

Article

An Overview of Kriging and Cokriging Predictors for Functional Random Fields

Ramón Giraldo ¹, Víctor Leiva ^{2,*} and Cecilia Castro ³

¹ Departamento de Estadística, Universidad Nacional de Colombia, Sede Bogotá, Bogotá 111321, Colombia; rgiraldoh@unal.edu.co

² School of Industrial Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso 2362807, Chile

³ Centre of Mathematics, Universidade do Minho, 4710-057 Braga, Portugal; cecilia@math.uminho.pt

* Correspondence: victor.leiva@pucv.cl or victorleivasanchez@gmail.com

Abstract: This article presents an overview of methodologies for spatial prediction of functional data, focusing on both stationary and non-stationary conditions. A significant aspect of the functional random fields analysis is evaluating stationarity to characterize the stability of statistical properties across the spatial domain. The article explores methodologies from the literature, providing insights into the challenges and advancements in functional geostatistics. This work is relevant from theoretical and practical perspectives, offering an integrated view of methodologies tailored to the specific stationarity conditions of the functional processes under study. The practical implications of our work span across fields like environmental monitoring, geosciences, and biomedical research. This overview encourages advancements in functional geostatistics, paving the way for the development of innovative techniques for analyzing and predicting spatially correlated functional data. It lays the groundwork for future research, enhancing our understanding of spatial statistics and its applications.

Keywords: functional data; geostatistics; kriging; non-stationarity; spatial prediction; stationarity

MSC: 62H11; 60G10



Citation: Giraldo, R.; Leiva, V.; Castro, C. An Overview of Kriging and Cokriging Predictors for Functional Random Fields. *Mathematics* **2023**, *11*, 3425. <https://doi.org/10.3390/math11153425>

Academic Editor: Jiansang Zhuang

Received: 13 July 2023

Revised: 27 July 2023

Accepted: 1 August 2023

Published: 7 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The functional data analysis (FDA) plays a critical role in various fields, as it involves data depicted as curves, surfaces, or high-dimensional objects. The broad applicability of the FDA across diverse domains, emphasizing its capacity to discern complex data patterns, was illustrated in [1]. Geostatistics, as detailed in [2] and foundational works like [3], is an area known for providing a comprehensive framework in the modeling of spatial dependence structure of data [4]. Geostatistics facilitates estimation, prediction, and uncertainty quantification of this type of data. The core concepts of geostatistics were outlined in [5], while the concept of geostatistical functional data analysis (GFDA) has its roots in early studies that aimed to combine geostatistics with FDA [6–9]. It was further formalized and refined in [10], highlighting its application in merging the spatial data domain with geostatistical techniques and FDA.

In light of the emphasis on geostatistics, it is essential to acknowledge the foundational works that have shaped the understanding of spatio-temporal geostatistical frameworks [11–13]. This understanding is evident when we consider the exploration of the physical geometry concept, which is pivotal for spatio-temporal geostatistical hydrology [14]. There are also advanced perspectives on environmental health modeling [15]. These insights have played a crucial role in refining geostatistical methodologies, especially when analyzing environmental variables [16]. Significant contributions include the integration of innovative approaches, such as the analysis of neural network residual data in unique environmental contexts [17]. Moreover, the recent fusion of quantitative analyses with spatio-temporal considerations highlights the evolving nature of geostatistical studies [18].

Practical implications of the GFDA span various sectors such as climate science [19], agriculture [20,21], oceanology [22], environmental monitoring [23], geology [24], epidemiology [25], and pollution [26], among others. Through the GFDA, we seek to enhance our comprehension of spatial patterns and relationships, leading to informed decision-making and increased understanding of complex spatial phenomena. The GFDA aligns with the kriging methods used for spatial modeling as outlined in [27,28], by extending traditional geostatistical methods, primarily designed for uni-and-multivariate data, to accommodate functional observations. In [29], it was postulated a proposition for a wavelet regression stemming from FDA based on spatial correlation. In [30], a conditional structure based on stationary functional processes [31] employing a random sampling design was presented, whereas in [32], the robustness for spatio-temporal data was discussed.

As the geostatistical field continues to evolve, recent advancements have addressed nuanced statistical aspects and their real-world applications. The intricacies of spatial autocorrelation, particularly concerning unorthodox random variables, have been explored, emphasizing their broader implications [33]. In the era of expansive datasets, the need for robust estimation techniques has become paramount, leading to the development of methods tailored for reduced rank models to accommodate large spatial data [34]. This extends into practical applications, as showcased by the integration of the spatial eigenvector methodology in regularized regression, enhancing predictive precision in domains such as property valuation [35]. The contemporary relevance of geostatistical tools is further underscored by their application in global challenges, such as the COVID-19 pandemic, where both Bayesian and nonparametric geostatistical models have been employed for data analysis [36].

The intent of this article is to furnish an exhaustive overview concerning the techniques utilized for spatial estimation via kriging and cokriging techniques in the realm of geospatially cohesive functional data. We aim to extend the analysis presented in [10] by exploring the breadth and depth of applicability of these techniques across numerous disciplines and their potential in enhancing the understanding of spatial patterns and relationships. Therefore, our primary objective is to deepen the GFDA framework, with a special emphasis on its synergy with kriging techniques, building on and expanding the foundational study presented in [10]. Our work aims to clearly convey this emphasis. In pursuit of our objective, we present a detailed review of various methods for kriging and cokriging prediction based on stationary and non-stationary functional random fields.

After this introduction, in Section 2, we expand on the investigation presented in [5], discussing the methodologies for stationary functional data, including ordinary kriging and continuous time-varying kriging. We also delve into cokriging prediction based on observations of stationary functional random fields and the functional kriging total model. In Section 3, we shift our focus to the prediction of non-stationary functional random fields. Here, we specifically discuss on universal kriging and residual kriging with external drift, areas that remain less explored in the literature.

Our comprehensive overview of interpolative and co-interpolative techniques for geospatial estimation of functional data represents a significant contribution to the advancement of GFDA. By enhancing the integration of the FDA and geostatistics, we hope to open new avenues for research and applications across a variety of fields. The methodologies presented here may provide a better understanding of spatial processes and patterns, leading to improved decision-making and management of spatially referenced functional observations [27,28].

This article unfolds under the following structure. Section 2 is dedicated to a rigorous exploration of kriging and cokriging predictions tailored for stationary functional random fields. We delve into the foundational methodologies and elucidate their broader implications. In Section 3, we pivot our attention towards kriging predictions designed for non-stationary functional random fields. Concluding our discussion, Section 4 serves as a platform for reflective synthesis of our findings, leading us to propose recommendations that might guide future research in the domain of GFDA.

2. Kriging and Cokriging Prediction for Stationary Functional Random Fields

In this section, we provide background on the GFDA under the assumption of stationarity [27]. This section provides a detailed exploration of how kriging and cokriging can be applied to stationary functional random fields. By examining several variations of the GFDA, it contributes to a nuanced understanding of how this analysis can adapt to different circumstances and requirements. It offers valuable insight into both the practical implementation of the GFDA and the theoretical considerations that drive its use.

