

$$\mathbf{\Sigma}_{ij}(\mathbf{h}) = a_{ij} \mathbf{\Sigma}(\mathbf{h}),$$

where the coefficients a_{ij} make up the ij-th entry of the matrix $\mathbf{A} = [a_{ij}]$, which must be positive definite. This is the simplest covariance model in the literature for the spatial data analysis, and it is the basis for the construction of the Linear Coregionalization Model.

3.3.2 Linear model of corregionalization

A popular model for multivariate random fields is the so-called linear model of coregionalization (LMC) (GOULARD; VOLTZ, 1992; BOURGAULT; MARCOTTE, 1991;

WACKERNAGEL, 2003; CHILÈS; DELFINER, 2012), which is built from univariate covariance functions. Thus, the resulting cross-covariance function takes the form:

$$\mathbf{\Sigma}_{ij}(\mathbf{h}) = \sum_{k=1}^{r} a_{ik}a_{jk}R_k(\mathbf{h}), \quad \text{with} \quad 1 \leq r \leq p,$$

where $R_k(.)$ are valid correlation functions.

The key ideia for the LMC is the overlap of spatial processes in order to induce a multivariate field. This approach is widely explored using the Bayesian approach for inference and prediction. Finley et al. (2015), Banerjee, Carlin, et al. (2003), Gelfand et al. (2004), Schmidt and Gelfand (2003) and Cecconi et al. (2016) are a few examples. We do not pursue Bayesian approaches here.

Despite a relatively simple construction, some authors (CRESSIE, 1993; GENTON; KLEIBER, 2015; GNEITING; KLEIBER, et al., 2010; GOOVAERTS et al., 1997) noticed difficulties to handling with the LMC, such as its lack of flexibility and its difficulty in estimating when the number of variables, and consequently, the number of parameters, increases. Genton and Kleiber (2015) presents two other constructions based on univariate models: convolution methods and latent dimensions.

3.3.3 Multivariate Matérn model

The multivariate Matérn covariance model presented in Gneiting, Kleiber, et al. (2010), considers that the ij-th entry for the covariance function of a multivariate stationary Gaussian random field with zero-mean vector, is given by:

$$\mathbf{\Sigma}(\mathbf{h}) = \left[\rho_{ij}\sigma_i\sigma_j M(\mathbf{h}|\nu_{ij}, 1/\phi_{ij})\right]_{i,j=1}^{p},$$

where $M(\mathbf{h}|\nu, 1/\phi)$, is the Matérn spatial correlation in the distance $|\mathbf{h}|$. When $i = j$, we have the marginal-covariance functions, with $\rho_{ii} = 1$, $\phi_{ii} = \phi_i$ being the marginal scale parameter and $\nu_{ii} = \nu_i$ denoting the marginal smoothness parameter. When $i \neq j$, we have the cross-covariance functions, with ρ_{ij} being the colocated correlation parameter, expressing the marginal correlation between the process Y_i and Y_j, ϕ_{ij} being the cross-scale parameter and ν_{ij} being the cross-smoothness parameter.

According to Gneiting, Kleiber, et al. (2010) a difficulty in this type of modeling is to find the conditions for the parameters ρ_{ij}, ν_{ij} and ϕ_{ij}, which result in a valid covariance model. Gneiting established sufficient and necessary conditions for the parameters of a bivariate random field, whose covariance functions are described by:

$$\begin{aligned}
\mathbf{\Sigma}_{11} &= \sigma_1^2 M(\mathbf{h}|\nu_1, 1/\phi_1), \\
\mathbf{\Sigma}_{22} &= \sigma_2^2 M(\mathbf{h}|\nu_2, 1/\phi_2), \\
\mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{21} &= \rho_{12}\sigma_1\sigma_2 M(\mathbf{h}|\nu_{12}, 1/\phi_{12}),
\end{aligned}$$

which is a valid structure if and only if,

$$\rho_{12}^2 \leq \frac{\Gamma(\nu_1+\frac{d}{2})\Gamma(\nu_2+\frac{d}{2})\Gamma(\nu_{12})^2}{\Gamma(\nu_1)\Gamma(\nu_2)\Gamma(\nu_{12}+\frac{d}{2})^2} \times \frac{\phi_{12}^{4\nu_{12}}}{\phi_1^{2\nu_1}\phi_2^{2\nu_2}} \inf_{t\geq 0} \frac{(1/\phi_{12}^2+t^2)^{2\nu_{12}+d}}{(1/\phi_1^2+t^2)^{\nu_1+\frac{d}{2}}(1/\phi_2^2+t^2)^{\nu_2+\frac{d}{2}}}.$$

The resulting model is called as *full bivariate Matérn model*.

3.3.4 Simplified models

Another important construction called as *separable* considers that the components of the multivariate random field share the same correlation structure (BEVILACQUA; ALEGRÍA, et al., 2016; VALLEJOS et al., 2020). Thus, if $R(\mathbf{h}|\mathbf{\Psi})$ is a univariate correlation function, the *separable* structure can be written as:

$$\mathbf{\Sigma}(\mathbf{h}) = [\rho_{ij}\sigma_i\sigma_j R(\mathbf{h}|\mathbf{\Psi})]_{i,j=1}^{p}, \quad \rho_{ii}=1. \tag{3.4}$$

The class of separable models appears as an alternative of estimation because it allows a simplification of more complex models. For example, considering the Matérn correlation function with ϕ and ν parameters, the *separable Matérn covariance model*, which is a simplified version of the *full* bivariate Matérn model, is:

$$\mathbf{\Sigma}(\mathbf{h}) = [\rho_{ij}\sigma_i\sigma_j M(\mathbf{h}|\nu, 1/\phi)]_{i,j=1}^{2}.$$

Acording to Bevilacqua, Alegría, et al. (2016) and Bevilacqua, Vallejos, et al. (2015) the separable model has some limitations because it cannot capture the different scales and smoothness of the variables under study.

Another simplification of the *full* bivariate Matérn model is when $\phi_{12}=0.5(\phi_1+\phi_2)$ and $\nu_{12}=0.5(\nu_1+\nu_2)$. Such specification is called *bivariate Matérn model with constraints* (BEVILACQUA; MORALES-OÑATE, 2018; VALLEJOS et al., 2020).

The `Geomodels` package of the R statistical software allows the estimation of some bivariate and univariate models (BEVILACQUA; MORALES-OÑATE, 2018).

4 SIMPLER COVARIANCE MODEL

This Chapter presents our proposed covariance specification for multivariate Gaussian random fields, which is based upon Martinez-Beneito (2013). We present the proof of its validity and discuss how to obtain the maximum likelihood estimates of the model parameters.

In Martinez's proposal (MARTINEZ-BENEITO, 2013), the results are presented for modeling multivariate mapping diseases problems based on Gaussian Markov random fields (GMRF), which are discretely indexed, following a Gaussian multivariate distribution with the additional restriction of conditional independence (RUE; HELD, 2005).

Our proposal extends Martinez's approach to construct a covariance function for Gaussian random fields that are continuously indexed, with several applications in geostatistical problems.

We present a simple construction that allows its generalization to larger dimensions more easily. The idea is to write the cross-covariance matrix as a product of matrices that induce variability within processes and between processes, and it is built upon the Kronecker products reformulation of covariance matrices. The resulting construction will be always positive definite for any parameter values in their usual domains.

4.1 MODEL SPECIFICATION

Consider a symmetric correlation matrix $\mathbf{\Sigma}_b$, with dimension $p \times p$, induces a spatial correlation between the processes, while the marginal-covariance functions $\mathbf{\Sigma}_{ii}$, for $i = 1, \ldots, p$, model the variability within each process. We specify the covariance matrix for the $\mathbf{Y}$ process considering the generalized Kronecker product, presented in Martinez-Beneito (2013). Thus, for the Gaussian random fields continuously indexed, the matrix-valued covariance function is defined by:

$$\mathbf{\Sigma}(\mathbf{h}) = \text{Bdiag}\left(\tilde{\mathbf{\Sigma}}_{11}, \tilde{\mathbf{\Sigma}}_{22}, \ldots, \tilde{\mathbf{\Sigma}}_{pp}\right)(\mathbf{\Sigma}_b \otimes \mathbf{I})\,\text{Bdiag}\left(\tilde{\mathbf{\Sigma}}_{11}^{\top}, \tilde{\mathbf{\Sigma}}_{22}^{\top}, \ldots, \tilde{\mathbf{\Sigma}}_{pp}^{\top}\right), \tag{4.1}$$

where, $\tilde{\mathbf{\Sigma}}_{ii}$ is the lower triangular matrix of the Cholesky decomposition of the matrix $\mathbf{\Sigma}_{ii}$, Bdiag represents the matrix in diagonal blocks of the matrices $\tilde{\mathbf{\Sigma}}_{11}$, $\tilde{\mathbf{\Sigma}}_{22}$, ..., $\tilde{\mathbf{\Sigma}}_{pp}$ and $\mathbf{I}$ is the identity matrix. The structure defined in (4.1) is very flexible, allowing different marginal-covariance functions for $\mathbf{\Sigma}_{ii}$ and different correlation structures for $\mathbf{\Sigma}_b$.

Without loss of generality, we will consider that the spatial correlation between the processes will be induced by the correlation matrix:

$$\mathbf{\Sigma}_b = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1p} \\ \rho_{12} & 1 & \dots & \rho_{2p} \\ \dots & \dots & \dots & \dots \\ \rho_{1p} & \rho_{2p} & \dots & 1 \end{pmatrix}, \tag{4.2}$$

where ρ_{ij}, $ij = 1, \dots, p$, is the correlation parameter between the variables i and j.

To quantify the variability within each process, different marginal-covariance structures could be used in (4.1). In a general way, we can write:

$$\mathbf{\Sigma}_{ii}(\mathbf{h}) = \sigma_i^2 R(\mathbf{h}|\mathbf{\Psi}_i), \tag{4.3}$$

where $R(\mathbf{h}|\mathbf{\Psi}_i)$ is a valid correlation function, with $\mathbf{\Psi}_i$ denoting the parameters vector that model the spatial dependence structure of the i-th component, for $i = 1, \dots, p$. For simplification and without loss of generality, we will consider for $R(\mathbf{h}|\mathbf{\Psi}_i)$, the Matérn correlation function. Thus, the marginal covariance function takes the form:

$$\mathbf{\Sigma}_{ii}(\mathbf{h}) = \sigma_i^2 M(\mathbf{h}|\nu_i, 1/\phi_i), \quad \text{for} \quad i = 1, 2, \dots, p. \tag{4.4}$$

The structure specified by (4.1), (4.2) and (4.4) will be called as *simpler multivariate Matérn* (MatSimpler) model. It accepts different marginal behaviors and it is able to handle with different smoothness and scale parameters for each variable. The *simpler multivariate exponential* (ExpSimpler) model is a particular case when the exponential correlation function is used. Ribeiro et al. (2021) illustrates the ExpSimpler model in bivariate analysis of meteorological data.

In Theorem 4.1.1, below, we prove the validity of our covariance specification for modeling multivariate spatial data.

Theorem 4.1.1. *Let $\mathbf{\Sigma}_{ii}$, for $i = 1, \dots, p$, the marginal covariance functions of dimension $n \times n$, $\mathbf{\Sigma}_b$ a valid spatial correlation function of dimension $p \times p$ and $\mathbf{I}$ the identity matrix of dimension $n \times n$, then the covariance function defined in (4.1) is a valid and full rank np specification for multivariate spatial data modeling.*

Proof. Since the marginal covariance functions, $\mathbf{\Sigma}_{ii}$, for $i = 1, \dots, p$, are symmetric positive definite matrix, the matrices $\tilde{\mathbf{\Sigma}}_{ii}$, resulting from the Cholesky decomposition, are lower triangular with positive diagonal elements and therefore, full rank (BANERJEE; ROY, 2014). Let's denote by $R_{\tilde{\mathbf{\Sigma}}_{ii}}$, the rank of $\tilde{\mathbf{\Sigma}}_{ii}$. Thus, rank$(\tilde{\mathbf{\Sigma}}_{ii}) = R_{\tilde{\mathbf{\Sigma}}_{ii}} = n$, for all i, and from the rank properties of the block-diagonal matrix, the rank of any block-diagonal

matrix is the sum of the ranks of its diagonal blocks (BANERJEE; ROY, 2014), that is:

$$\text{rank}\left[\text{Bdiag}\left(\tilde{\boldsymbol{\Sigma}}_{11}^{\top}, \tilde{\boldsymbol{\Sigma}}_{22}^{\top}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}^{\top}\right)\right] = \sum_{i=1}^{p} R_{\tilde{\boldsymbol{\Sigma}}_{ii}} = np,$$

therefore, Bdiag $\left(\tilde{\boldsymbol{\Sigma}}_{11}^{\top}, \tilde{\boldsymbol{\Sigma}}_{22}^{\top}, ..., \tilde{\boldsymbol{\Sigma}}_{pp}^{\top}\right)$ is a full np rank matrix.

On the other hand, since $\boldsymbol{\Sigma}_b$ and $\mathbf{I}$ are positive definite matrices, it follows by the kronecker product properties that $(\boldsymbol{\Sigma}_b \otimes \mathbf{I})$ is also a positive definite matrix (HARDY; STEEB, 2019). With $\boldsymbol{\Sigma}_b$ of dimension p and $\mathbf{I}$ of dimension n, the resulting kronecker product between them will be a positive definite matrix of dimension np.

Now, for simplicity of notation, let's denote by $\mathbf{A}$ and $\mathbf{B}$ the respective $(\boldsymbol{\Sigma}_b \otimes \mathbf{I})$ and Bdiag $\left(\tilde{\boldsymbol{\Sigma}}_{11}^{\top}, \tilde{\boldsymbol{\Sigma}}_{22}^{\top}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}^{\top}\right)$ matrices. Since $\mathbf{A}$ is a positive definite matrix and $\mathbf{B}$ is a full rank matrix, follows that $\mathbf{B}^{\top}\mathbf{A}\mathbf{B}$ preserves not only the rank but also the positive definiteness (GENTLE, 2017; PETERSEN; PEDERSEN, et al., 2008). To visualize this, let $\mathbf{a} \neq \mathbf{0}$, any vector of dimension np, and let $\mathbf{z} = \mathbf{B}\mathbf{a}$, where $\mathbf{z} \neq \mathbf{0}$, because $\mathbf{B}$ is a full rank matrix. Using the positive definite matrix definition, follows:

$$\begin{aligned}
\mathbf{a}^{\top}\left(\mathbf{B}^{\top}\mathbf{A}\mathbf{B}\right)\mathbf{a} &= (\mathbf{a}\mathbf{B})^{\top}\mathbf{A}\left(\mathbf{B}\mathbf{a}\right) \\
&= \mathbf{z}^{\top}\mathbf{A}\mathbf{z} \\
&> 0
\end{aligned}$$

□

The result holds if variables are observed in different numbers of sample locations, since the incomplete data can be treated as missing information and this does not imply any additional complexity and the proof of theorem 4.1.1 remains valid. The dimensions of the resulting covariance matrix will depend on the number of sample locations for each variable considered in the analysis. If we consider marginal covariance functions with dimension n_i, for $i = 1, \ldots, p$, the resulting covariance matrix will have dimension $N = \sum_{i=1}^{p} n_i$.

Considering that $\boldsymbol{\Sigma}$ is a valid covariance specification for any valid choice of marginal-covariance and correlation functions, the proposed model allows its smoothness, scale, correlation and variance parameters to vary in their usual domains, favoring the inferential process and allowing the model parameters to be more easily interpreted.

In our covariance specification, considering the correlation structure defined in (4.2), the matrix-valued covariance function $\boldsymbol{\Sigma}(\mathbf{h})$, can be written in terms of the cross-covariance matrices between the process. Thus, for the ij-th component, with $1 \leq i \neq j \leq p$, the cross-covariance function is, $\boldsymbol{\Sigma}_{ij}(\mathbf{h}) = \rho_{ij}\tilde{\Sigma}_{ii}(\mathbf{h})\tilde{\Sigma}_{jj}(\mathbf{h})^{\top}$ (Eq. A.5). When $i = j$, $\rho_{ii} = 1$ and we have the marginal covariance matrices $\boldsymbol{\Sigma}_{ii} = \tilde{\Sigma}_{ii}\tilde{\Sigma}_{ii}^{\top}$. In Appendix A we detail the matrix calculations that justify these results.