2.1. Context

Consider the stochastic functional process $\zeta_s(t)$, where s belongs to a domain $D \subseteq \mathbb{R}^d$ (commonly, $d = 2$). This process is indexed by time t within T . For each specific s in D , $\zeta_s(t)$ acts as a functional variable. Observations are made at locations s_1, \dots, s_n within D , represented as $\zeta_{s_1}(t), \dots, \zeta_{s_n}(t)$. The goal is to estimate the functional variable $\zeta_{s_0}(t)$ at an unsampled location s_0 .

For every $t \in T$, we operate under the assumption of a second-order stationary and isotropic process, ensuring consistent statistical properties throughout the domain. Specifically, for the process $\zeta_s(t)$, its expected value is $E(\zeta_s(t)) = \mu(t)$, for all $t \in T$ and $s \in D$; its variance is $\text{Var}(\zeta_s(t)) = \sigma^2(t)$; and its covariance is $\text{Cov}(\zeta_{s_i}(t), \zeta_{s_j}(t)) = \text{Cov}(h; t)$, which is a function of the distance $h = \|s_i - s_j\|$, for any locations $s_i, s_j \in D$. Then, its semi-variogram is defined by $(1/2)\text{Var}(\zeta_{s_i}(t) - \zeta_{s_j}(t))$ and denoted by $q(h; t) = q_{s_i s_j}(t)$.

2.2. Standard Kriging for Spatially Correlated Functional Data

The kriging estimator for ζ_{s_0} is given in vectorial notation as [27]

$$\hat{\zeta}_{s_0}(t) = \boldsymbol{\eta}^\top \cdot \boldsymbol{\zeta}(t), \tag{1}$$

where $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)^\top$ represents the coefficient vector and $\boldsymbol{\zeta}(t) = (\zeta_{s_1}(t), \dots, \zeta_{s_n}(t))^\top$ denotes the vector of observed values. The symbol “ \cdot ” used in (1) indicates the scalar product between the vectors $\boldsymbol{\eta}$ and $\boldsymbol{\zeta}(t)$. Optimal values for $\boldsymbol{\eta}$ as defined in (1) are determined by solving the subsequent system presented as

$$\begin{pmatrix} \int_T q_{s_1 s_1}(t) dt & \dots & \int_T q_{s_1 s_n}(t) dt & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \int_T q_{s_n s_1}(t) dt & \dots & \int_T q_{s_n s_n}(t) dt & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \\ -\nu \end{pmatrix} = \begin{pmatrix} \int_T q_{s_0 s_1}(t) dt \\ \vdots \\ \int_T q_{s_0 s_n}(t) dt \\ 1 \end{pmatrix}, \tag{2}$$

where ν is the Lagrange multiplier corresponding to the unbiasedness constraint.

The integral representation $q(h) = \int_T q_{s_i s_j}(t) dt$ stated in (2), with $h = \|s_i - s_j\|$, is known as the trace-semi-variogram [27]. For the existence of this integral, certain conditions are vital. Specifically, T should be bounded, and the corresponding stochastic functional process ζ must be compactly-supported. Holding these conditions ensures that the trace-semi-variogram provides an accurate representation of the spatial dependence structure.

Given the pairwise function stated as $F_{ij}(t) = (\zeta_{s_i}(t) - \zeta_{s_j}(t))^2$, defined over the time domain T , and the set expressed as $n(h) = \{(s_i, s_j): \|s_i - s_j\| = h\}$, the function related to the semi-variogram can be formulated as

$$\hat{q}(h) = \frac{1}{2|n(h)|} \sum_{(i,j) \in n(h)} \int_T F_{ij}(t) dt,$$

where $|n(h)|$ represents the total number of distinct pairs in $n(h)$. This formulation of the semi-variogram emphasizes the significance of spatial relationships within the data. By integrating over the time domain and considering pairwise differences, we gain a more comprehensive insight into the spatial dependencies inherent in the dataset.

2.3. Continuous Temporal Fluctuating Interpolation for Functional Data

We provide an overview of the theoretical framework based on [28], considering functional parameters depicted as $\eta_i(t)$, where i spans the set from 1 to n . This leads to

$$\widehat{\zeta}_{s_0}(t) = \boldsymbol{\eta}(t)^\top \cdot \boldsymbol{\zeta}(t) \tag{3}$$

with $\boldsymbol{\eta}(t) = (\eta_1(t), \dots, \eta_n(t))^\top$ being the time-dependent coefficients and the vector of observed values being $\boldsymbol{\zeta}(t) = (\zeta_{s_1}(t), \dots, \zeta_{s_n}(t))^\top$.

Thus, the time-varying nature of the coefficient vector $\boldsymbol{\eta}(t)$ arises from the functional character of the data. As the observed data vary over the time domain T , the weight or importance assigned to each observation must be flexible, adjusting over time. This variation captures the continuous shifts in data behavior across the domain.

In contrast, traditional ordinary kriging primarily deals with scalar spatial data, where each spatial location s has a static value. Our approach, focused on functional data, captures intrinsic variability over time, distinguishing it from traditional methods.

The functional weights, $\eta_i(t)$ namely, change over T , denoting a dynamic weighting approach different from the traditional method where weights remain constant. The concept of minimizing the integrand term-by-term might seem similar to minimizing the overall integral, but this is not always the case in FDA. Given potential interdependencies of functions over T , our approach explicitly addresses the inherent integrated characteristics of functional data, rather than analyzing each temporal instance in isolation.

Having established the nature of functional weights, we next consider the problem of determining these weights mathematically. State the functional parameters $\eta_i(t)$ as mentioned in (3) by solving the minimization problem formulated as

$$\underset{\eta_1(t), \dots, \eta_n(t)}{\text{minimize}} \left\{ \int_T \text{Var}(\widehat{\zeta}_{s_0}(t) - \zeta_{s_0}(t)) dt \right\}.$$

Suppose that we can represent each observed function with a combination of K basis functions, B_l , for $l \in \{1, \dots, K\}$, as suggested in [1]. This representation can be mathematically articulated as

$$\begin{aligned} \zeta_{s_j}(t) &= \mathbf{a}_j^\top \mathbf{B}(t) \\ &= \sum_{l=1}^K a_{jl} B_l(t), \quad \forall j \in \{1, \dots, n\}. \end{aligned} \tag{4}$$

The assumptions concerning stationarity are:

- (i) The expected value of \mathbf{a}_j is consistent across locations. In more specific terms, we have $E(\mathbf{a}_j) = \mathbf{v}_j$ for every location j , where \mathbf{v}_j is a location-independent constant vector.
- (ii) The covariance matrix $\boldsymbol{\Sigma}_{ij}$ depends solely on the relative distance or difference between locations i and j , rather than their absolute positions.

To ensure stationarity within our framework, the linear model of co-regionalization is subject to certain restrictions. Specifically, this model assumes that the covariances between different components of the multivariate field depend on the spatial difference only through a set of shared univariate functions.

Given the premise that $\zeta_{s_j}(t)$, for each j within the set $\{1, \dots, n\}$, represents spatially dependent stochastic functions, the assumption is that matrix \mathbf{A} , constituted by coefficients denoted a_{jl} , is characterized as

$$\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1K} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nK} \end{pmatrix} = (\mathbf{a}_1, \dots, \mathbf{a}_K).$$

Note that \mathbf{A} constitutes a K -variate random field [37] of mean $E(\mathbf{a}_l) = \mathbf{v}_l$ and matrix of covariance Σ defined as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \dots & \Sigma_{1K} \\ \vdots & \ddots & \vdots \\ \Sigma_{K1} & \dots & \Sigma_{KK} \end{pmatrix}, \tag{5}$$

where $\Sigma_{lr} = \text{Cov}(\mathbf{a}_l, \mathbf{a}_r)$. Here, a_{lr} are assumed to be samples from the spatial random field \mathbf{a}_l , for l in the set $\{1, \dots, K\}$. We employ the framework of a co-regionalization linear model, as detailed in [37,38], for the estimation of the matrix presented in (5).

When stating unbiasedness for making a prediction using (3), each parameter $\eta_i(t)$ is represented as a linear combination of the basis functions B_l with coefficients b_{il} . Mathematically, we have that

$$\eta_i(t) = \sum_{l=1}^K b_{il} B_l(t) = \mathbf{b}_i^\top \mathbf{B}(t), \quad i \in \{1, \dots, n\}. \tag{6}$$

By using the formulas defined in (4) and (6), the predictor stated in (3) is now formulated as

$$\hat{\zeta}_{s_0}(t) = \sum_{i=1}^n \mathbf{b}_i^\top \mathbf{B}(t) \mathbf{a}_i^\top \mathbf{B}(t) = \sum_{i=1}^n \mathbf{b}_i^\top \mathbf{B}(t) \mathbf{B}^\top(t) \mathbf{a}_i.$$

The constraint of unbiasedness, expanded using the constant equal to one, is stated as

$$\sum_{l=1}^K c_l B_l(t) = \mathbf{c}^\top \mathbf{B}(t) = 1.$$

In this context, \mathbf{c} is a coefficient vector ensuring that the unbiasedness constraint is satisfied for all values of t by making sure the linear combination of the basis functions $B_l(t)$ yields a constant value equal to one. This constraint may be presented by

$$\sum_{i=1}^n \mathbf{b}_i^\top \mathbf{B}(t) = \mathbf{c}^\top \mathbf{B}(t), \quad \forall t \in T,$$

which is analogous to $\sum_{i=1}^n \mathbf{b}_i = \mathbf{c}$, or to $\sum_{i=1}^n b_{i1} = c_1, \dots, \sum_{i=1}^n b_{iK} = c_K$.

While constructing the dispersion of the objective function, it is pertinent to note that the coefficient vector \mathbf{a}_0 corresponds to the functional expansion of $\zeta_{s_0}(t)$, where $\zeta_{s_0}(t) = \mathbf{a}_0^\top \mathbf{B}(t) = \sum_{l=1}^K a_{0l} B_l(t)$. Thus, we have that

$$\begin{aligned} \text{Var}(\hat{\zeta}_{s_0}(t) - \zeta_{s_0}(t)) &= \text{Var}\left(\sum_{i=1}^n \mathbf{b}_i^\top \mathbf{B}(t) \mathbf{B}^\top(t) \mathbf{a}_i\right) + \mathbf{B}^\top(t) \text{Var}(\mathbf{a}_0) \mathbf{B}(t) \\ &\quad - 2 \sum_{i=1}^n \mathbf{b}_i^\top \mathbf{B}(t) \mathbf{B}^\top(t) \text{Cov}(\mathbf{a}_i, \mathbf{a}_0) \mathbf{B}(t) \\ &= \sum_{i=1}^n \mathbf{b}_i^\top \mathbf{B}(t) \mathbf{B}^\top(t) \text{Var}(\mathbf{a}_i) \mathbf{B}(t) \mathbf{B}^\top(t) \mathbf{b}_i \\ &\quad + 2 \sum_{i < j} \mathbf{b}_i^\top \mathbf{B}(t) \mathbf{B}^\top(t) \text{Cov}(\mathbf{a}_i, \mathbf{a}_j) \mathbf{B}(t) \mathbf{B}^\top(t) \mathbf{b}_j \\ &\quad + \mathbf{B}^\top(t) \text{Var}(\mathbf{a}_0) \mathbf{B}(t) \\ &\quad - 2 \sum_{i=1}^n \mathbf{b}_i^\top \mathbf{B}(t) \mathbf{B}^\top(t) \text{Cov}(\mathbf{a}_0, \mathbf{a}_i) \mathbf{B}(t). \end{aligned} \tag{7}$$

From the expression given in (7), for $i < j$, with i, j in the set $\{1, \dots, n\}$, we get

$$\text{Cov}(\mathbf{a}_i, \mathbf{a}_j) = \begin{pmatrix} \text{Cov}(a_{i1}, a_{j1}) & \dots & \text{Cov}(a_{i1}, a_{jK}) \\ \vdots & \ddots & \vdots \\ \text{Cov}(a_{iK}, a_{j1}) & \dots & \text{Cov}(a_{iK}, a_{jK}) \end{pmatrix}, \quad i, j \in \{1, \dots, n\}. \quad (8)$$

For $i = j$, the expression given in (8) denotes a matrix associated with \mathbf{a}_i , wherein the diagonal components correspond to the variances of the components \mathbf{a}_i .

Consider the definitions, for $i, j \in \{1, \dots, n\}$, stated as

$$\begin{aligned} Q_i &= \int_T \mathbf{B}(t)\mathbf{B}^\top(t)\text{Var}(\mathbf{a}_i)\mathbf{B}(t)\mathbf{B}^\top(t)dt, \\ Q_{ij} &= \int_T \mathbf{B}(t)\mathbf{B}^\top(t)\text{Cov}(\mathbf{a}_i, \mathbf{a}_j)\mathbf{B}(t)\mathbf{B}^\top(t)dt, \\ D &= \int_T \mathbf{B}^\top(t)\text{Var}(\mathbf{a}_0)\mathbf{B}(t)dt, \\ J_i &= \int_T \mathbf{B}(t)\mathbf{B}^\top(t)\text{Cov}(\mathbf{a}_0, \mathbf{a}_i)\mathbf{B}(t)dt. \end{aligned}$$

where the matrix D is determined through $\widehat{\text{Var}}(\mathbf{a}_0)$, which stems from the established linear model of co-regionalization.