The Theorem 4.1.2 shows that the separable model is a particular case of the simpler covariance model class.

Theorem 4.1.2. *Let $Y_1(\mathbf{s}), Y_2(\mathbf{s}), \ldots, Y_p(\mathbf{s})$ a p-dimensional random field with the same spatial dependence structure for all $Y_i(\mathbf{s})$, $i = 1, \ldots, p$, then the simpler covariance model specified by (4.1), (4.2) and (4.3) is reduced to the class of separable models defined in 3.3.4.*

Proof. From the results presented in Appendix A and considering the general structure of the marginal covariance matrices, defined in (4.3), we see that in the particular case where the process components share the same spatial dependency structure, that is, $\boldsymbol{\Psi}_i = \boldsymbol{\Psi}_j = \boldsymbol{\Psi}$, for all $i, j = 1, \ldots, p$, the simpler covariance specification reduces to the class of separable models, ie,

$$\begin{aligned}
\boldsymbol{\Sigma}_{ij}(\mathbf{h}) &= \rho_{ij}\tilde{\Sigma}_{ii}(\mathbf{h})\tilde{\Sigma}_{jj}(\mathbf{h})^\top \\
&= \rho_{ij}\sigma_i\tilde{R}(\mathbf{h}|\boldsymbol{\Psi})\sigma_j\tilde{R}(\mathbf{h}|\boldsymbol{\Psi})^\top \\
&= \rho_{ij}\sigma_i\sigma_j\tilde{R}(\mathbf{h}|\boldsymbol{\Psi})\tilde{R}(\mathbf{h}|\boldsymbol{\Psi})^\top \\
&= \rho_{ij}\sigma_i\sigma_j R(\mathbf{h}|\boldsymbol{\Psi}), \quad \text{for} \quad i, j = 1, \ldots, p \quad \text{and} \quad \rho_{ii} = 1,
\end{aligned}$$

which is the same structure defined in 3.4. Here, $\tilde{R}(\mathbf{h}|\boldsymbol{\Psi})$ is the lower triangular Cholesky decomposition of the correlation matrix, $R(\mathbf{h}|\boldsymbol{\Psi})$. □

To illustrate the behavior of the proposed model, we simulate a bivariate Mat-Simpler model on a regular unit grid. We set the parameter values as $\sigma_1 = 0.5$, $\sigma_2 = 1.0$, $\nu_1 = 0.5$, $\nu_2 = 0.8$ $\phi_1 = 0.1$, $\phi_2 = 0.2$ and $\rho_{12} = -0.7$ or 0.7, aiming to illustrate different smoothness and variability in the data and situations with positive and negative correlations between the variables. Fig. 9 shows the corresponding simulations.

4.2 ESTIMATION AND INFERENCE

For the estimation process, let $N = np$ and $\mathbf{Y} = \{\mathbf{Y}_1^\top, \ldots, \mathbf{Y}_p^\top\}^\top$, be the $N \times 1$ stacked vector of response variables, with $N \times 1$ mean-vector $\boldsymbol{\mu} = \{\boldsymbol{\mu}_1^\top, \ldots, \boldsymbol{\mu}_p^\top\}^\top$, where $\boldsymbol{\mu}_i = \mathbf{X}_i\boldsymbol{\beta}_i$ denotes the $n \times 1$ vector of expected values for the response variable $\mathbf{Y}_i$, $i = 1, \ldots, p$, with $\mathbf{X}_i$ being a $n \times k_i$ design matrix composed of k_i covariates and, consequently, $\boldsymbol{\beta}_i$ denotes a $k_i \times 1$ regression parameter vector.

We will denote the set of parameters to be estimated by $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\lambda}^\top)^\top$, where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \ldots, \boldsymbol{\beta}_p^\top)^\top$ denotes the regression parameters vector and $\boldsymbol{\lambda} = (\rho_1, \ldots, \rho_{p(p-1)/2}, \sigma_1, \ldots, \sigma_p, \nu_1,$ is the covariance specification parameters vector. Considering $\mathbf{y}$ the stacked vector of observed values, the log-likelihood function for $\boldsymbol{\theta}$ is given by:

$$\mathcal{L}(\boldsymbol{\theta}; \mathbf{y}) = -\frac{1}{2}\left[N\ln(2\pi) + \ln|\boldsymbol{\Sigma}(\boldsymbol{\lambda})| + (\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\beta}))^\top\boldsymbol{\Sigma}(\boldsymbol{\lambda})^{-1}(\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\beta}))\right]. \qquad (4.5)$$

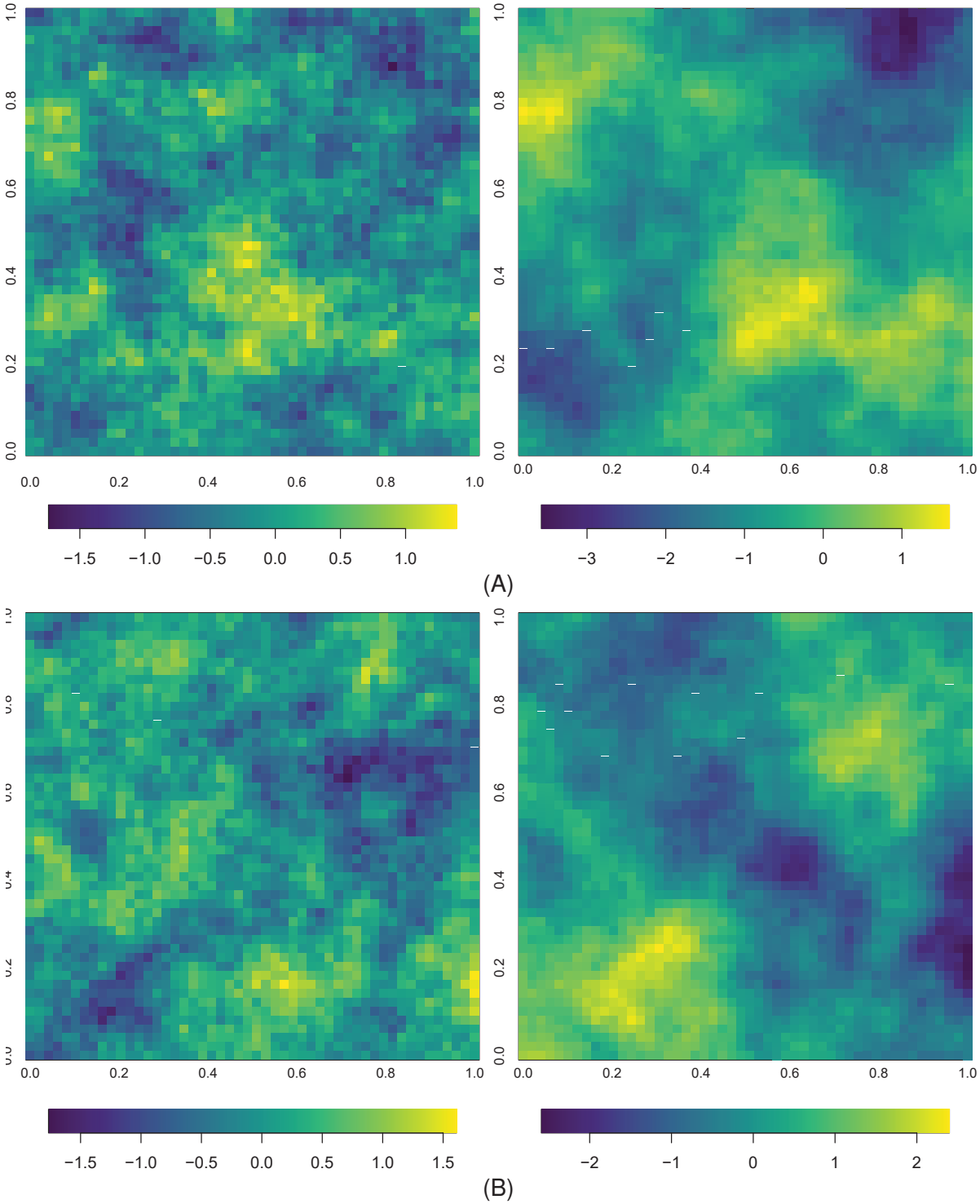

FIGURE 9 – Simulation of a bivariate MatSimpler model with $\sigma_1 = 0.5$, $\sigma_2 = 1.0$, $\nu_1 = 0.5$, $\nu_2 = 0.8$, $\phi_1 = 0.1$, $\phi_2 = 0.2$ and (A) $\rho_{12} = 0.7$ (B) $\rho_{12} = -0.7$

The vector of expected values, $\boldsymbol{\mu}(\beta)$, depends on the regression parameters while the covariance specification $\boldsymbol{\Sigma}(\boldsymbol{\lambda})$ depends on the covariance parameters vector, $\boldsymbol{\lambda}$. We obtain the maximum likelihood estimators by maximizing the function in (4.5) with respect to the $\boldsymbol{\theta}$ parameter vector.

Based on matrix properties (WAND, 2002; BONAT; PETTERLE, et al., 2021), the score functions with respect to the β and $\boldsymbol{\lambda}$ parameters, are given, respectively, by:

$$\mathcal{U}_{\boldsymbol{\beta}} = \mathbf{D}^{\top}\boldsymbol{\Sigma}(\boldsymbol{\lambda})^{-1}\mathbf{r}(\boldsymbol{\beta}),$$

$$\mathcal{U}_{\boldsymbol{\lambda}}(\boldsymbol{\theta}) = -\frac{1}{2}\left\{\boldsymbol{\Sigma}(\boldsymbol{\lambda})^{-1} - \boldsymbol{\Sigma}(\boldsymbol{\lambda})^{-1}\mathbf{r}(\boldsymbol{\beta})\mathbf{r}(\boldsymbol{\beta})^{\top}\boldsymbol{\Sigma}(\boldsymbol{\lambda})^{-1}\right\}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\lambda})}{\partial\boldsymbol{\lambda}}, \tag{4.6}$$

with $\mathbf{r}(\boldsymbol{\beta}) = (\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\beta}))$ and $\mathbf{D} = \frac{\partial\boldsymbol{\mu}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} = \text{Bdiag}(\mathbf{X}_1, \ldots, \mathbf{X}_p)$.

We achieve the maximum likelihood estimator of β by solving $\mathcal{U}_{\boldsymbol{\beta}}$, which results in:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{D}^{\top}\boldsymbol{\Sigma}(\boldsymbol{\lambda})^{-1}\mathbf{D}\right)^{-1}\left(\mathbf{D}^{\top}\boldsymbol{\Sigma}(\boldsymbol{\lambda})^{-1}\mathbf{y}\right).$$

Making similar calculations, we find the Fisher information matrix which, for β, is given by:

$$\mathcal{F}_{\boldsymbol{\beta}} = \mathbf{D}^{\top}\boldsymbol{\Sigma}(\boldsymbol{\lambda})^{-1}\mathbf{D}.$$

For $\boldsymbol{\lambda}$, the $(i,j)^{th}$ entry of the Fisher information matrix, is given by:

$$[\mathcal{F}_{\boldsymbol{\lambda}}]_{ij} = \frac{1}{2}\text{tr}\left[\mathbf{W}_{\lambda_i}\boldsymbol{\Sigma}(\boldsymbol{\lambda})\mathbf{W}_{\lambda_j}\boldsymbol{\Sigma}(\boldsymbol{\lambda})\right],$$

where $\mathbf{W}_{\lambda_i} = -\partial\boldsymbol{\Sigma}(\boldsymbol{\lambda})^{-1}/\partial\lambda_i$.

Considering $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^{\top}, \hat{\boldsymbol{\lambda}}^{\top})^{\top}$ the maximum likelihood estimator of $\boldsymbol{\theta}$ parameter, we have that its asymptotic distribution, with respect to the increase in the density of points in a regular grid with a fixed domain, is $\hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta}, \mathcal{F}_{\theta}^{-1})$, where $\mathcal{F}_{\boldsymbol{\theta}} = \begin{pmatrix} \mathcal{F}_{\boldsymbol{\beta}} & \mathbf{0} \\ \mathbf{0} & \mathcal{F}_{\boldsymbol{\lambda}} \end{pmatrix}$ denotes the Fisher information matrix of $\boldsymbol{\theta}$. This corresponds to the infill asymptotic regime (CRESSIE, 1993).

The maximum likelihood estimates of $\boldsymbol{\lambda}$ can be found through Newton's scoring iterative algorithm (BONAT; PETTERLE, et al., 2021):

$$\boldsymbol{\lambda}^{i+1} = \boldsymbol{\lambda}^{i} - \alpha\mathcal{F}_{\boldsymbol{\lambda}}^{-1}\mathcal{U}_{\boldsymbol{\lambda}}(\tilde{\boldsymbol{\theta}}),$$

where $\tilde{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^{\top}, \boldsymbol{\lambda}^{\top})^{\top}$ and α controls the step length.

In section 4.5 through a simulation study we evaluate the properties of the maximum likelihood estimators for the model in a finite sample scenario.

4.2.1 Derivatives of the covariance matrix

We present in this section the derivative of matrix Σ used in calculations of (4.6). Let ρ_r, for $r = 1, \ldots, p(p-1)/2$, denoting the correlation parameters of $\boldsymbol{\Sigma}_b$, σ_i^2, ϕ_i and ν_i,

denoting the variance, scale and smoothness parameters of the marginal-covariance matrix, $\boldsymbol{\Sigma}_{ii}$, for $i = 1, \ldots, p$.

The partial derivative of the matrix-valued covariance function $\boldsymbol{\Sigma}$, with respect to each correlation parameter, ρ_r, is given by:

$$\frac{\partial \boldsymbol{\Sigma}}{\partial \rho_r} = \text{Bdiag}\left(\tilde{\boldsymbol{\Sigma}}_{11}, \tilde{\boldsymbol{\Sigma}}_{22}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}\right)\left(\frac{\partial \boldsymbol{\Sigma}_b}{\partial \rho_r} \otimes \mathbf{I}\right)\text{Bdiag}\left(\tilde{\boldsymbol{\Sigma}}_{11}^{\top}, \tilde{\boldsymbol{\Sigma}}_{22}^{\top}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}^{\top}\right).$$

To obtain the partial derivative with respect to variance parameter, σ_i^2, we will use matrix properties, that is,

$$\begin{aligned}\frac{\partial \boldsymbol{\Sigma}}{\partial \sigma_i^2} =& \text{Bdiag}\left(\mathbf{0}, \ldots, \frac{\partial \tilde{\boldsymbol{\Sigma}}_{ii}}{\partial \sigma_i^2}, \ldots, \mathbf{0}\right)(\boldsymbol{\Sigma}_b \otimes \mathbf{I})\,\text{Bdiag}\left(\tilde{\boldsymbol{\Sigma}}_{11}^{\top}, \tilde{\boldsymbol{\Sigma}}_{22}^{\top}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}^{\top}\right) \\ &+\text{Bdiag}\left(\tilde{\boldsymbol{\Sigma}}_{11}, \tilde{\boldsymbol{\Sigma}}_{22}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}\right)(\boldsymbol{\Sigma}_b \otimes \mathbf{I})\,\text{Bdiag}\left(\mathbf{0}, \ldots, \frac{\partial \tilde{\boldsymbol{\Sigma}}_{ii}^{\top}}{\partial \sigma_i^2}, \ldots, \mathbf{0}\right).\end{aligned} \tag{4.7}$$

An analogous procedure to the Eq. (4.7) can be used to obtain the derivatives with respect to ϕ_i and ν_i. Thus, to obtain the derivatives of $\boldsymbol{\Sigma}$ with respect to each parameter, we must calculate the parcial derivatives in (4.7). Using the result of partial derivatives of Cholesky's factorization (SÄRKKÄ, 2013; BONAT; JØRGENSEN, 2016), follows:

$$\begin{aligned}\frac{\partial \tilde{\boldsymbol{\Sigma}}_{ii}}{\partial \sigma_i^2} &= \tilde{\boldsymbol{\Sigma}}_{ii}\Phi\left(\tilde{\boldsymbol{\Sigma}}_{ii}^{-1}\frac{\partial \boldsymbol{\Sigma}_{ii}}{\partial \sigma_i^2}\tilde{\boldsymbol{\Sigma}}_{ii}^{-1}\right), \\ \frac{\partial \tilde{\boldsymbol{\Sigma}}_{ii}}{\partial \phi_i} &= \tilde{\boldsymbol{\Sigma}}_{ii}\Phi\left(\tilde{\boldsymbol{\Sigma}}_{ii}^{-1}\frac{\partial \boldsymbol{\Sigma}_{ii}}{\partial \phi_i}\tilde{\boldsymbol{\Sigma}}_{ii}^{-1}\right), \\ \frac{\partial \tilde{\boldsymbol{\Sigma}}_{ii}}{\partial \nu_i} &= \tilde{\boldsymbol{\Sigma}}_{ii}\Phi\left(\tilde{\boldsymbol{\Sigma}}_{ii}^{-1}\frac{\partial \boldsymbol{\Sigma}_{ii}}{\partial \nu_i}\tilde{\boldsymbol{\Sigma}}_{ii}^{-1}\right),\end{aligned}$$

where $\Phi(.)$ is the strictly lower triangular part of the argument and half of its diagonal.

4.3 PREDICTION

For the case of multivariate spatial data, spatial prediction is a generalization of the univariate case that consists of predicting $\mathbf{Y}$ at some unknown location, $\mathbf{s}_0$, based on other sample information, $\mathbf{s}_i$, for $i = 1, \ldots, n$ (VER HOEF; CRESSIE, 1993; BIVAND et al., 2008; PEBESMA, 2004).

Let $\boldsymbol{\Sigma}_{\mathbf{Y}_1}$ be the covariance matrix of $\mathbf{Y}_1 = \mathbf{Y}(\mathbf{s}_i)$, for $i = 1, \ldots, n$, $\boldsymbol{\Sigma}_{\mathbf{Y}_0}$ be the covariance matrix of $\mathbf{Y}_0 = \mathbf{Y}(\mathbf{s}_0)$, $\boldsymbol{\Sigma}_{\mathbf{Y}_1\mathbf{Y}_0}$ be the covariance matrix between $\mathbf{Y}_1$ and $\mathbf{Y}_0$ and $\mathbf{d}_0 = \text{Bdiag}(\mathbf{x}_1(\mathbf{s}_0), \ldots, \mathbf{x}_p(\mathbf{s}_0))$. Then, the best linear unbiased predictor for $\mathbf{Y}_0$ is:

$$E(\mathbf{Y}_0|\mathbf{Y}_1) = E(\mathbf{Y}_0) + \boldsymbol{\Sigma}^{\top}_{\mathbf{Y}_1\mathbf{Y}_0}\boldsymbol{\Sigma}^{-1}_{\mathbf{Y}_1}(\mathbf{Y}_1 - E(\mathbf{Y}_1)),$$

with prediction covariance matrix:

$$Cov(\mathbf{Y}_0|\mathbf{Y}_1) = \boldsymbol{\Sigma}_{\mathbf{Y}_0} - \boldsymbol{\Sigma}^{\top}_{\mathbf{Y}_1\mathbf{Y}_0}\boldsymbol{\Sigma}^{-1}_{\mathbf{Y}_1}\boldsymbol{\Sigma}_{\mathbf{Y}_1\mathbf{Y}_0}.$$

We can replace the unknown parameters of the model by their respective maximum likelihood estimators (MARTINS et al., 2016) and, with this, we obtain the stacked vector prediction for the p variables in the $\mathbf{s}_0$ unobserved locations.

4.4 COMPUTATIONAL RESOURCES

In our computational implementation, we used the `R` statistical software (R CORE TEAM, 2021) for the estimation, simulation and prediction procedures. For the implementation of the proposed covariance matrix specification, we used the `kronecker` function, basic to `R`, which enables the calculation of the kronecker product between the matrices $\boldsymbol{\Sigma}_b$ and $\mathbf{I}$. We use the `Matrix` (BATES; MAECHLER, 2021) package to perform matrix operations more efficiently, such as the Cholesky decomposition of the marginal-covariance matrices, through the `chol` function, and the construction of the diagonal block matrix Bdiag $\left(\tilde{\boldsymbol{\Sigma}}_{11}, \tilde{\boldsymbol{\Sigma}}_{22}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}\right)$, through the `bdiag` function.

To calculate the marginal-covariance functions, we use the `matern` function from `geoR` (RIBEIRO JR et al., 2020) package. Furthermore, the `crossprod` and `tcrossprod` functions allowed a more efficient calculation of products between matrices. To calculate the log-likelihood function, used in the estimation process, we use the `dmvn` function and for simulations we use the `rmvn` function, both from `mvnfast` package (FASIOLO, 2016) that provides computationally efficient methods related to the multivariate normal distribution. For the log-likelihood function optimization, we use `optim` function. R codes are available in on-line supplementary material.

4.5 SIMULATION STUDY

This section presents a simulation study to evaluate the behavior of the maximum likelihood estimators for the model parameters of the MatSimpler model. It also presents some computational time results for the proposed covariance specification.

4.5.1 Behavior of estimators

We explored a bivariate random field in three different scenarios to illustrate different situations that could occur in practice, exemplifying the flexibility of the proposed model in each case. For each scenario, 500 samples of a bivariate zero-mean stationary isotropic Gaussian random field, of sizes 100, 225, 400 and 625, for each variable, were simulated in a regular unit grid. In all scenarios, we set the scale parameters values in $\phi_1 = 0.05$ and $\phi_2 = 0.1$. For the correlation parameter between the variables the values considered are $\rho_{12} = (-0.7, -0.4, 0.0, 0.4, 0.7)$, aiming to illustrate different correlation structures that could occur in practice between the variables, with values ranging from a strong negative correlation ($\rho_{12} = -0.7$) to a strong positive correlation ($\rho_{12} = 0.7$) and including the no correlation case where $\rho_{12} = 0$. Table 4 summarizes the simulated scenarios.

TABLE 4 – Parameter values for each simulated scenario

Scenarios	Situation	Parameters						
		ϕ_1	ϕ_2	ν_1	ν_2	σ_1	σ_2	ρ_{12}
1	Less smoothness and less variability	0.05	0.1	0.3	0.4	0.5	1.0	-0.7
								-0.4
2	Less smoothness and greater variability	0.05	0.1	0.3	0.4	1.5	2.0	0.0
								0.4
3	Greater smoothness and less variability	0.05	0.1	0.7	1.0	0.5	1.0	0.7

Scenario 1 considers the variables have smaller variability, for which we consider relatively small values for the variance parameters: $\sigma_1^2 = 0.25$ and $\sigma_2^2 = 1.0$. Also in this scenario, we consider smaller smoothness for the variables, with values equal to $\nu_1 = 0.3$ and $\nu_2 = 0.4$. When $\nu = 0.5$ the Matérn correlation function reduces to the exponential correlation function. Thus, this scenario represents lesser smoothness marginal behavior for the variables when compared to the exponential correlation model.

Scenario 2 keeps $\nu_1 = 0.3$ and $\nu_2 = 0.4$, and considers a greater variability with variances values equal to $\sigma_1^2 = 2.25$ and $\sigma_2^2 = 4.00$.

Scenario 3 considers a situation of smaller variability, with variance values fixed at $\sigma_1^2 = 0.25$ and $\sigma_2^2 = 1.0$, and considers higher smoothness values for variables: $\nu_1 = 0.7$ and $\nu_2 = 1.0$. In this scenario, we have marginal behaviors that are smoother in comparison with the exponential correlation model.

The practical ranges are different considering the different scenarios and marginal behaviors of the components of the process. For scenarios 1 and 2, this range is 0.12 for variable 1 and 0.27 for variable 2, and so these choices for phis are generating fields that are weakly associated for variable 1 and moderately associated for variable 2. In scenario 3, this range is 0.17 for variable 1 and 0.4 for variable 2, which

also generate fields that are weakly and moderately associated for variables 1 and 2 respectively, however, for this scenario the fields are slightly more associated than compared to scenarios 1 and 2.

The Figure 10 illustrates the marginal covariances for each simulated scenario, which decays to zero as distance increases. Figure 17 in Appendix B presents the cross-covariances for each simulated scenario and each correlation parameter.

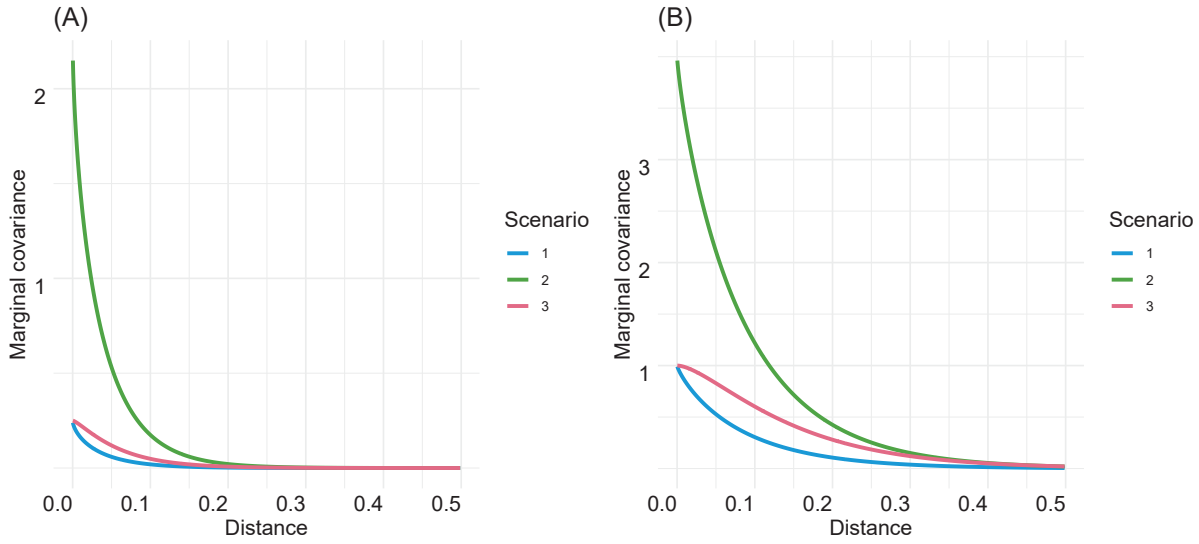

FIGURE 10 – Marginal covariances for (A) first variable and (B) second variable considering the simulated scenarios

Fig. 11 shows the estimated bias plus and minus the expected standard error for estimators of the model for each simulated scenario, based on the asymptotic approximation of the maximum likelihood estimators. To facilitate the visualization in Fig. 11 we follow Bonat and Jørgensen (2016) and Petterle et al. (2019), considering, for each parameter, standardized scales with respect to the expected standard error of the sample size 100, that is, for each parameter, the estimated bias and the limits of the intervals are divided by the expected standard error obtained on the sample of size 100.

The true reference scales, that is, the standard errors obtained in the sample of size 100 used in the standardization of the scales for each scenario and ρ correlation parameter, are presented in table 5.

Fig. 18 in Appendix B presents the expected standard errors for each parametric simulated scenario. Standard errors and biases gets closer to zero as the sample size increases for the considered scenarios. In all scenarios, there appears to be a small overestimate for the smoothness and scale parameters, especially for smaller samples. As expected the results for negative correlations mirrors the positive ones.

Fig. 12 presents the coverage rates for the model parameters, which were obtained by constructing, for each simulated sample and parameter, confidence intervals based on the asymptotic approximation with 95% confidence, and calculating

TABLE 5 – Expected standard errors obtained in the sample of size 100 used in the standardization of the scales for each scenario and ρ correlation parameter

	ρ	Parameters						
		ϕ_1	ϕ_2	ν_1	ν_2	s_1	s_2	ρ_{12}
Scenario 3	-0.7	0.0644	0.0554	1.8791	0.5690	0.0371	0.1164	0.0517
	-0.4	0.0825	0.0599	2.1313	0.7123	0.0382	0.1226	0.0851
	0.0	0.0797	0.0695	1.8846	1.0078	0.0385	0.1302	0.1013
	0.4	0.0738	0.0629	1.8726	0.8101	0.0380	0.1264	0.0848
	0.7	0.0609	0.0564	1.8478	0.5861	0.0372	0.1166	0.0514
Scenario 2	-0.7	0.1176	0.1080	1.8929	0.5261	0.1087	0.1765	0.0510
	-0.4	0.1392	0.1184	2.3429	0.8077	0.1112	0.1780	0.0850
	0.0	0.1316	0.1274	2.4503	0.8717	0.1107	0.1748	0.1008
	0.4	0.1590	0.1045	2.3952	0.7988	0.1091	0.1756	0.0849
	0.7	0.1160	0.0960	2.0315	0.6067	0.1086	0.1690	0.0511
Scenario 1	-0.7	0.1307	0.1075	1.3409	0.5391	0.0363	0.0872	0.0509
	-0.4	0.1458	0.1276	1.7767	0.6768	0.0371	0.0878	0.0849
	0.0	0.1322	0.1465	1.8321	0.7773	0.0368	0.0880	0.1009
	0.4	0.1325	0.1092	1.6620	0.7458	0.0361	0.0855	0.0841
	0.7	0.1045	0.1063	1.6395	0.6705	0.0361	0.0859	0.0511

the proportion of intervals that contained the true value of the parameter. The solid, dashed, and dotted lines represent the coverage rates for the three simulated scenarios, considering each value of the ρ_{12} parameter. The coverage rates vary similarly between simulated scenarios, but without a clear pattern, being close to the 95% level for the simulations performed.

Coverage rates for the scale parameters are slightly above to the 95% level and for the smoothness parameters, these rates tend to fall below to this level, especially for smaller samples. For both scale and smoothness parameters, the coverage rates trend to this level as sample sizes increase. For the variance parameters, these rates fluctuate around the 95% level, especially for σ_1 which represents a situation of smaller variability.

4.5.2 Comparing estimation times

As already mentioned, the MatSimpler model can be extended for more than two variables with relative ease. The computational estimation time will depend on the number of variables and sample locations considered in the analysis. To illustrate the computational time spent on estimation, we implement the generic model for p variables

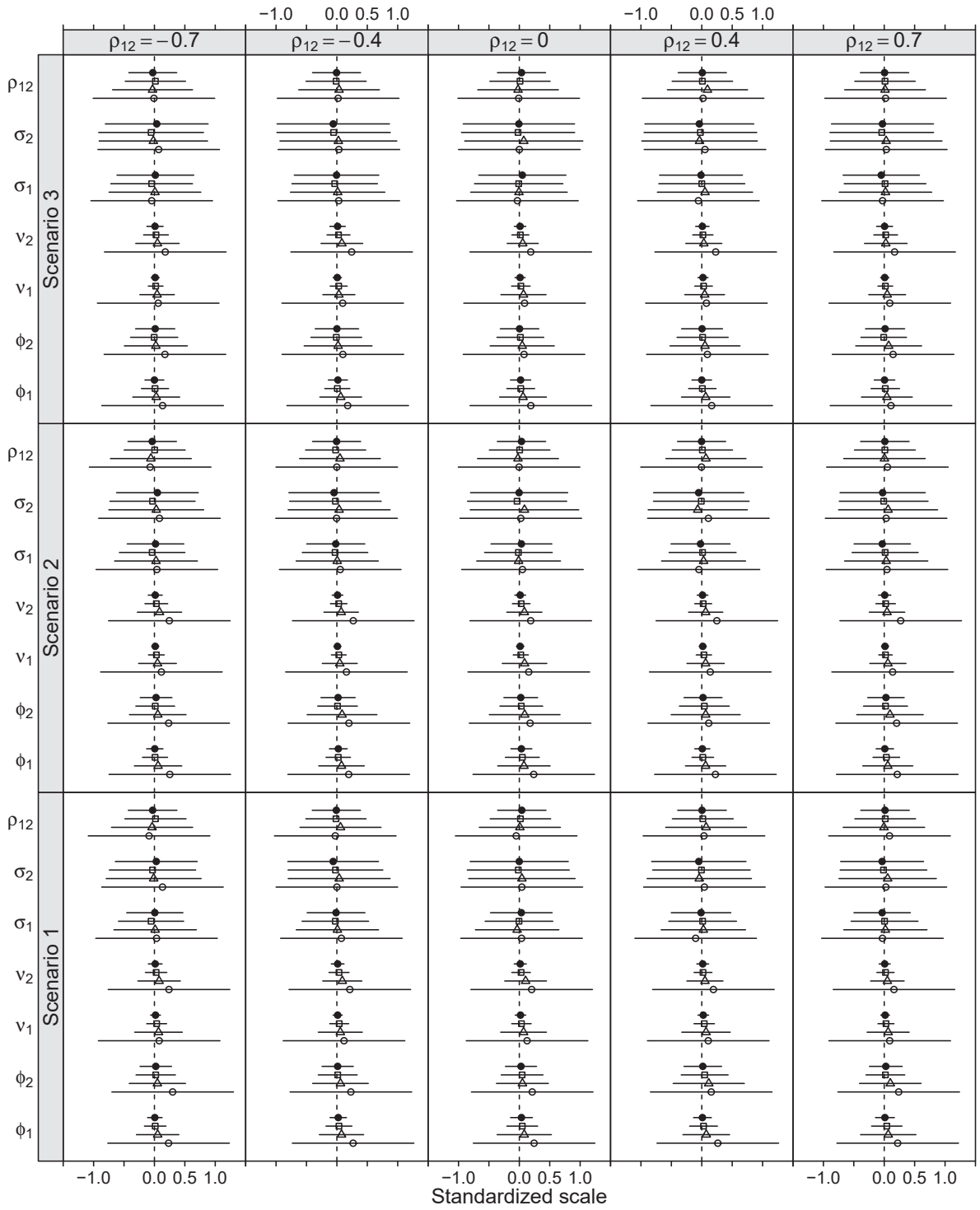

FIGURE 11 – Estimated bias and confidence interval on a standardized scale for each scenario and sample size (○, 100; △, 225; □, 400; ●, 625) for the parameters of the MatSimpler model

and n sample locations and consider the MatSimpler model for different sample sizes and numbers of variables.