Taking into account K Lagrange constraints represented as $\mathbf{v}^\top = (v_1, \dots, v_K)$, the objective function to be optimized is articulated as

$$\text{minimize}_{\mathbf{b}_1, \dots, \mathbf{b}_n} \left\{ \sum_{i=1}^n \mathbf{b}_i^\top Q_i \mathbf{b}_i + 2 \sum_{i < j} \mathbf{b}_i^\top Q_{ij} \mathbf{b}_j + D - 2 \sum_{i=1}^n \mathbf{b}_i^\top J_i + 2\mathbf{v}^\top \sum_{i=1}^n (\mathbf{b}_i - \mathbf{c}) \right\}. \quad (9)$$

By defining $\boldsymbol{\beta} = (\mathbf{b}_1^\top, \dots, \mathbf{b}_n^\top, \mathbf{v}^\top)^\top$, the expression established in (9) is reframed as

$$\text{minimize}_{\boldsymbol{\beta}} \left\{ \boldsymbol{\beta}^\top Q \boldsymbol{\beta} + D - 2\boldsymbol{\beta}^\top J \right\}, \quad (10)$$

with a structure presented by

$$Q = \begin{pmatrix} Q_1 & \dots & Q_{1n} & \mathbf{I} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{n1} & \dots & Q_{nn} & \mathbf{I} \\ \mathbf{I} & \dots & \mathbf{I} & \mathbf{0} \end{pmatrix}, \quad J = \begin{pmatrix} J_1 \\ \vdots \\ J_n \\ \mathbf{c} \end{pmatrix}. \quad (11)$$

The unit matrix \mathbf{I} given in (11) has rank K . When optimizing the formulation stated in (10) in terms of $\boldsymbol{\beta}$, one obtains $2Q\boldsymbol{\beta} - 2J = \mathbf{0}$, which implies

$$J = Q\boldsymbol{\beta}, \quad Q^{-1}J = \widehat{\boldsymbol{\beta}}.$$

The plug-in estimate for the integrated prediction variance is represented as

$$\widehat{\sigma}_{s_0}^2 = \int_T \text{Var}(\widehat{\zeta}_{s_0}(t) - \zeta_{s_0}(t))dt.$$

The consolidated predictive variance $\widehat{\sigma}_{s_0}^2$ offers insight into the variability associated with forecasting a complete curve.

2.4. Cokriging Prediction Based on Observations of Stationary Functional Random Fields

This subsection is based on [23]. Let $\{Z(s), s \in D\}$ be a multivariate random field, where $Z(s) \in \mathbb{R}^p$ and $D \subset \mathbb{R}^d$. Assume we have an observation $\{Z(s_1), \dots, Z(s_n)\}$ and that

$$Z(s_i) = \begin{pmatrix} Z_1(s_i) \\ \vdots \\ Z_p(s_i) \end{pmatrix} = \begin{pmatrix} \mu_1(s_i) \\ \vdots \\ \mu_p(s_i) \end{pmatrix} + \begin{pmatrix} \varepsilon_1(s_i) \\ \vdots \\ \varepsilon_p(s_i) \end{pmatrix}, \quad i \in \{1, \dots, n\}.$$

Let $\mu_j(s)$ denote the mean of variable j , which is considered to be constant across all $s \in D$, and let $\varepsilon_j(s)$ represent the random error. The cokriging predictor of variable j at location s_0 is expressed as

$$\widehat{Z}_j(s_0) = \sum_{i=1}^n \eta_{i1} Z_1(s_i) + \dots + \sum_{i=1}^n \eta_{ip} Z_p(s_i) = \sum_{i=1}^n \sum_{j=1}^p \eta_{ij} Z_j(s_i).$$

Now, let us shift our attention to a stochastic process with functional outputs, defined as $\{\zeta_s(t), s \in D, t \in T \subset \mathbb{R}\}$. With the available observations $\{\zeta_{s_1}(t), \dots, \zeta_{s_n}(t)\}$, we estimate the value of the transformation at the specific pair $(s_0 \in D, v \in T)$.

The challenge in this context is to adapt the traditional cokriging predictor for these curves (functional data). This adaptation involves utilizing functional analogs in place of the usual parameters and shifting from random variables to a series of functional variables, like $\zeta_{s_i}(t)$, with i spanning from 1 to n . Details of the adaptation are elaborated in the expressions stated in (12) and (13) and presented next.

- Parameters

Multivariate cokriging	\Rightarrow	Cokriging employing curves	
$\eta_{11} \dots \eta_{1p}$	\Rightarrow	$\eta_1^v(t), t \in T$	(12)
\vdots	\vdots	\vdots	
$\eta_{n1} \dots \eta_{np}$	\Rightarrow	$\eta_n^v(t), t \in T$	

- Variables

Multivariate cokriging	\Rightarrow	Cokriging employing curves	
$Z_1(s_1) \dots Z_p(s_1)$	\Rightarrow	$\zeta_{s_1}(t), t \in T$	(13)
\vdots	\vdots	\vdots	
$Z_1(s_n) \dots Z_p(s_n)$	\Rightarrow	$\zeta_{s_n}(t), t \in T$	

Hence, within the realm of functional data, the cokriging predictor for $\zeta_{s_0}(v)$ is depicted by

$$\widehat{\zeta}_{s_0}(v) = \sum_{i=1}^n \int_T \eta_i^v(t) \zeta_{s_i}(t) dt. \tag{14}$$

For each $v \in T$, the functional coefficients $\eta_i^v(t)$ given in (14) adhere to traditional geostatistical requirements like unbiasedness and minimal prediction variance.

In addressing the complexities of our functional data, we adopt a strategy rooted in basis functions. This involves representing both variables and coefficients through expansions, for $i \in \{1, \dots, n\}$, given by

$$\zeta_{s_i}(t) = \sum_{l=1}^K a_{il} B_l(t) = \mathbf{a}_i^\top \mathbf{B}(t), \tag{15}$$

$$\eta_i^v(t) = \sum_{l=1}^K b_{ilv} B_l(t) = \mathbf{b}_{iv}^\top \mathbf{B}(t), \tag{16}$$

where K denotes the total number of basis functions. The choice of an optimal K is data-dependent and can be determined using cross-validation techniques. By utilizing the expansions given by (15) and (16), the predictor stated in (14) undergoes a transformation, yielding

$$\widehat{\zeta}_{s_0}(v) = \sum_{i=1}^n \int_T \mathbf{b}_{iv}^\top \mathbf{B}(t) \mathbf{B}^\top(t) \mathbf{a}_i dt = \sum_{i=1}^n \mathbf{b}_{iv}^\top \mathbf{W} \mathbf{a}_i. \tag{17}$$

The integral product is defined as

$$\mathbf{W} = \int_T \mathbf{B}(t) \mathbf{B}^\top(t) dt. \tag{18}$$

For commonly adopted orthonormal bases, like the Fourier series, the Gram matrix \mathbf{W} stated in (18) is recognized to be the identity matrix as referenced in (18). Conversely, for bases like B-splines, the determination of \mathbf{W} requires numerical integration. Assuming stationary behavior in the random functions, we deduce that

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1K} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nK} \end{pmatrix} = (\mathbf{a}_1, \dots, \mathbf{a}_K).$$