We illustrate the case in which the MatSimpler model is reduced to the separable structure, so that the variables share the same spatial dependence structure. We

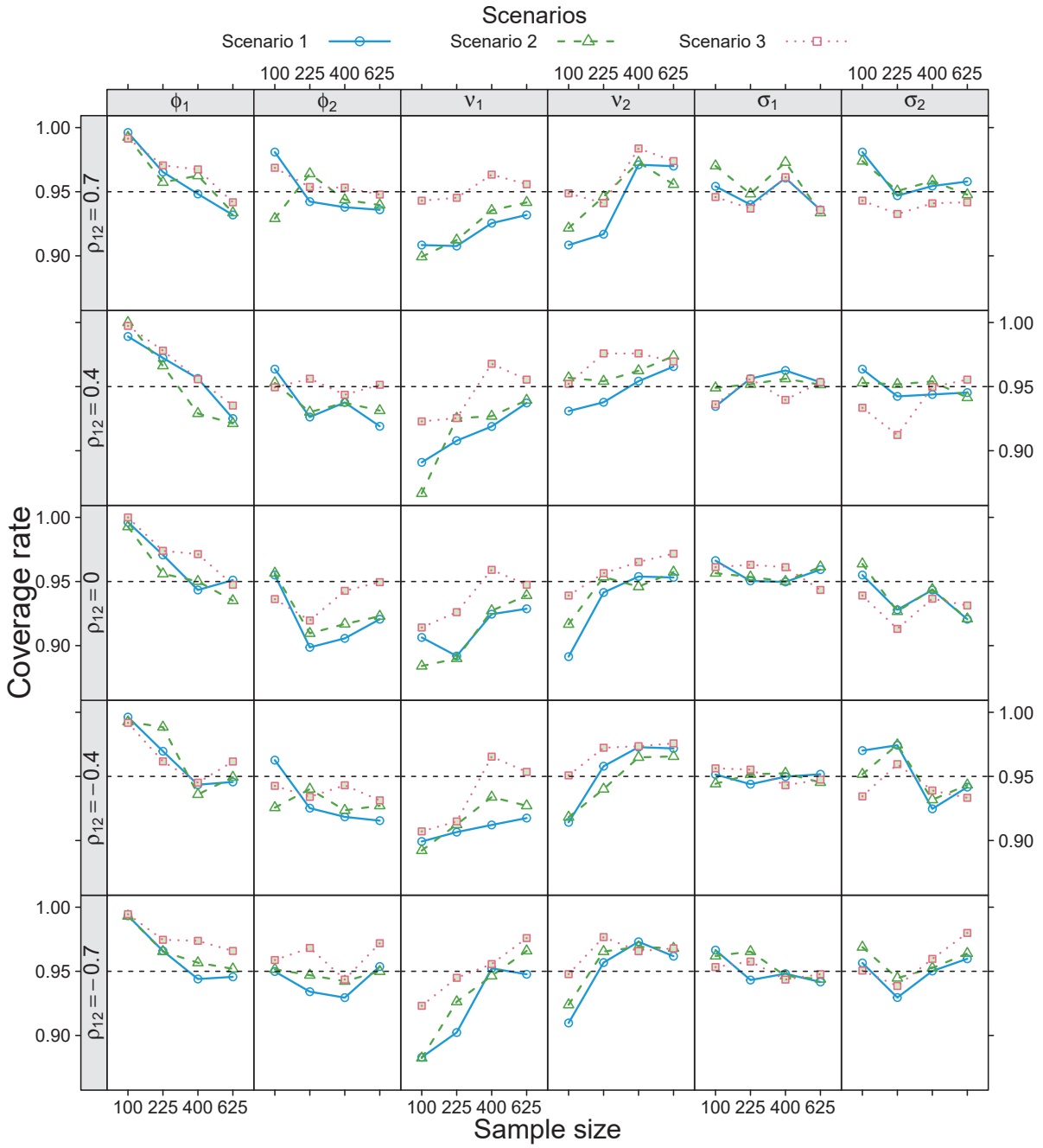

FIGURE 12 – Coverage rates of $\{\phi_1, \phi_2, \nu_1, \nu_2, \sigma_1, \sigma_2\}$ parameters for a bivariate Mat-Simpler model considering different simulated parametric scenarios and different sample sizes

consider the number of variables, p, ranging from 2 to 6 and the number of sample locations, $n = (100, 225, 400, 625, 900)$, taken in a unit square grid.

We set $\phi_i = 0.2$, $\nu_i = 0.5$, $\sigma_i = 0.3$, for all $i = 1, \ldots, p$. In this case, the variables have a smoothness that corresponds to the exponential model. The correlation parameters were chosen between -0.7 and 0.7, such that the resulting Σ_b was a valid structure and in order to represent different spatial correlations between the variables.

Figure 13 presents the MatSimpler model estimation times for all number of

variables and sample locations and Figure 19 in Appendix B illustrates the marginal and cross-covariance functions for the bivariate simulation, with $\rho_{12} = 0.75$. The results shows that the estimation computational time increases with the number of variables and sample locations, which is due to the Cholesky decompositions of the marginal-covariance matrices. These calculations are necessary to obtain the proposed specification.

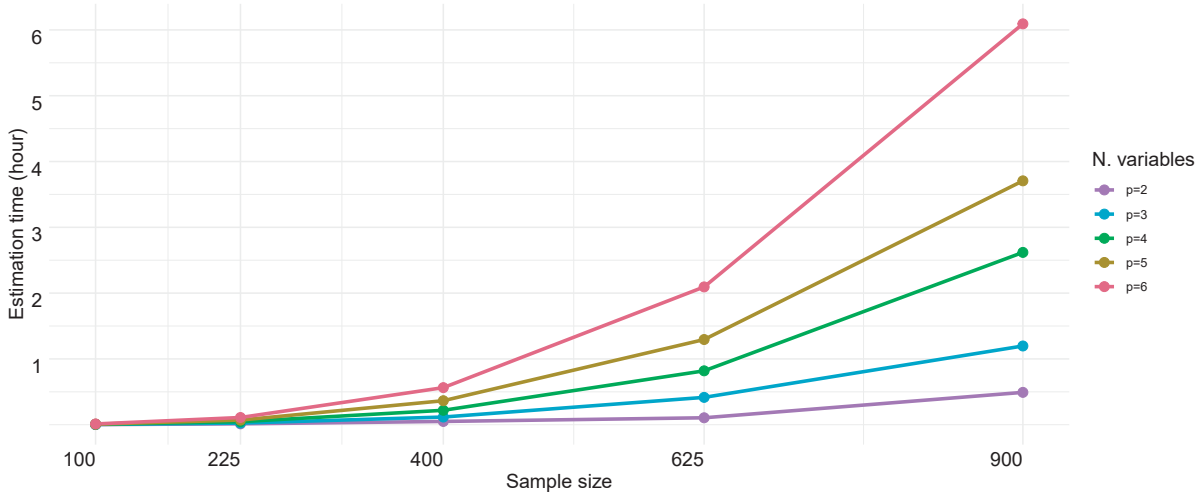

FIGURE 13 – Estimation times for MatSimpler model considering different sample sizes and number of variables

We also compare the estimation computational times for the bivariate case of the MatSimpler model with three other literature models, the *bivariate Matérn model with constraints* (MatConstr), the *bivariate separable Matérn model* (MatSep) and the LMC model. The data were simulated from the MatConstr model. We set $\phi_i = 0.2$, $\nu_i = 0.5$, $\sigma_i = 0.3$, for $i = 1, 2$, and $\rho_{12} = 0.8$. The MatConstr, MatSep and LMC models were estimated by `GeoModels` package, in which we consider the standard likelihood function. For all models we used the Nelder-Mead optimizer and convergence was successful in all scenarios. Fig. 14 illustrates the comparisons between models.

Based on the simulations carried out, we see that the MatSimpler model, with 7 parameters, presents estimation times much lower than the MatConstr model which have the same number of parameters. Furthermore, our model presents estimation times very close to the model with separable structure, MatSep (with 5 parameters), especially for sample locations less than 400, and to the LMC (with 6 parameters). The R codes for Figures 13 and 14 are available in the supplementary materials.

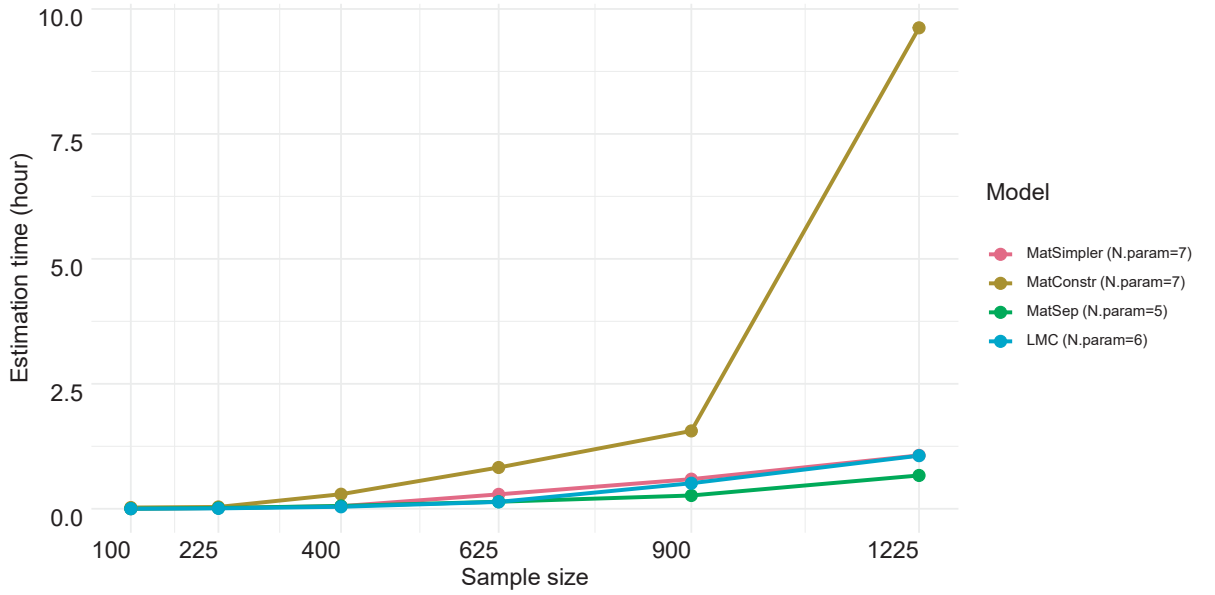

FIGURE 14 – Estimation times for the MatSimpler, MatConstr, MatSep and LMC models, considering simulated data from the MatConstr model for different sample sizes

5 RESULTS AND DISCUSSION

In this chapter, we illustrate the application of our approach in analyzing literature datasets. For the first one, the soill dataset, aim is to model the hydrogen content and the catium exchange capability (CTC). We compare results obtained fitting the MatSimpler and other models. The second one, a dataset on heavy metal concentration, is used to illustrate the model's flexibility by fitting a four-variate dataset.

5.1 EXAMPLE 1: SOIL DATA

The *soil250* dataset from `geoR` package (RIBEIRO JR et al., 2020) contains some soil chemistry properties measured on a regular grid with 10x25 points spaced by 5 meters. We study the relation between the hydrogen content and the catium exchange capability (CTC). These data illustrate a scenario with strong correlation between the variables under study. Fig. 15 shows circle plots of the hydrogen (left panel) and CTC (right panel) data separately. The coordinates are divided by a constant to easy the visualizations. In Appendix C, Fig. 20 shows the histograms of each variable and Fig. 21 shows the scatterplots of each variable against each spatial coordinate, used for check spatial trends.

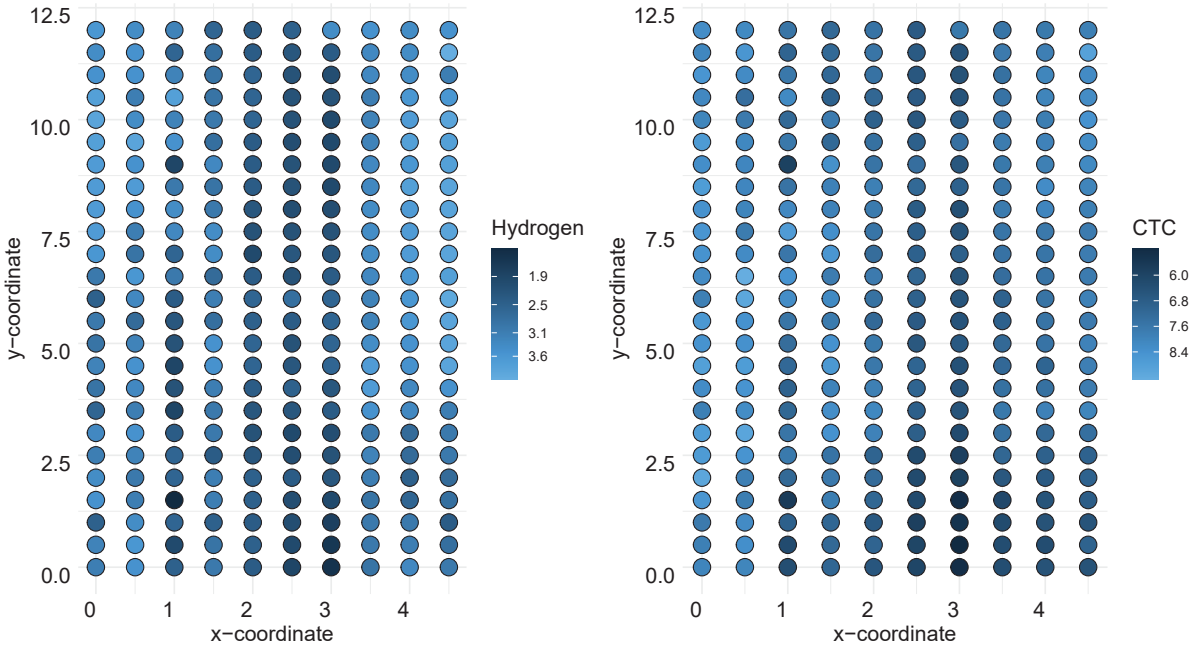

FIGURE 15 – Circle plot of Hydrogen content (left panel) and CTC (right panel) for soil250 data

We proceed with the residuals of a linear regression with constant trend as a realization of a zero-mean bivariate isotropic stationary Gaussian random field. The

sample correlation between the variables was 0.68. Standard deviations are 0.60 for hydrogen content and 0.77 for the CTC.

The models considered in the estimation step are the MatSimpler, MatConstr, MatInd, LMC and MatSep. For MatSimpler and MatConstr the vector of parameters to be estimated is $\boldsymbol{\lambda} = (\phi_1, \phi_2, \nu_1, \nu_2, \sigma_1, \sigma_2, \rho_{12})$, for the separable model, MatSep, we have $\phi_1 = \phi_2$ and $\nu_1 = \nu_2$ and for the independent model we have $\rho_{12} = 0$. The LMC model is parameterized by $\boldsymbol{\lambda} = (a_{11}, a_{12}, a_{22}, a_{21}, \phi_1, \phi_2)$, where a_{11}, a_{12}, a_{22} and a_{21} are elements of the corregionalization matrix and ϕ_1 and ϕ_2 are parameters of the exponential correlation model.

Table 6 presents the parameter estimates, the maximized log-likelihood (LL), the Akaike information criterion (AIC), the Bayesian information criterion (BIC) values and the elapsed estimation times, highlighting the lowest time for the MatSimpler model, especially when compared to the MatConstr model which has the same number of parameters. The MatConstr, MatInd, LMC and MatSep models were estimated using the `GeoFit` function from the `GeoModels` package (BEVILACQUA; MORALES-OÑATE, 2018) for which the standard likelihood function and the Euclidean distance were considered. The Nelder-Mead optimizer is the algorithm of choice for all model fits. The MatSimpler model exhibits the highest log-likelihood value when compared to other models. We remark that, given the strong correlation between the hydrogen and CTC content, the independent model is not expected to perform as well, as it fixes the correlation parameter at $\rho_{12} = 0$. The R codes are available in the supplementary materials.

Considering that the model with separable structure can be seen as a particular case of the MatSimpler model, we performed a likelihood ratio test to verify if the more parsimonious model would be more suitable in this case. Thus, the test statistic is

$$\lambda_0 = -2[LL_{\text{MatSep}} - LL_{\text{MatSimpler}}] = -2[-166.298 + 161.551] = 9.494.$$

With a .05 significance level and $7 - 5 = 2$ degrees of freedom, we have a p-value of 0.008 and so we reject the null hypothesis that the separable model would be more suitable in this case. We can say that the more flexible model, MatSimpler, fits to the proposal of modeling the spatial dependence between the variables.

The empirical and fitted marginal and cross-covariance functions based on the estimates of each model are presented in Figure 22 (Appendix C), where we can see that for hydrogen, the MatSimpler model presented an estimate of the scale parameter slightly lower than the other models. The CTC content presented a bigger scale estimate than hydrogen, for all models. The smoothness estimates for both variables was slightly above 0.5, indicating slightly smoother behaviors than the exponential correlation model, with the exception of the separable and independent models that showed less smooth

TABLE 6 – Parameter estimates of each model for soil250 data

Estimates	Models				
	MatSimpler	MatConstr	MatInd	LMC	MatSep
a_{11}	-	-	-	0.615	-
a_{12}	-	-	-	-0.067	-
a_{22}	-	-	-	0.544	-
a_{21}	-	-	-	0.608	-
$\hat{\phi}_1$	1.077	1.415	1.849	1.156	1.978
$\hat{\phi}_2$	1.994	1.483	2.340	2.816	
ν_1	0.543	0.509	0.398	-	0.495
ν_2	0.513	0.562	0.422	-	
$\hat{\sigma_1}$	0.627	0.671	0.648	-	0.774
$\hat{\sigma_2}$	0.920	0.850	0.848	-	0.891
$\hat{\rho}_{12}$	0.823	0.811	-	-	0.816
LL	-161.551	-166.143	-301.813	-164.688	-166.298
AIC	337.103	346.285	615.626	341.376	342.595
BIC	361.753	375.787	640.914	366.663	363.668
Elapsed time	59.301	406.421	169.630	97.002	113.752
N. Param	7	7	6	6	5

behaviors. The models presented close estimates for the common parameters. Figure 23 (Appendix C) presented the image plot for MatSimpler model fits.

Finally, we conducted a small cross-validation study, evaluating the predictive behavior by a random training selection of 200 locations (80% of the data), from which we estimate the models under study and compute the mean absolute error (MAE), the root mean square error (RMSE) and the normalized mean square error (NMSE) for each model using co-kriging predictor for the 50 remaining locations (20% of the data). These measures are defined as,

$$\text{MAE}_i = \frac{1}{50}\sum_{k=1}^{50}|Y_i(\mathbf{s}_k) - \hat{Y}_i(\mathbf{s}_k)|,$$

$$\text{RMSE}_i = \sqrt{\frac{1}{50}\sum_{k=1}^{50}(Y_i(\mathbf{s}_k) - \hat{Y}_i(\mathbf{s}_k))^2},$$

$$\text{NMSE}_i = \frac{\text{RMSE}_i}{\max(\hat{Y}_i(\mathbf{s}_k)) - \min(\hat{Y}_i(\mathbf{s}_k))},$$

where $\hat{Y}_i(\mathbf{s}_k)$ is the cokriging predictor of the variable $Y_i(\mathbf{s}_k)$, with $i = \text{H}, \text{CTC}$, representing hydrogen and CTC content, respectively. We repeated the same process 150 times,

calculating the log-likelihood, MAE_i, RMSE_i and NMSE_i values, for each variable each time.

We also consider the energy score (ES) (GNEITING; STANBERRY, et al., 2008), which is a multivariate generalization of the continuous ranked probability score (CRPS), and the variogram score of order 0.5 (VS) (SCHEUERER; HAMILL, 2015). The `scoringRules` (JORDAN et al., 2019) package provides tools for comparative evaluation of probabilistic models, including the CRPS, ES and VS, among others.

Table 7 presents the summary of this measures. In general, all models presented equivalent results in terms of predictive capacity. Our proposal presented better results considering the VS, especially when compared to models with the Matérn correlation function. The VS is seen as a predictive diagnostic measure that is sensitive to various types of miscalibration of multivariate forecasts (SCHEUERER; HAMILL, 2015).

TABLE 7 – Mean for log-likelihood and forecasts measures considering 150 splits of data into training (80%) and test (20%) for each model and each variable for soil250 data

Models	LL	MAE_H	RMSE_H	NMSE_H	MAE_CTC	RMSE_CTC	NMSE_CTC	ES	VS
MatSimpler	-144.267	0.278	0.369	0.166	0.315	0.417	0.187	4.961	1196.695
MatConstr	-148.533	0.271	0.364	0.163	0.308	0.411	0.184	4.969	1202.291
MatInd	-255.602	0.272	0.364	0.163	0.309	0.412	0.185	4.939	1203.011
LMC	-147.542	0.272	0.363	0.163	0.307	0.410	0.184	4.924	1185.765
MatSep	-148.916	0.270	0.364	0.163	0.307	0.410	0.184	4.986	1205.838

5.2 EXAMPLE 2: HEAVY METAL CONCENTRATION DATA

To illustrate the application of the MatSimpler model to a 4-variate set, we consider the data *meuse* dataset from the `sp` package, which contains topsoil heavy metal concentrations, collected in a plain flood of the river Meuse an area of approximately 15m x 15m. We studied the relationship between the variables cadmium, copper, lead, zinc, which have a strong sample correlation with each other. Fig. 16 shows the concentrations of each metal along the region under study.

One of the important predictors of heavy metal concentrations is the distance to the Meuse river. Thus, we consider such trend we take residuals from a linear fit of the metal concentrations on such distance. Considering $\hat{\rho}_{ij}$ the sample correlation between the variables i and j, for $i, j = 1, \ldots, 4$, the correlations between them are: $\hat{\rho}_{12} = 0.6510$, $\hat{\rho}_{13} = 0.6165$, $\hat{\rho}_{14} = 0.6585$, $\hat{\rho}_{23} = 0.6970$, $\hat{\rho}_{24} = 0.7466$, $\hat{\rho}_{34} = 0.9392$, with standard deviations: $\hat{s}_1 = 0.8627$, $\hat{s}_2 = 0.3343$, $\hat{s}_3 = 0.4625$ and $\hat{s}_4 = 0.4339$. The histograms of each quantity are shown in Fig. 24 in Appendix C.

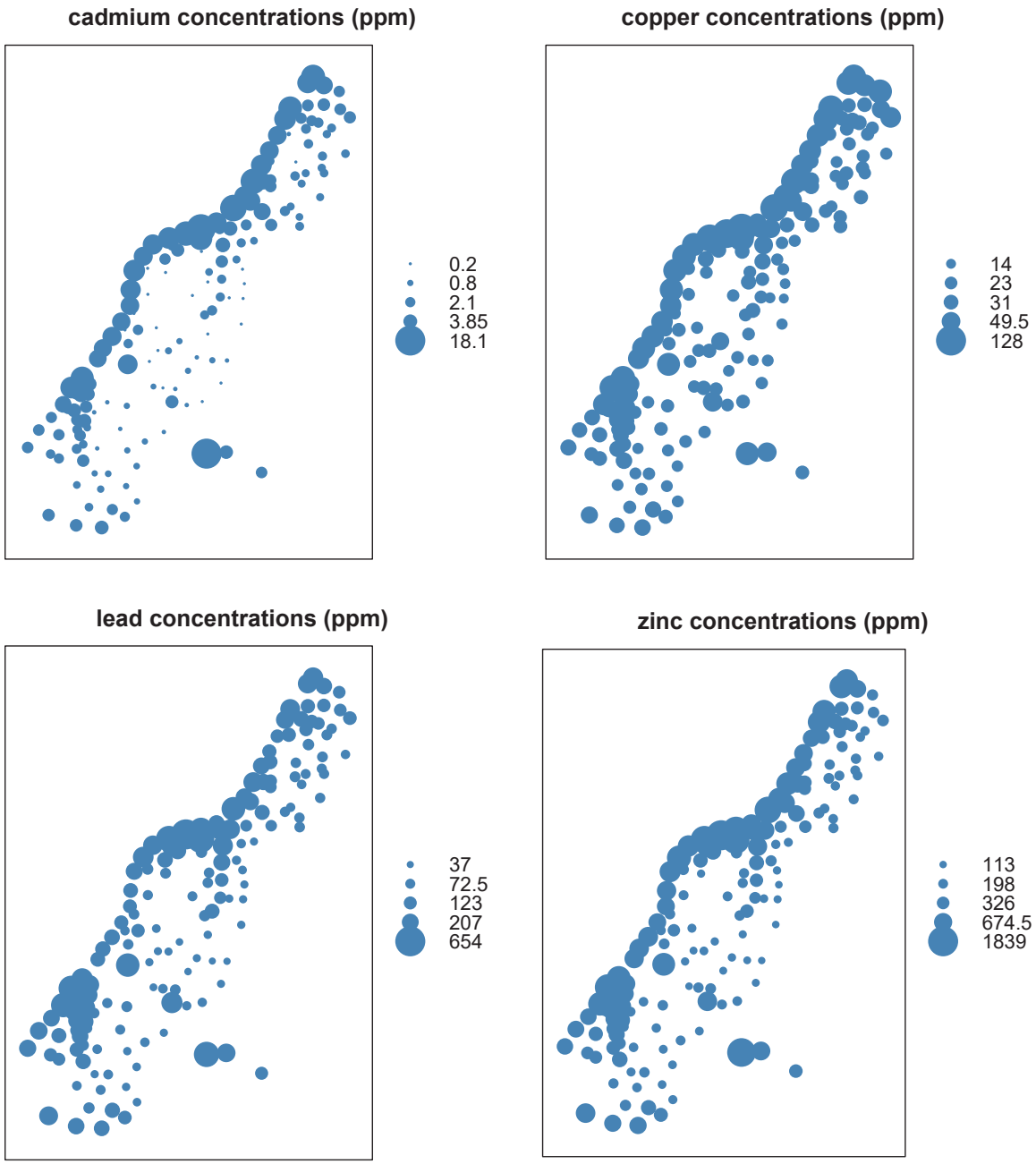

FIGURE 16 – Circle plot of cadmium, copper, lead and zinc concentration, for Meuse data

The parametric estimates as well as the standard errors for the MatSimpler model are presented in Table 8. All variables showed less smooth behavior than the exponential correlation model ($v_i < 0.5$, for $i, \ldots, 4$). The estimates were very close to the sample ones. We present in Figure 25 (Appendix C) the marginal covariance functions, to illustrate the behavior of each variable. We noticed that the lead and zinc variables present very similar marginal behaviors. The variables cadmium and cooper present the largest and smallest variability, respectively.

The MatSimpler model was able to capture the individual variability of each

variable as well as the estimated correlations between them.

TABLE 8 – Parameter estimates for Meuse data

	Estimates					
LL = -89.28	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\phi}_4$	$\hat{\nu}_1$	$\hat{\nu}_2$
Value	0.4260	0.4449	0.4098	0.3401	0.1021	0.1952
sd	0.2925	0.2258	0.1983	0.1557	0.0506	0.0591
	$\hat{\nu}_3$	$\hat{\nu}_4$	$\hat{s}_1$	$\hat{s}_2$	$\hat{s}_3$	$\hat{s}_4$
Value	0.2063	0.2409	0.8897	0.3610	0.4786	0.4596
sd	0.0555	0.0656	0.0595	0.0288	0.0378	0.0348
	$\hat{\rho}_{12}$	$\hat{\rho}_{13}$	$\hat{\rho}_{14}$	$\hat{\rho}_{23}$	$\hat{\rho}_{24}$	$\hat{\rho}_{34}$
Value	0.6701	0.6242	0.7260	0.6728	0.7530	0.9376
sd	0.0447	0.0495	0.0384	0.0444	0.0351	0.0098

SUPPLEMENTARY MATERIALS

Supplementary materials for this work are available at: `https://angelicamariatortola.github.io/academic/thesis.html`.

6 CONCLUSION

We presented a covariance specification for multivariate random fields which is built by the product of matrices for continuously indexed data. The model is simple whilst flexible, allowing for different marginal-covariance functions and correlation structures. Its simple construction and the possibility that its parameters vary in their usual domains, facilitates the specification, estimation, computation and generalization for more than two variables.

The analyzed data sets illustrate the ability to deal with different covariance structures, allowing for different variability, smoothness and scales parameters. It also highlights the reduced computational times in comparison with other models. A simulation study with different combinations of model parameters suggests maximum likelihood estimators are asymptotically unbiased and consistent.

The major computational burden is the Cholesky decomposition, depending on the number of sample locations and response variables considered in the analysis. We illustrate the growth of computational time for up to six response variables and 900 sample locations. In this scenario, estimation time was approximately six hours (Fig. 13). Precise times will be hardware dependent. However, comparing it with some literature models for the bivariate case (Fig. 14), our proposal was competitive, presenting smaller estimation times, specially when compared to the MatConstr model, which has the same number of parameters.

Table 9 presents a general comparison between the models, highlighting some main characteristics: if the model in question has a more restricted parametric space, if it allows the estimation of a correlation parameter between the variables under study, if it allows different parameters of smoothness and scale (in the case of the Matérn model), it allows different families of correlation, it allows that different variables can be represented by different families of correlation and the proportion of computational time of estimation. We observed the numerous advantages of the Simpler model, due to its flexibility and computational efficiency.

Future directions include the study of estimation methods for the proposed approach that are more efficient, so that the estimation times will not be so penalized by the increase in variables and sampling locations. The organization and construction of an R package that involves the proposed approach shall be considered. Furthermore, the presented proposal opens options for future research in the context of non-Gaussian modeling and asymmetric data for multivariate spatial problems.

TABLE 9 – Comparisons between models, considering different characteristics

Charac. / Models	Complexity[1]	ρ_{ij}[2]	Different correlations[3]	Different models[4]	Cross-cov. from marginal-cov.[5]	One model by variable[6]	Time[7]
Separable model		X		X	X		6,92%
LMC			X	X			11,03%
Simpler model		X	X	X	X	X	11,09%
Matérn with Constr.	X	X	X				100%

[1] Constraints in parametric space
[2] Able to estimate correlation parameter between variables
[3] Allows different correlation structures for each variable (considering the Matérn model)
[4] Supports different correlation families: Matérn, Spherical, Cauchy, ...
[5] Cross-covariance function specified from the marginal-covariance functions
[6] Allows different variables to be represented by different correlation families
[7] Proportion of estimation times with respect to maximum time (Figure 20 - Ap. C) for n=1225.