This establishes a K -variate random field with its mean and covariance matrix respectively expressed as

$$\mathbf{E}(\mathbf{a}_l) = \begin{pmatrix} \mathbf{E}(a_{1j}) \\ \vdots \\ \mathbf{E}(a_{nl}) \end{pmatrix} = \begin{pmatrix} \vartheta_l \\ \vdots \\ \vartheta_l \end{pmatrix}, \quad l \in \{1, \dots, K\}, \tag{19}$$

and

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \cdots & \boldsymbol{\Sigma}_{1K} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_{K1} & \cdots & \boldsymbol{\Sigma}_{KK} \end{pmatrix}, \tag{20}$$

where $\boldsymbol{\Sigma}_{lr} = \text{Cov}(\mathbf{a}_l, \mathbf{a}_r)$ depicts the covariance between \mathbf{a}_l and \mathbf{a}_r . Consequently, employing the formula given in (19), we derive

$$\mathbf{E}(\mathbf{a}_i) = \begin{pmatrix} \mathbf{E}(a_{i1}) \\ \vdots \\ \mathbf{E}(a_{iK}) \end{pmatrix} = \begin{pmatrix} \vartheta_1 \\ \vdots \\ \vartheta_K \end{pmatrix} = \boldsymbol{\vartheta}, \quad i \in \{1, \dots, n\}.$$

Hence, the mean predictor presented in (17) is characterized as

$$\mathbf{E}(\widehat{\zeta}_{s_0}(v)) = \sum_{i=1}^n \mathbf{b}_{iv}^\top \mathbf{W} \mathbf{E}(\mathbf{a}_i) = \sum_{i=1}^n \mathbf{b}_{iv}^\top \mathbf{W} \boldsymbol{\vartheta}.$$

Moreover, the mean unobserved function at s_0 (location) and v (time) is formulated by

$$\mathbf{E}(\zeta_{s_0}(v)) = \sum_{l=1}^K \mathbf{E}(a_{0l}) B_l(v) = \mathbf{B}^\top(v) \mathbf{E}(\mathbf{a}_0) = \mathbf{B}^\top(v) \boldsymbol{\vartheta}.$$

Therefore, the suggested predictor holds with the unbiasedness condition if

$$\sum_{i=1}^n \mathbf{b}_{iv}^\top = \mathbf{B}^\top(v) \mathbf{W}^{-1}. \tag{21}$$

While Gram matrices are generally positive semi-definite, ensuring their strict positive definiteness can be challenging. However, in the context of our specific Gram matrix \mathbf{W} defined in (18), its positive definiteness is guaranteed. This arises from the fact that the functions $B_l(t)$, for $l \in \{1, \dots, K\}$, constitute a basis set and are linearly independent. As a direct consequence, the inverse matrix \mathbf{W}^{-1} exists, ensuring the well-defined nature of the expression presented in (21).

To identify the optimal linear unbiased estimator, the functional parameters within our proposed predictor are obtained by addressing the optimization problem stated as

$$\underset{\eta_1^v, \dots, \eta_n^v}{\text{minimize}} \{ \text{Var}(\widehat{\zeta}_{s_0}(v) - \zeta_{s_0}(v)) \}, \tag{22}$$

subject to $E(\widehat{\zeta}_{s_0}(v)) = E(\zeta_{s_0}(v))$.

Obtaining the variance from the objective function formulated in (22), we find that

$$\begin{aligned} \text{Var}(\widehat{\zeta}_{s_0}(v) - \zeta_{s_0}(v)) &= \text{Var}(\widehat{\zeta}_{s_0}(v)) + \text{Var}(\zeta_{s_0}(v)) - 2\text{Cov}(\widehat{\zeta}_{s_0}(v), \zeta_{s_0}(v)) \\ &= \sum_{i=1}^n \mathbf{b}_{iv}^\top \mathbf{W} \text{Var}(\mathbf{a}_i) \mathbf{W}^\top \mathbf{b}_{iv} + 2 \sum_{i < j} \mathbf{b}_{iv}^\top \mathbf{W} \text{Cov}(\mathbf{a}_i, \mathbf{a}_j) \mathbf{W}^\top \mathbf{b}_{jv} \\ &\quad + \mathbf{B}^\top(v) \text{Var}(\mathbf{a}_0) \mathbf{B}(v) - 2 \sum_{i=1}^n \mathbf{b}_{iv}^\top \mathbf{W} \text{Cov}(\mathbf{a}_0, \mathbf{a}_i) \mathbf{B}(v). \end{aligned} \tag{23}$$

In this context, $\text{Var}(\mathbf{a}_i)$ and $\text{Var}(\mathbf{a}_0)$ denote the variances of \mathbf{a}_i and \mathbf{a}_0 , respectively. Similarly, $\text{Cov}(\mathbf{a}_i, \mathbf{a}_j)$ signifies the covariance between \mathbf{a}_i and \mathbf{a}_j , while $\text{Cov}(\mathbf{a}_0, \mathbf{a}_i)$ represents the covariance of \mathbf{a}_0 with \mathbf{a}_i . These metrics can be computed from Σ given in (20) such as previously determined. However, multivariate geostatistics [37], and specifically a linear model of co-regionalization [37], can be employed to estimate these matrices. It should be noted that, due to the stationary random functions in \mathbf{A} , $\text{Var}(\mathbf{a}_i)$ and $\text{Var}(\mathbf{a}_0)$ are identical. Furthermore, $\text{Cov}(\mathbf{a}_0, \mathbf{a}_i)$ is solely dependent on the distances and not on their locations. As a result, they may be directly obtained using Σ . Defining

$$\mathbf{M}_i = (\mathbf{W} \text{Var}(\mathbf{a}_i) \mathbf{W}^\top), \quad \mathbf{M}_{ij} = (\mathbf{W} \text{Cov}(\mathbf{a}_i, \mathbf{a}_j) \mathbf{W}^\top), \quad \mathbf{N}_i(v) = (\mathbf{W} \text{Cov}(\mathbf{a}_0, \mathbf{a}_i) \mathbf{B}(v)), \quad \mathbf{D}(v) = \mathbf{B}^\top(v) \text{Var}(\mathbf{a}_0) \mathbf{B}(v),$$

the expression given in (23) may be presented as

$$\text{Var}(\widehat{\zeta}_{s_0}(v) - \zeta_{s_0}(v)) = \sum_{i=1}^n \mathbf{b}_{iv}^\top \mathbf{M}_i \mathbf{b}_{iv} + 2 \sum_{i < j} \mathbf{b}_{iv}^\top \mathbf{M}_{ij} \mathbf{b}_{jv} + \mathbf{D}(v) - 2 \sum_{i=1}^n \mathbf{b}_{iv}^\top \mathbf{N}_i(v). \tag{24}$$

From the expression formulated in (24) and assuming the Lagrange multipliers $\mathbf{v}_v^\top = (v_{1v}, \dots, v_{Kv})$, the optimization problem formulated in (22) may be stated as

$$\underset{\mathbf{b}_{1v}, \dots, \mathbf{b}_{nv}, \mathbf{v}_v}{\text{minimize}} \left\{ \sum_{i=1}^n \mathbf{b}_{iv}^\top \mathbf{M}_i \mathbf{b}_{iv} + 2 \sum_{i < j} \mathbf{b}_{iv}^\top \mathbf{M}_{ij} \mathbf{b}_{jv} + \mathbf{D}(v) - 2 \sum_{i=1}^n \mathbf{b}_{iv}^\top \mathbf{N}_i(v) + 2 \mathbf{v}_v^\top \left(\sum_{i=1}^n \mathbf{b}_{iv} - \mathbf{W}^{-1} \mathbf{B}(v) \right) \right\}.$$

Considering $\boldsymbol{\beta}_v = (\mathbf{b}_{1v}^\top, \dots, \mathbf{b}_{nv}^\top, \mathbf{v}_v^\top)^\top$, we have

$$\underset{\boldsymbol{\beta}_v}{\text{minimize}} \left\{ \boldsymbol{\beta}_v^\top \mathbf{M} \boldsymbol{\beta}_v + \mathbf{D}(v) - 2 \boldsymbol{\beta}_v^\top \mathbf{N}(v) \right\}, \tag{25}$$

with

$$M = \begin{pmatrix} M_1 & \dots & M_{1n} & I \\ \vdots & \ddots & \vdots & \vdots \\ M_{n1} & \dots & M_{nn} & I \\ I & \dots & I & \mathbf{0} \end{pmatrix}, \quad N(v) = \begin{pmatrix} N_1(v) \\ \vdots \\ N_n(v) \\ W^{-1}B(v) \end{pmatrix}.$$

Solving the equation given in (25) for β_v , we find $2M\beta_v - 2N(v) = \mathbf{0}$, which implies that $M\beta_v = N(v)$ and then $\hat{\beta}_v = M^{-1}N(v)$.

For the prediction variance, we use the notation established as

$$\sigma_{s_0}^2(v) = \text{Var}(\hat{\zeta}_{s_0}(v) - \zeta_{s_0}(v)),$$

with its plug-in estimate being given by

$$\hat{\sigma}_{s_0}^2(v) = \hat{\beta}_v^\top M \hat{\beta}_v + D(v) - 2\hat{\beta}_v^\top N(v),$$

where the matrix $D(v)$ is derived from an estimated value of $\widehat{\text{Var}}(a_0)$ via the fitted linear model of co-regionalization.

2.5. Holistic Functional Kriging Approach

Next, we synthesize the findings presented in [39]. The predictor is delineated by

$$\hat{\zeta}_{s_0}(t) = \sum_{i=1}^n \int_T \eta_i(v, t) \zeta_{s_i}(v) dv. \tag{26}$$

This formulation extends the predictor introduced in (14). The functional parameter $\eta_i(v, t)$ gauges the effect of i -th location at v on an unsampled function at location t . The formulation is in harmony with the linear framework designed for responses with functional structure, often denoted as the holistic model detailed in [1].

In the realm of geostatistical analyses, we present two distinct cokriging methodologies: the traditional multivariate cokriging and the more nuanced functional cokriging based on curves. The pivotal distinction between these methodologies hinges upon their treatment of parameters. In the traditional multivariate approach, each spatial location is associated with a set of fixed cokriging coefficients, representing the influence of each observed variable on the prediction. Conversely, in the functional cokriging model, these fixed coefficients evolve into functions, varying across a specified domain t . This introduces an element of flexibility, enabling the model to capture more intricate spatial relationships. While the former model operates directly upon observed variables, the latter delves into a functional setting, emphasizing the impact of each observed function over a domain. The functional cokriging provides an enhanced analytical perspective, aligning closely with the model tailored to functional data. This pivotal distinction between functional cokriging and other methods merits heightened attention to deepen the understanding of the model structures and implications. To delineate functional variables, we employ the basis functions denoted as $B(v)$. These are instrumental for expanding our variables and parameters into known terms. Mathematically, the functional variable $\zeta_{s_i}(v)$ and the functional parameter $\eta_i(v, t)$ can be articulated using these basis functions as

$$\zeta_{s_i}(v) = \sum_{l=1}^K a_{il} B_l(v) = \mathbf{B}^\top(v) \mathbf{a}_i, \tag{27}$$

$$\eta_i(v, t) = \sum_{j=1}^K \sum_{l=1}^K c_{jl}^i B_j(v) B_l(t) = \mathbf{B}^\top(t) \mathbf{C}_i \mathbf{B}(v), \quad i \in \{1, \dots, n\}. \tag{28}$$

The coefficients c_{jl}^i in C_i allow us to model interactions between various basis functions, pivotal to capturing the data intricacy. The matrix C_i is structured as

$$C_i = \begin{pmatrix} c_{11}^i & \dots & c_{K1}^i \\ \vdots & \ddots & \vdots \\ c_{1K}^i & \dots & c_{KK}^i \end{pmatrix}, \quad i \in \{1, \dots, n\}.$$

Now, considering the prediction, $\widehat{\zeta}_{s_0}(t)$ namely, which estimates the expected function value at point t , we can express it as

$$\begin{aligned} \widehat{\zeta}_{s_0}(t) &= \sum_{i=1}^n \int_T \mathbf{B}^\top(t) C_i \mathbf{B}(v) \mathbf{B}^\top(v) \mathbf{a}_i dv \\ &= \sum_{i=1}^n \mathbf{B}^\top(t) C_i \mathbf{F} \mathbf{a}_i = \mathbf{B}^\top(t) \sum_{i=1}^n C_i \mathbf{F} \mathbf{a}_i, \end{aligned} \tag{29}$$

where

$$\mathbf{F} = \begin{pmatrix} \int_T B_1^2(v) dv & \dots & \int_T B_1(v) B_K(v) dv \\ \vdots & \ddots & \vdots \\ \int_T B_K(v) B_1(v) dv & \dots & \int_T B_K^2(v) dv \end{pmatrix}.$$

The matrix F encapsulates the inner products of the basis functions, playing a pivotal role in merging the basis functions and the coefficients in the model. In the case of an orthogonal basis, such as the Fourier basis, the matrix F reduces to an identity matrix. For alternative basis functions, like B-spline, it becomes necessary to compute F potentially through the application of numerical integration methods.

Leveraging the vec operator, which vectorizes a matrix by stacking its columns into a single column vector, the formulation presented in (28) can be articulated as

$$\eta_i(v, t) = (\text{vec}(C_i))^\top \mathbf{D}(t) \mathbf{B}(v), \quad i \in \{1, \dots, n\},$$

with

$$\mathbf{D}(t) = \begin{pmatrix} \mathbf{B}(t) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}(t) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{B}(t) \end{pmatrix}.$$

Therefore, the predictor stated in (29) can be also expressed as

$$\begin{aligned} \widehat{\zeta}_{s_0}(t) &= \sum_{i=1}^n \int_T (\text{vec}(C_i))^\top \mathbf{D}(t) \mathbf{B}(v) \mathbf{B}^\top(v) \mathbf{a}_i dv \\ &= \sum_{i=1}^n (\text{vec}(C_i))^\top \mathbf{D}(t) \mathbf{F} \mathbf{a}_i. \end{aligned} \tag{30}$$

Next, unbiasedness is addressed. Assuming that the coefficients, as delineated in the formula presented in (27), for $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, K\}$, are structured in matrix form as