REFERENCES

ALEGRÍA, A.; PORCU, E.; FURRER, R. Asymmetric matrix-valued covariances for multivariate random fields on spheres. **Journal of Statistical Computation and Simulation**, v. 88, n. 10, p. 1850–1862, 2018. Cit. on pp. 11, 12.

ALEGRÍA, A.; PORCU, E.; FURRER, R.; MATEU, J. Covariance functions for multivariate Gaussian fields evolving temporally over planet earth. **Stochastic Environmental Research and Risk Assessment**, v. 33, n. 8, p. 1593–1608, 2019. Cit. on p. 11.

APANASOVICH, T. V.; GENTON, M. G.; SUN, Y. A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. **Journal of the American Statistical Association**, v. 107, n. 497, p. 180–193, 2012. Cit. on p. 11.

ARIA, M.; CUCCURULLO, C. bibliometrix: An R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, v. 11, n. 4, p. 959–975, 2017. Cit. on p. 14.

BANERJEE, S.; CARLIN, B. P.; GELFAND, A. E. **Hierarchical modeling and analysis for spatial data**. Boca Raton: Chapman and Hall/CRC, 2003. Cit. on p. 24.

BANERJEE, S.; ROY, A. **Linear algebra and matrix analysis for statistics**. Boca Raton: CRC Press, 2014. Cit. on pp. 27, 28.

BATES, D.; MAECHLER, M. **Matrix: Sparse and Dense Matrix Classes and Methods**. R Foundation for Statistical Computing. Vienna, Austria, 2021. R package version 1.3-4. Available from: <https://CRAN.R-project.org/package=Matrix>. Cit. on p. 33.

BEVILACQUA, M.; ALEGRÍA, A., et al. Composite likelihood inference for multivariate Gaussian random fields. **Journal of agricultural, biological, and environmental statistics**, v. 21, n. 3, p. 448–469, 2016. Cit. on pp. 11, 12, 25.

BEVILACQUA, M.; DIGGLE, P.; PORCU, E. Families of covariance functions for bivariate random fields on spheres. **Spatial Statistics**, v. 40, n. 100448, p. 1–29, 2020. Cit. on pp. 11, 12.

BEVILACQUA, M.; FASSÒ, A., et al. Covariance tapering for multivariate Gaussian random fields estimation. **Statistical Methods & Applications**, v. 25, n. 1, p. 21–37, 2016. Cit. on p. 12.

BEVILACQUA, M.; MORALES-OÑATE, V. **GeoModels: A Package for Geostatistical Gaussian and non Gaussian Data Analysis**. R Foundation for Statistical Computing. Vienna, Austria, 2018. R package version 1.0.3-4. Available from: <https://vmoprojs.github.io/GeoModels-page/>. Cit. on pp. 25, 42.

BEVILACQUA, M.; VALLEJOS, R.; VELANDIA, D. Assessing the significance of the correlation between the components of a bivariate Gaussian random field. **Environmetrics**, v. 26, n. 8, p. 545–556, 2015. Cit. on pp. 12, 25.

BIVAND, R. S.; PEBESMA, E. J.; GÓMEZ-RUBIO, V. **Applied spatial data analysis with R**. New York: Springer, 2008. Cit. on p. 32.

BONAT, W. H.; PETTERLE, R. R., et al. Modelling multiple outcomes in repeated measures studies: Comparing aesthetic eyelid surgery techniques. **Statistical Modelling**, v. 21, n. 6, p. 564–582, 2021. Cit. on p. 31.

BONAT, W. H.; JØRGENSEN, B. Multivariate covariance generalized linear models. **Journal of the Royal Statistical Society. Series C (Applied Statistics)**, v. 65, n. 5, p. 649–675, 2016. Cit. on pp. 32, 35.

BOURGAULT, G.; MARCOTTE, D. Multivariable variogram and its application to the linear model of coregionalization. **Mathematical Geology**, v. 23, n. 7, p. 899–928, 1991. Cit. on pp. 11, 23.

CECCONI, L. et al. Preferential sampling and Bayesian geostatistics: Statistical modeling and examples. **Statistical methods in medical research**, v. 25, n. 4, p. 1224–1243, 2016. Cit. on p. 24.

CHILÈS, J.-P.; DELFINER, P. **Geostatistics: Modeling Spatial Uncertainty**. New York: Wiley, 2012. Cit. on pp. 18, 19, 21–24.

CRESSIE, N. **Statistics for spatial data**. New York: Wiley, 1993. Cit. on pp. 12, 24, 31.

DIGGLE, P.; RIBEIRO JR, P. J. **Model-based Geostatistics**. New York: Springer, 2007. Cit. on pp. 19, 21, 22.

EMERY, X.; PORCU, E. Simulating isotropic vector-valued Gaussian random fields on the sphere through finite harmonics approximations. **Stochastic Environmental Research and Risk Assessment**, v. 33, n. 8, p. 1659–1667, 2019. Cit. on p. 11.

EMERY, X.; PORCU, E.; BISSIRI, P. A semiparametric class of axially symmetric random fields on the sphere. **Stochastic Environmental Research and Risk Assessment**, v. 33, n. 10, p. 1863–1874, 2019. Cit. on p. 11.

FASIOLO, M. **An introduction to mvnfast**. R Foundation for Statistical Computing. Vienna, Austria, 2016. R package version 0.1.6. Available from: <https://CRAN.R-project.org/package=mvnfast>. Cit. on p. 33.

FINLEY, A. O.; BANERJEE, S.; GELFAND, A. E. spBayes for Large Univariate and Multivariate Point-Referenced Spatio-Temporal Data Models. **Journal of Statistical Software, Articles**, v. 63, n. 13, p. 1–28, 2015. Cit. on p. 24.

GELFAND, A. E. et al. Nonstationary multivariate process modeling through spatially varying coregionalization. **Test**, v. 13, n. 2, p. 263–312, 2004. Cit. on p. 24.

GENTLE, J. E. **Matrix algebra: Theory, Computations, and Applications in Statistics**. New York: Springer, 2017. Cit. on p. 28.

GENTON, M. G.; KLEIBER, W. Cross-covariance functions for multivariate geostatistics. **Statistical Science**, v. 30, n. 2, p. 147–163, 2015. Cit. on pp. 11, 23, 24.

GNEITING, T. Correlation functions for atmospheric data analysis. **Quarterly Journal of the Royal Meteorological Society**, v. 125, n. 559, p. 2449–2464, 1999. Cit. on p. 22.

GNEITING, T.; KLEIBER, W.; SCHLATHER, M. Matérn cross-covariance functions for multivariate random fields. **Journal of the American Statistical Association**, v. 105, n. 491, p. 1167–1177, 2010. Cit. on pp. 11, 12, 24.

GNEITING, T.; STANBERRY, L. I., et al. Assessing probabilistic forecasts of multivariate quantities, with an application to ensemble predictions of surface winds. **Test**, v. 17, n. 2, p. 211–235, 2008. Cit. on p. 44.

GOOVAERTS, P. et al. **Geostatistics for natural resources evaluation**. New York: Oxford University Press, 1997. Cit. on pp. 12, 24.

GOULARD, M.; VOLTZ, M. Linear coregionalization model: tools for estimation and choice of cross-variogram matrix. **Mathematical Geology**, v. 24, n. 3, p. 269–286, 1992. Cit. on pp. 11, 23.

HARDY, Y.; STEEB, W.-H. **Matrix Calculus, Kronecker Product and Tensor Product: A Practical Approach to Linear Algebra, Multilinear Algebra and Tensor Calculus with Software Implementations**. Singapore: World Scientific, 2019. Cit. on p. 28.

IP, R. H.; LI, W. A class of valid Matérn cross-covariance functions for multivariate spatio-temporal random fields. **Statistics & Probability Letters**, v. 130, p. 115–119, 2017. Cit. on p. 12.

JORDAN, A.; KRÜGER, F.; LERCH, S. Evaluating Probabilistic Forecasts with scoringRules. **Journal of Statistical Software**, v. 90, n. 12, p. 1–37, 2019. Cit. on p. 44.

KLEIBER, W. Coherence for multivariate random fields. **Statistica Sinica**, p. 1675–1697, 2017. Cit. on p. 12.

MACNAB, Y. C. Linear models of coregionalization for multivariate lattice data: Order-dependent and order-free cMCARs. **Statistical methods in medical research**, v. 25, n. 4, p. 1118–1144, 2016. Cit. on p. 11.

MACNAB, Y. C. Some recent work on multivariate Gaussian Markov random fields. **Test**, v. 27, n. 3, p. 497–541, 2018. Cit. on p. 11.

MARTINEZ-BENEITO, M. A. A general modelling framework for multivariate disease mapping. **Biometrika**, v. 100, n. 3, p. 539–553, 2013. Cit. on p. 26.

MARTINEZ-BENEITO, M. A. Some links between conditional and coregionalized multivariate Gaussian Markov random fields. **Spatial Statistics**, v. 40, n. 100383, p. 1–17, 2020. Cit. on p. 11.

MARTINS, A. B. T.; BONAT, W. H.; RIBEIRO JR, P. J. Likelihood analysis for a class of spatial geostatistical compositional models. **Spatial Statistics**, v. 17, p. 121–130, 2016. Cit. on p. 33.

MATÉRN, B. **Spatial variation**. Berlin: Springer, 1986. Cit. on p. 21.

PEBESMA, E. J. Multivariable geostatistics in S: the gstat package. **Computers & Geosciences**, v. 30, p. 683–691, 2004. Cit. on p. 32.

PETERSEN, K. B.; PEDERSEN, M. S., et al. The matrix cookbook. **Technical University of Denmark**, 2008. Cit. on p. 28.

PETTERLE, R. R.; BONAT, W. H.; SCARPIN, C. T. Quasi-beta longitudinal regression model applied to water quality index data. **Journal of Agricultural, Biological and Environmental Statistics**, v. 24, n. 2, p. 346–368, 2019. Cit. on p. 35.

PORCU, E. et al. Radial basis functions with compact support for multivariate geostatistics. **Stochastic environmental research and risk assessment**, v. 27, n. 4, p. 909–922, 2013. Cit. on p. 12.

PRAUSE, A.; STELAND, A., et al. Estimation of the asymptotic variance of univariate and multivariate random fields and statistical inference. **Electronic Journal of Statistics**, v. 12, n. 1, p. 890–940, 2018. Cit. on p. 12.

QADIR, G. A.; EUÁN, C.; SUN, Y. Flexible modeling of variable asymmetries in cross-covariance functions for multivariate random fields. **Journal of Agricultural, Biological and Environmental Statistics**, v. 26, n. 1, p. 1–22, 2021. Cit. on p. 11.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2021. Available from: <https://www.R-project.org/>. Cit. on pp. 12, 14, 33.

RIBEIRO, A. M. T.; RIBEIRO JR, P. J.; BONAT, W. H. Comparison of exponential covariance functions for bivariate geostatistical data. **Revista Brasileira de Biometria**, v. 39, n. 1, p. 89–102, 2021. Cit. on p. 27.

RIBEIRO JR, P. J. et al. **geoR: Analysis of Geostatistical Data**. R Foundation for Statistical Computing. Vienna, Austria, 2020. R package version 1.8-1. Available from: <https://CRAN.R-project.org/package=geoR>. Cit. on pp. 33, 41.

RUE, H.; HELD, L. **Gaussian Markov random fields: theory and applications**. Boca Raton: CRC press, 2005. Cit. on p. 26.

SALVAÑA, M. L. O.; GENTON, M. G. Nonstationary cross-covariance functions for multivariate spatio-temporal random fields. **Spatial Statistics**, v. 37, n. 100411, p. 1–24, 2020. Cit. on p. 11.

SÄRKKÄ, S. **Bayesian filtering and smoothing**. Cambridge: Cambridge University Press, 2013. Cit. on p. 32.

SCHEUERER, M.; HAMILL, T. M. Variogram-based proper scoring rules for probabilistic forecasts of multivariate quantities. **Monthly Weather Review**, v. 143, n. 4, p. 1321–1334, 2015. Cit. on p. 44.

SCHMIDT, A. M.; GELFAND, A. E. A Bayesian coregionalization approach for multivariate pollutant data. **Journal of geophysical research**, v. 108, n. D24, p. 1–9, 2003. Cit. on p. 24.

TEICHMANN, J. et al. Modeling and Fitting of Three-Dimensional Mineral Microstructures by Multinary Random Fields. **Mathematical Geosciences**, v. 53, n. 5, p. 877–904, 2020. Cit. on p. 11.

VALLEJOS, R.; OSORIO, F.; BEVILACQUA, M. **Spatial relationships between two georeferenced variables: With applications in R**. New York: Springer, 2020. Cit. on pp. 11, 12, 25.

VER HOEF, J. M.; CRESSIE, N. Multivariable spatial prediction. **Mathematical Geology**, v. 25, n. 2, p. 219–240, 1993. Cit. on p. 32.

WACKERNAGEL, H. **Multivariate geostatistics: an introduction with applications**. Berlin: Springer, 2003. Cit. on pp. 11, 18–21, 23.

WAND, M. Vector differential calculus in statistics. **The American Statistician**, v. 56, n. 1, p. 55–62, 2002. Cit. on p. 31.

Appendix

APPENDIX A – SECTION 4.1: MODEL SPECIFICATION RESULTS

MATRIX CALCULATIONS FOR OBTAINING THE MATRIX-VALUED COVARIANCE SPECIFICATION $\boldsymbol{\Sigma}(\mathbf{h})$:

Considering $\boldsymbol{\Sigma}_{ii}$ the marginal covariance function of the i-th component, for $i = 1, \ldots, p$, $\tilde{\boldsymbol{\Sigma}}_{ii}$ the lower triangular matrix of the Cholesky decomposition of the matrix $\boldsymbol{\Sigma}_{ii}$, $\boldsymbol{\Sigma}_b$ the correlation function defined by 4.2, and $\mathbf{I}$ the identity matrix of dimension n, each matrix term of 4.1 takes the form:

$$\text{Bdiag}\left(\tilde{\boldsymbol{\Sigma}}_{11}, \tilde{\boldsymbol{\Sigma}}_{22}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}\right) = \begin{pmatrix} \tilde{\boldsymbol{\Sigma}}_{11} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \tilde{\boldsymbol{\Sigma}}_{22} & \ldots & \mathbf{0} \\ \ldots & \ldots & \ldots & \ldots \\ \mathbf{0} & \mathbf{0} & \ldots & \tilde{\boldsymbol{\Sigma}}_{pp} \end{pmatrix}, \tag{A.1}$$

$$\text{Bdiag}\left(\tilde{\boldsymbol{\Sigma}}_{11}^{\top}, \tilde{\boldsymbol{\Sigma}}_{22}^{\top}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}^{\top}\right) = \begin{pmatrix} \tilde{\boldsymbol{\Sigma}}_{11}^{\top} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \tilde{\boldsymbol{\Sigma}}_{22}^{\top} & \ldots & \mathbf{0} \\ \ldots & \ldots & \ldots & \ldots \\ \mathbf{0} & \mathbf{0} & \ldots & \tilde{\boldsymbol{\Sigma}}_{pp}^{\top} \end{pmatrix}, \tag{A.2}$$

$$(\boldsymbol{\Sigma}_b \otimes \mathbf{I}) = \left(\begin{array}{cccc|cccc|c|cccc} 1 & 0 & \ldots & 0 & \rho_{12} & 0 & \ldots & 0 & \ldots & \rho_{1p} & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \rho_{12} & \ldots & 0 & \ldots & 0 & \rho_{1p} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & \rho_{12} & \ldots & 0 & 0 & \ldots & \rho_{1p} \\ \hline \rho_{12} & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 & \ldots & \rho_{2p} & 0 & \ldots & 0 \\ 0 & \rho_{12} & \ldots & 0 & 0 & 1 & \ldots & 0 & \ldots & 0 & \rho_{2p} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \rho_{12} & 0 & 0 & \ldots & 1 & \ldots & 0 & 0 & \ldots & \rho_{2p} \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \hline \rho_{1p} & 0 & \ldots & 0 & \rho_{2p} & 0 & \ldots & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & \rho_{1p} & \ldots & 0 & 0 & \rho_{2p} & \ldots & 0 & \ldots & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \rho_{1p} & 0 & 0 & \ldots & \rho_{2p} & \ldots & 0 & 0 & \ldots & 1 \end{array}\right). \tag{A.3}$$