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} a_{11} & \dots & a_{1K} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nK} \end{pmatrix} \\ &= (\mathbf{a}_1, \dots, \mathbf{a}_K). \end{aligned}$$

We can consider A as emanating from a K multivariate random field. In this context, the expected value of a_j is $E(a_j) = E(a_{1j}, \dots, a_{nj})^\top = (\vartheta_j, \dots, \vartheta_j)^\top$, for all j in the set $\{1, \dots, K\}$. The covariance matrix is then presented as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \dots & \Sigma_{1K} \\ \vdots & \ddots & \dots \\ \Sigma_{K1} & \dots & \Sigma_{KK} \end{pmatrix}, \tag{31}$$

where $\Sigma_{ij} = \text{Cov}(a_i, a_j)$. If a_i is a $K \times 1$ vector with elements a_{i1}, \dots, a_{iK} , its expected value is defined as

$$E(a_i^\top) = \boldsymbol{\vartheta}^\top = (\vartheta_1, \dots, \vartheta_K).$$

Thus, the expected function value at an unsampled location s_0 is defined as

$$\begin{aligned} E(\zeta_{s_0}(t)) &= E\left(\sum_{l=1}^K a_{0l} B_l(t)\right) \\ &= E(\mathbf{B}^\top(t) \mathbf{a}_0) \\ &= \mathbf{B}^\top(t) E(\mathbf{a}_0) \mathbf{B}^\top(t) \boldsymbol{\vartheta}. \end{aligned} \tag{32}$$

By obtaining the expectation of the formula detailed in (29), we obtain

$$\begin{aligned} E(\widehat{\zeta}_{s_0}(t)) &= \mathbf{B}^\top(t) \sum_{i=1}^n \text{Cov}_i F E(a_i) \\ &= \mathbf{B}^\top(t) \sum_{i=1}^n \text{Cov}_i F \boldsymbol{\vartheta}. \end{aligned} \tag{33}$$

Note that Cov_i is the covariance matrix of observation i in the estimation process. From the expressions given in (32) and (33), the predictor presented in (26) is unbiased if and only if

$$\mathbf{B}^\top(t) \sum_{i=1}^n \text{Cov}_i F \boldsymbol{\vartheta} = \mathbf{B}^\top(t) \boldsymbol{\vartheta}$$

and

$$\sum_{i=1}^n \text{Cov}_i F \boldsymbol{\vartheta} = \boldsymbol{\vartheta}.$$

Assuming both F and $\sum_{i=1}^n \text{Cov}_i$ hold the condition of being full rank matrices, such a condition is analogous to $\sum_{i=1}^n \text{Cov}_i = F^{-1}$, or, by introducing the vec operation which linearly maps a matrix into a column vector by stacking its columns, the above can be equivalently stated as

$$\sum_{i=1}^n \text{vec}(\text{Cov}_i) = \text{vec}(F^{-1}).$$

Next, minimum variance and parameters estimation is discussed. Given the expressions presented in (27) and (30), the functional parameters within the proposed predictor are calculated (estimated) by optimizing the problem established as

$$\underset{\text{Cov}_1, \dots, \text{Cov}_n}{\text{minimize}} \left\{ \int_T \text{Var} \left(\sum_{i=1}^n (\text{vec}(\text{Cov}_i))^\top D(t) F a_i - \mathbf{B}^\top(t) a_0 \right) dt \right\}, \tag{34}$$

subject to $\sum_{i=1}^n \text{vec}(\text{Cov}_i) = \text{vec}(F^{-1})$.

By expanding the variance established in (34), we obtain

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^n (\text{vec}(\text{Cov}_i))^T D(t) F a_i - B^T(t) a_0\right) &= \sum_{i=1}^n \left((\text{vec}(\text{Cov}_i))^T D(t) F \right) \text{Var}(a_i) \left((\text{vec}(\text{Cov}_i))^T D(t) F \right)^T \\ &+ 2 \sum_{i < j} \left((\text{vec}(\text{Cov}_i))^T D(t) F \right) \text{Cov}(a_i, a_j) \left((\text{vec}(\text{Cov}_j))^T D(t) F \right)^T \\ &+ B^T(t) \text{Var}(a_0) B(t) \\ &- 2 \sum_{i=1}^n \left((\text{vec}(\text{Cov}_i))^T D(t) F \right) \text{Cov}(a_i, a_0) B(t). \end{aligned} \tag{35}$$

In the expression provided in (35), for $i < j$ where $i, j \in \{1, \dots, n\}$, the covariance matrix between vectors a_i and a_j is defined as

$$\text{Cov}(a_i, a_j) = \begin{pmatrix} \text{Cov}(a_{i1}, a_{j1}) & \dots & \text{Cov}(a_{i1}, a_{jK}) \\ \vdots & \ddots & \vdots \\ \text{Cov}(a_{iK}, a_{j1}) & \dots & \text{Cov}(a_{iK}, a_{jK}) \end{pmatrix}.$$

Moreover, it is noteworthy that the variance-covariance matrix of vector a_i can be directly obtained as a special case of the above matrix by setting $i = j$.

Observe that $\text{Var}(a_i) = \text{Var}(a_\bullet)$, for all i . This is due to the stationarity assumption, which implies that the variance does not vary with location. Similarly, the covariance $\text{Cov}(a_i, a_j) = \text{Cov}(a_j, a_i)$ is symmetric and does not depend on the order of i and j . If Σ presented in (31) has been previously estimated, then $\text{Var}(a_i)$ and $\text{Cov}(a_i, a_j)$ are known.

Now, we can utilize multivariate geostatistics [37]. Specifically, we employ a linear co-regionalization framework to estimate these matrices. From (23) and defining

$$\begin{aligned} Q_\bullet &= \int_T D(t) F \text{Var}(a_\bullet) F^T D^T(t) dt = \left(F \text{Var}(a_\bullet) F^T \otimes F \right), \\ Q_{ij} &= \int_T D(t) F \text{Cov}(a_i, a_j) F^T D^T(t) dt = \left(F \text{Cov}(a_i, a_j) F^T \otimes F \right), \\ K &= \left(\int_T \phi^T(t) \text{Var}(a_\bullet) \phi(t) dt \right), \\ J_i &= \int_T D(t) F \text{Cov}(a_i, a_0) \phi(t) dt = \text{vec}(F \text{Cov}(a_i, a_0) F)^T, \end{aligned}$$

where \otimes is the Kronecker product of matrices, the objective function presented in (34) can be rewritten as

$$\underset{C_1, \dots, C_n, \nu}{\text{minimize}} \left\{ \sum_{i=1}^n (\text{vec}(C_i))^T Q_\bullet \text{vec}(C_i) + 2 \sum_{i < j} (\text{vec}(C_i))^T Q_{ij} \text{vec}(C_j) + K - 2 \sum_{i=1}^n (\text{vec}(C_i))^T J_i + 2\nu^T \left(\sum_{i=1}^n \text{vec}(C_i) - \text{vec}(F^{-1}) \right) \right\}. \tag{36}$$

Then, we define $\beta = (\text{vec}(C_1)^T, \dots, \text{vec}(C_n)^T, \nu^T)^T$. By substituting β into the expression stated in (36), it simplifies to the expression given by

$$\underset{\hat{\beta}}{\text{minimize}} \left\{ \hat{\beta}^T Q \hat{\beta} + K - 2 \hat{\beta}^T J \right\}, \tag{37}$$

where

$$Q = \begin{pmatrix} Q_\bullet & \dots & Q_{1n} & I \\ \vdots & \ddots & \vdots & \vdots \\ Q_{n1} & \dots & Q_\bullet & I \\ I & \dots & I & 0 \end{pmatrix}, \quad J = \begin{pmatrix} J_1 \\ \vdots \\ J_n \\ \text{vec}(F^{-1}) \end{pmatrix}. \tag{38}$$

By considering the formula presented in (37), we obtain that

$$2Q\beta - 2J = 0, \tag{39}$$

which implies that $Q\beta = J$ and then $\hat{\beta} = Q^{-1}J$.