Making the product between A.1 and A.3, we have:

$$\text{Bdiag}\left(\tilde{\boldsymbol{\Sigma}}_{11}, \tilde{\boldsymbol{\Sigma}}_{22}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}\right)(\boldsymbol{\Sigma}_b \otimes \mathbf{I}) = \begin{pmatrix} \tilde{\boldsymbol{\Sigma}}_{11} & \rho_{12}\tilde{\boldsymbol{\Sigma}}_{11} & \ldots & \rho_{1p}\tilde{\boldsymbol{\Sigma}}_{11} \\ \rho_{12}\tilde{\boldsymbol{\Sigma}}_{22} & \tilde{\boldsymbol{\Sigma}}_{22} & \ldots & \rho_{2p}\tilde{\boldsymbol{\Sigma}}_{22} \\ \ldots & \ldots & \ldots & \ldots \\ \rho_{1p}\tilde{\boldsymbol{\Sigma}}_{pp} & \rho_{2p}\tilde{\boldsymbol{\Sigma}}_{pp} & \ldots & \tilde{\boldsymbol{\Sigma}}_{pp} \end{pmatrix}, \tag{A.4}$$

and from the product between A.4 and A.2 the resulting covariance specification, takes the matrix form:

$$\begin{aligned} \boldsymbol{\Sigma}(\mathbf{h}) &= \text{Bdiag}\left(\tilde{\boldsymbol{\Sigma}}_{11}, \tilde{\boldsymbol{\Sigma}}_{22}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}\right)(\boldsymbol{\Sigma}_b \otimes \mathbf{I})\,\text{Bdiag}\left(\tilde{\boldsymbol{\Sigma}}_{11}^{\top}, \tilde{\boldsymbol{\Sigma}}_{22}^{\top}, \ldots, \tilde{\boldsymbol{\Sigma}}_{pp}^{\top}\right) \\ &= \begin{pmatrix} \tilde{\boldsymbol{\Sigma}}_{11} & \rho_{12}\tilde{\boldsymbol{\Sigma}}_{11} & \ldots & \rho_{1p}\tilde{\boldsymbol{\Sigma}}_{11} \\ \rho_{12}\tilde{\boldsymbol{\Sigma}}_{22} & \tilde{\boldsymbol{\Sigma}}_{22} & \ldots & \rho_{2p}\tilde{\boldsymbol{\Sigma}}_{22} \\ \ldots & \ldots & \ldots & \ldots \\ \rho_{1p}\tilde{\boldsymbol{\Sigma}}_{pp} & \rho_{2p}\tilde{\boldsymbol{\Sigma}}_{pp} & \ldots & \tilde{\boldsymbol{\Sigma}}_{pp} \end{pmatrix} \begin{pmatrix} \tilde{\boldsymbol{\Sigma}}_{11}^{\top} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \tilde{\boldsymbol{\Sigma}}_{22}^{\top} & \ldots & \mathbf{0} \\ \ldots & \ldots & \ldots & \ldots \\ \mathbf{0} & \mathbf{0} & \ldots & \tilde{\boldsymbol{\Sigma}}_{pp}^{\top} \end{pmatrix} \\ &= \begin{pmatrix} \tilde{\boldsymbol{\Sigma}}_{11}\tilde{\boldsymbol{\Sigma}}_{11}^{\top} & \rho_{12}\tilde{\boldsymbol{\Sigma}}_{11}\tilde{\boldsymbol{\Sigma}}_{22}^{\top} & \ldots & \rho_{1p}\tilde{\boldsymbol{\Sigma}}_{11}\tilde{\boldsymbol{\Sigma}}_{pp}^{\top} \\ \rho_{12}\tilde{\boldsymbol{\Sigma}}_{22}\tilde{\boldsymbol{\Sigma}}_{11}^{\top} & \tilde{\boldsymbol{\Sigma}}_{22}\tilde{\boldsymbol{\Sigma}}_{22}^{\top} & \ldots & \rho_{2p}\tilde{\boldsymbol{\Sigma}}_{22}\tilde{\boldsymbol{\Sigma}}_{pp}^{\top} \\ \ldots & \ldots & \ldots & \ldots \\ \rho_{1p}\tilde{\boldsymbol{\Sigma}}_{pp}\tilde{\boldsymbol{\Sigma}}_{11}^{\top} & \rho_{2p}\tilde{\boldsymbol{\Sigma}}_{pp}\tilde{\boldsymbol{\Sigma}}_{22}^{\top} & \ldots & \tilde{\boldsymbol{\Sigma}}_{pp}\tilde{\boldsymbol{\Sigma}}_{pp}^{\top} \end{pmatrix} \end{aligned}$$

that is, the cross-covariance specification between the ij-th process is defined as:

$$\boldsymbol{\Sigma}_{ij}(\mathbf{h}) = \rho_{ij}\tilde{\boldsymbol{\Sigma}}_{ii}\tilde{\boldsymbol{\Sigma}}_{jj}^{\top}, \quad \rho_{ii} = 1. \tag{A.5}$$

When $i = j$, the cross-covariance function in A.5 reduces to the marginal covariance function, $\boldsymbol{\Sigma}_{ii}(\mathbf{h})$.

APPENDIX B – SECTION 4.5: SIMULATION STUDY RESULTS

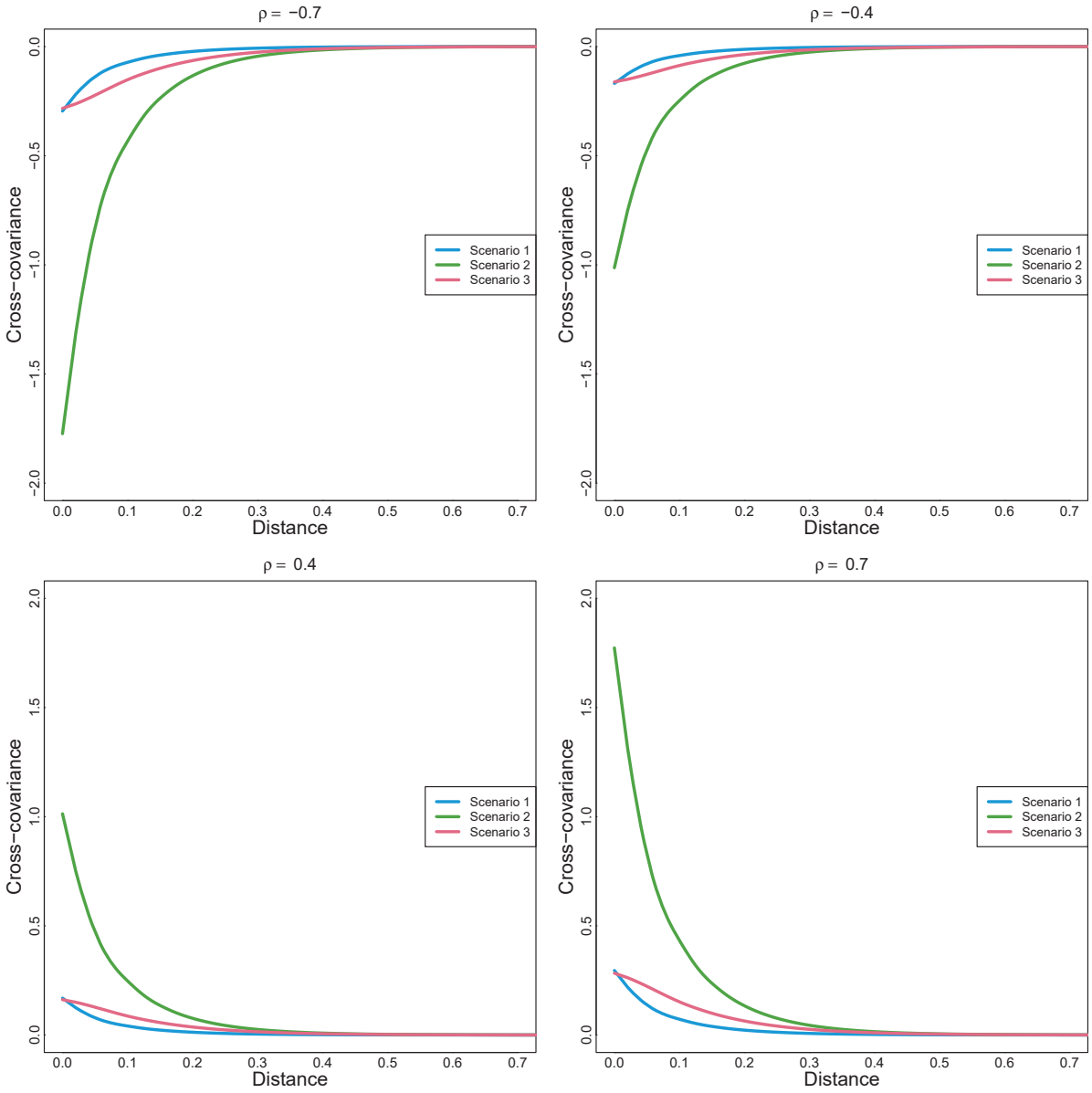

FIGURE 17 – Cross-covariances for considering each simulated scenario and each correlation parameter

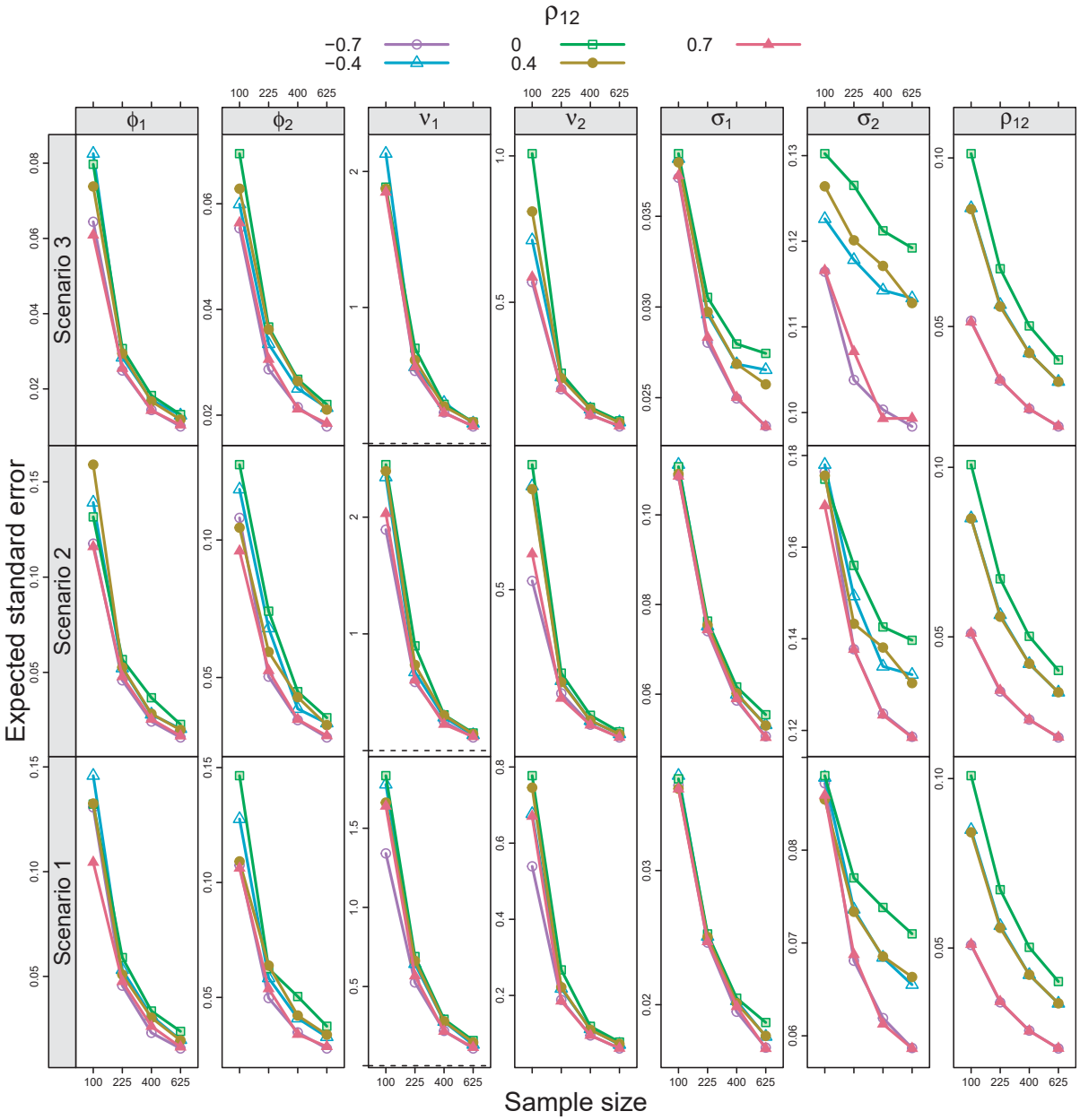

FIGURE 18 – Expected standard error for MatSimpler model parameters considering different simulated parametric scenarios and different sample sizes. Note that figures are on different numerical scales

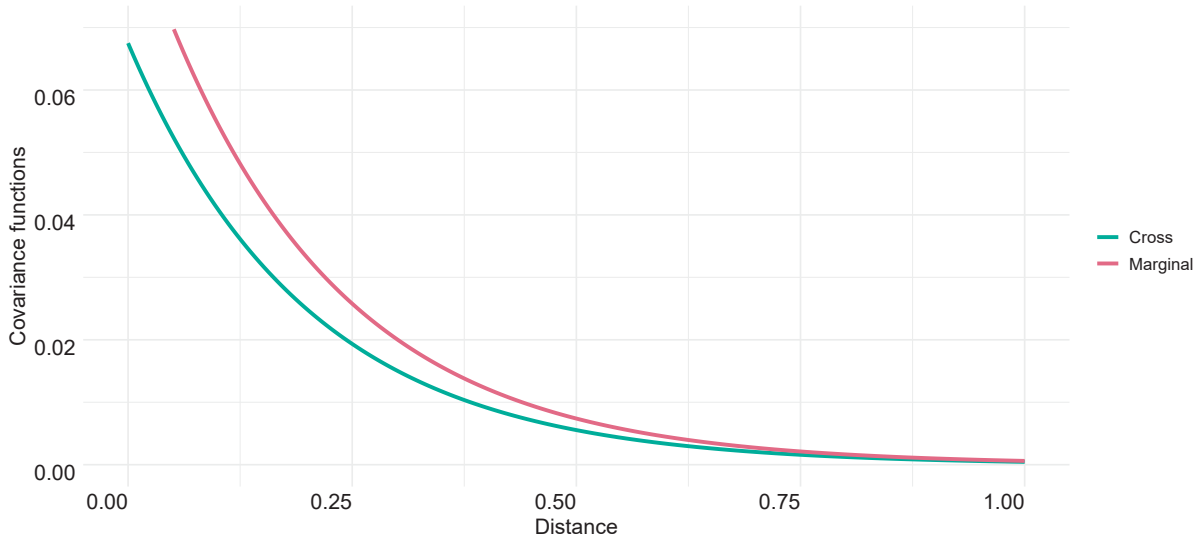

FIGURE 19 – Marginal and cross-covariance funtions for bivariate MatSimpler model simulation with $\phi_1 = \phi_2 = 0.2$, $\nu_1 = \nu_2 = 0.5$, $\sigma_1 = \sigma_2 = 0.3$ and $\rho_{12} = 0.75$

APPENDIX C – SECTION 5.2: DATA ANALYSIS RESULTS

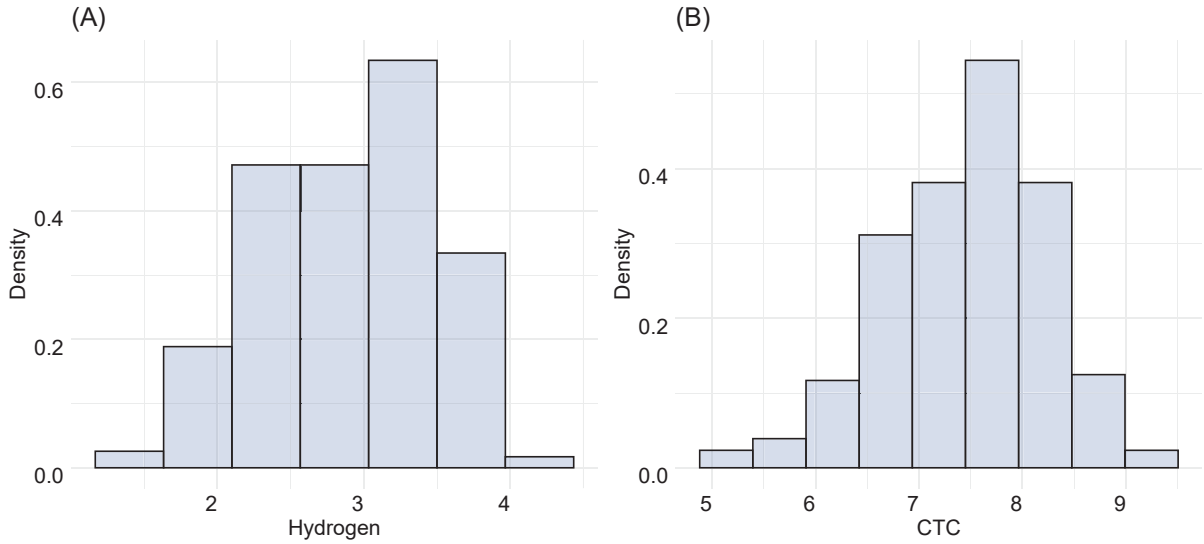

FIGURE 20 – Histogram of (A) Hydrogen content and (B) CTC for soil250 data

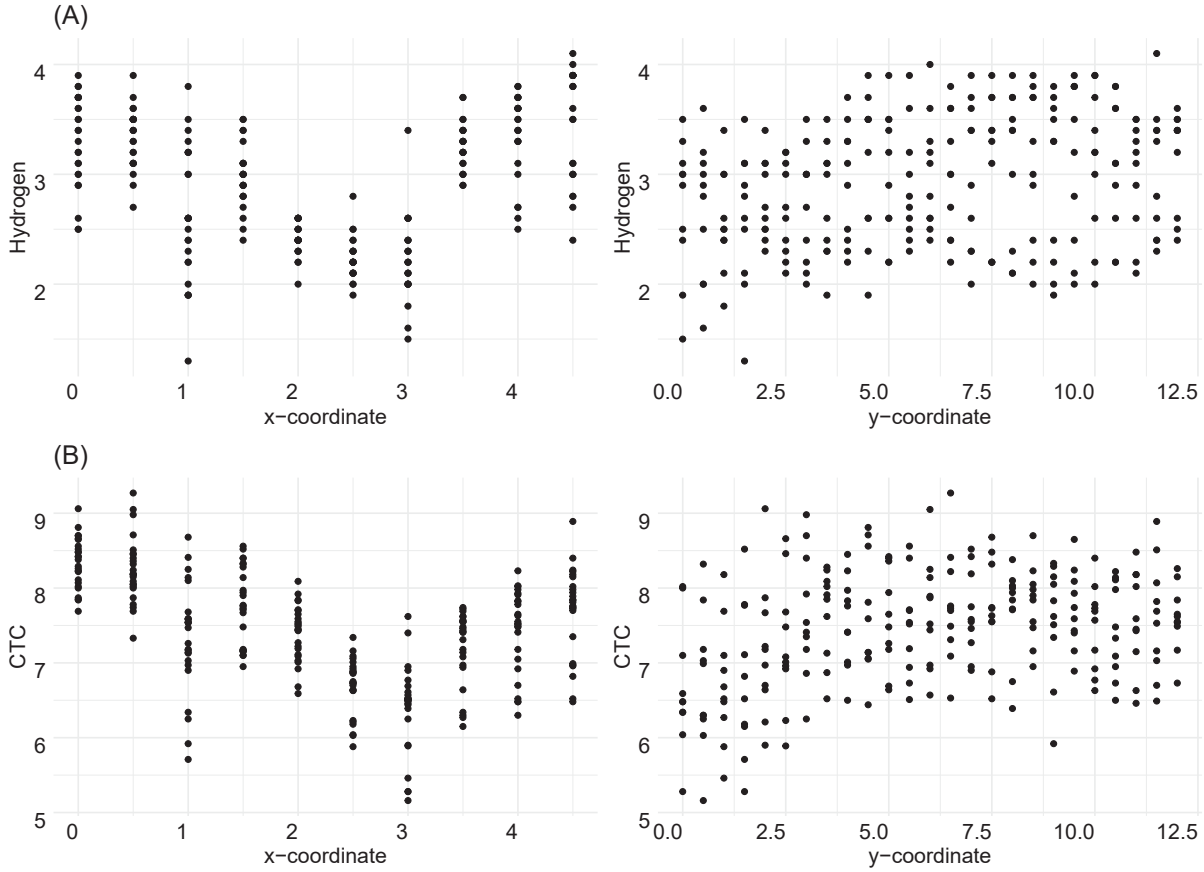

FIGURE 21 – Scatterplots of (A) Hydrogen content and (B) CTC against the coordinates for soil250 data

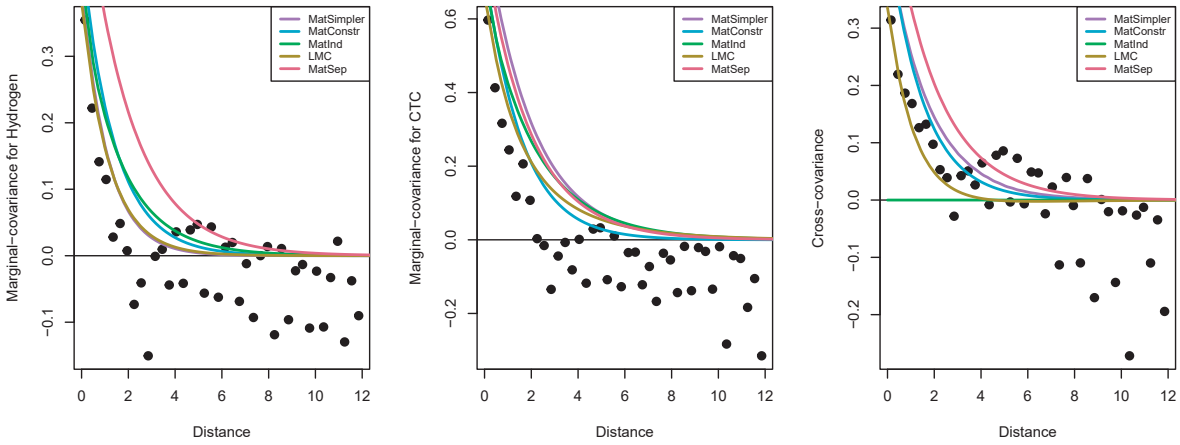

FIGURE 22 – Empirical covariance and cross-covariance functions for Hydrogen and CTC content, with maximum likelihood estimated models for soil data250

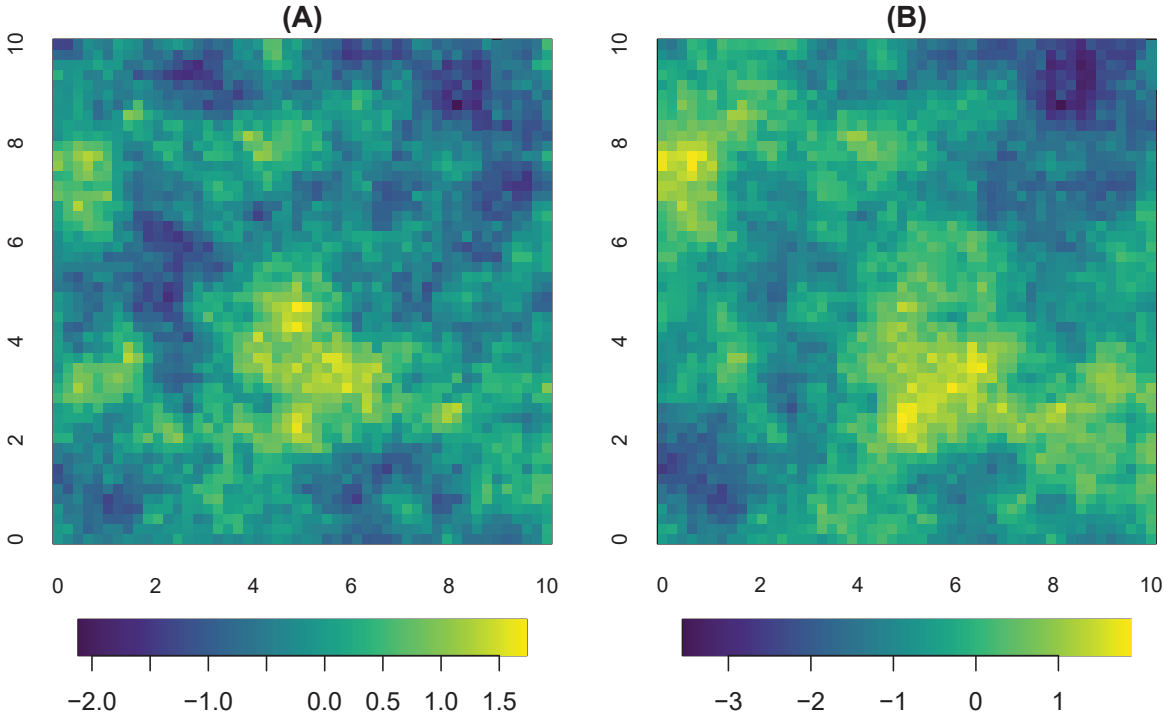

FIGURE 23 – Image plot for estimated covariance functions for (A) Hydrogen and (B) CTC content, considering MatSimpler model likelihood estimates, for soil data250

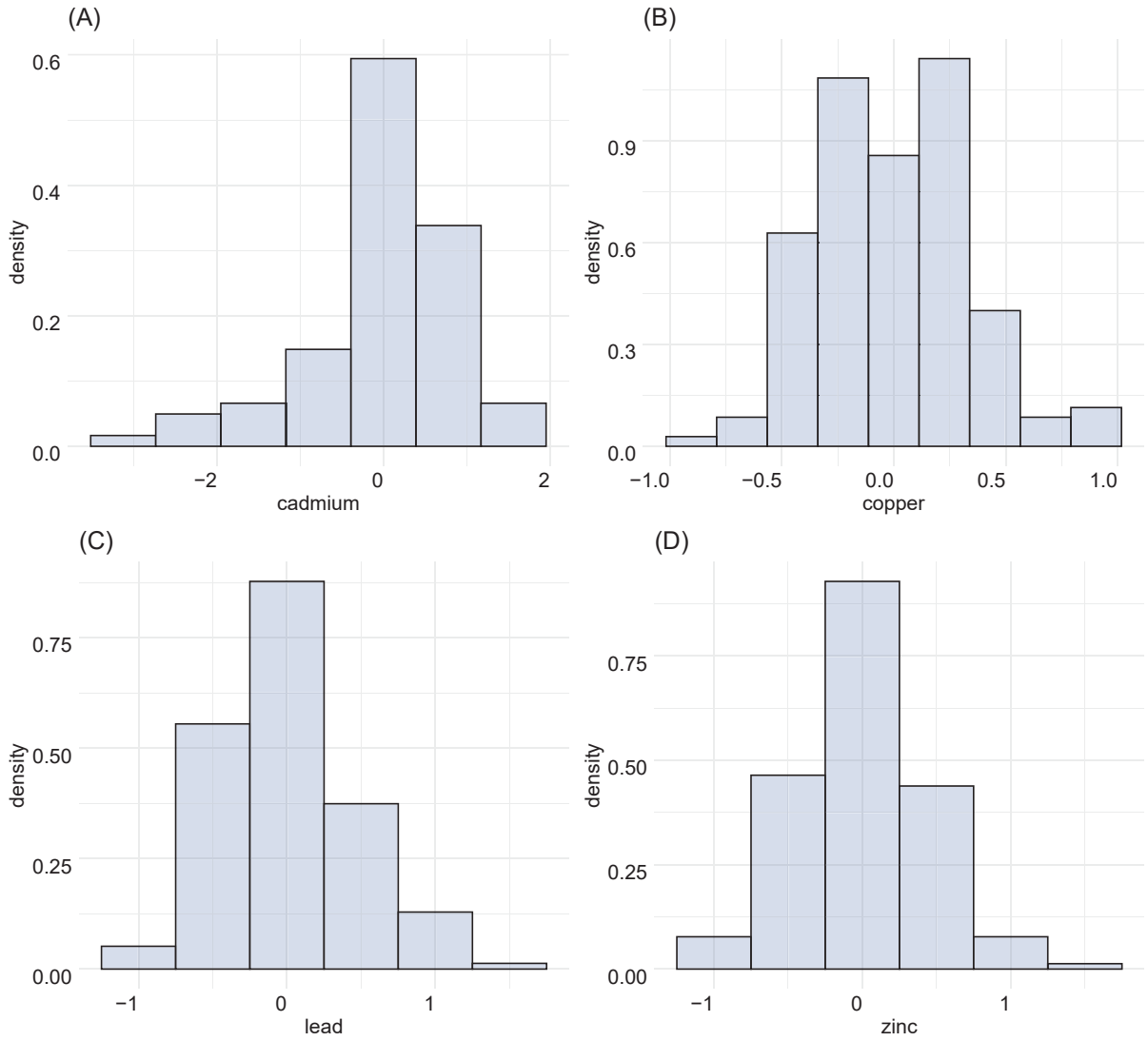

FIGURE 24 – Histogram of the residuals after removing tendencies with respect to distance of (A) cadmium, (B) copper, (C) lead and (D) zinc concentration, for Meuse data

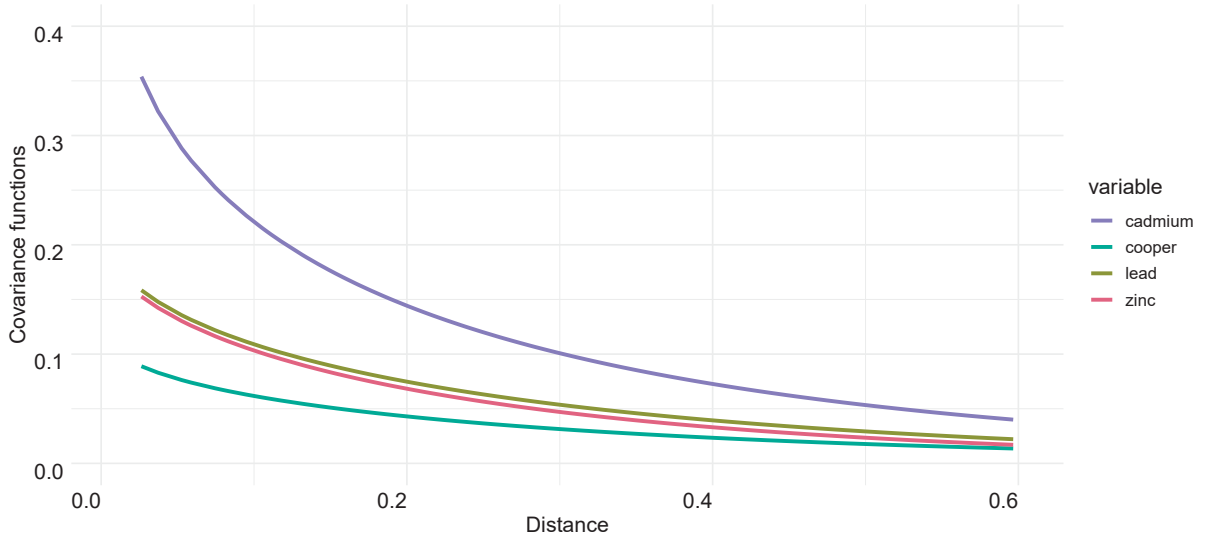

FIGURE 25 – Marginal covariance functions of each variable, for Meuse data