Empirically, we begin by evaluating a linear co-regionalization scheme for the multivariate random field A . This provides an estimate for the matrix outlined in (31), leading directly to estimators for the matrices $\text{Var}(a_i)$ and $\text{Cov}(a_i, a_j)$. After obtaining Q and J as stated in (38), we can use these matrices in conjunction with the formula given in (39) to estimate C_1, \dots, C_n . This, in turn, allows us to determine the functional parameters as described in (28). Alternatively, the cumulative prediction variance, denoted as $\sigma_{s_0}^2$, can be directly estimated using the expression formulated as $\sigma_{s_0}^2 = \int_T \text{Var}(\hat{\zeta}_{s_0}(t) - \zeta_{s_0}(t)) dt$. This estimation is expressed as $\hat{\sigma}_{s_0}^2 = \hat{\beta}^\top Q \hat{\beta} + K - 2\hat{\beta}^\top J$, with the matrix K derived using $\widehat{\text{Var}}(a_\bullet)$, which, in turn, can be sourced from a co-regionalization linear model. The cumulative prediction variance, $\sigma_{s_0}^2$ namely, serves as a holistic measure of the uncertainty associated with approximating the entire curve. Utilizing the estimated parameters and drawing from the equation outlined in (23), we can derive a variance function for individual prediction points.

3. Kriging Prediction for Non-Stationary Functional Random Fields

Non-stationarity in functional data poses unique challenges in geostatistical analysis. To address this type of non-stationarity, various kriging methods have been proposed and refined over the years. In this section, we focus on two core methodologies: universal kriging and residual kriging for functional data [40,41]. By understanding the underpinnings of these approaches, we aim to provide researchers and practitioners with a robust toolkit to address the intricacies posed by non-stationary functional random fields.

3.1. Formulation for Kriging of Geospatial Functional Observations

Let us begin by defining the formulation for kriging of geospatial functional data. Consider the $n \times 1$ functional vector $\zeta_s(t) = (\zeta_{s_1}(t), \dots, \zeta_{s_n}(t))^\top$, in which every component is a data point of a functional stochastic variable $\zeta_s(t)$ within a given region $D \subset \mathbb{R}^d$, where d is typically either 2 or 3. Let us denote $\Sigma = \text{Var}(\zeta_s(t))$, $\sigma_0^2 = \text{Var}(\zeta_{s_0}(t))$, and $c = \text{Cov}(\zeta_s(t), \zeta_{s_0}(t))$, where $c \in \mathbb{R}^n$ and $t \in (0, \infty)$.

Assuming a linear geospatial trend, we can write

$$\zeta_s(t) = \alpha(t) + \beta_1(t)x(s) + \beta_2(t)y(s) + \varepsilon_s(t) = \mathbf{X}(s)\beta(t) + \varepsilon(t),$$

where $\alpha(t)$ is the functional intercept, while $\beta_1(t)$ and $\beta_2(t)$ are functional coefficients corresponding to the spatial coordinates $x(s)$ and $y(s)$, respectively, for every $s \in D$ and $t \in (0, \infty)$. Additionally, $E(\varepsilon_s(t)) = 0$ and $\text{Var}(\varepsilon_s(t)) = \sigma^2(t)$. The functional kriging estimator for $\zeta_{s_0}(t)$ is then given by

$$\hat{\zeta}_{s_0}(t) = \sum_{i=1}^n \eta_i \zeta_{s_i}(t) = \boldsymbol{\eta}^\top \boldsymbol{\zeta}_s(t). \tag{40}$$

The elements of the weight vector $\boldsymbol{\eta}$ are chosen so that the resulting estimator is both unbiased and minimizes its variance. The estimator stated in (40) remains unbiased under the conditions presented as

$$\begin{aligned} E(\hat{\zeta}_{s_0}(t)) &= E(\zeta_{s_0}(t)), \\ \boldsymbol{\eta}^\top \mathbf{X}(s)\beta(t) &= \mathbf{X}^\top(s_0)\beta(t), \\ \boldsymbol{\eta}^\top \mathbf{X}(s) &= \mathbf{X}^\top(s_0). \end{aligned}$$

To determine the values of η_i , for $i \in \{1, \dots, n\}$, we consider the formulation stated as

$$\underset{\eta_1, \dots, \eta_n}{\text{minimize}} \{ \text{MSE}(s_0) \},$$

subject to $\boldsymbol{\eta}^\top \mathbf{X}(s) = \mathbf{X}^\top(s_0)$, where $\text{MSE}(s_0) = \int_T \text{Var}(\widehat{\zeta}_{s_0}(t) - \zeta_{s_0}(t)) dt$ represents the mean squared error (MSE).

In the stationary scenario, the variance is expressed as

$$\begin{aligned} \text{Var}(\widehat{\zeta}_{s_0}(t) - \zeta_{s_0}(t)) &= \boldsymbol{\eta}^\top \text{Var}(\boldsymbol{\zeta}(t)) \boldsymbol{\eta} - 2\boldsymbol{\eta}^\top \text{Cov}(\boldsymbol{\zeta}(t), \zeta_{s_0}(t)) + \sigma_0^2(t) \\ &= \boldsymbol{\eta}^\top \boldsymbol{\Sigma} \boldsymbol{\eta} - 2\boldsymbol{\eta}^\top \mathbf{c}(t) + \sigma_0^2(t). \end{aligned}$$

Consider the formula given by

$$\phi(\boldsymbol{\eta}, \mathbf{v}) = \int_T (\boldsymbol{\eta}^\top \boldsymbol{\Sigma} \boldsymbol{\eta})(t) dt - 2 \int_T \boldsymbol{\eta}^\top \mathbf{c}(t) dt + \int_T \sigma_0^2(t) dt - 2(\boldsymbol{\eta}^\top \mathbf{X}(s) - \mathbf{X}^\top(s_0)) \mathbf{v},$$

where \mathbf{v} denotes the Lagrange multipliers associated with the unbiasedness constraint. The optimal values of ϕ are achieved when the partial derivatives $\partial\phi/\partial\boldsymbol{\eta}$ and $\partial\phi/\partial\mathbf{v}$ vanish.

Taking derivatives with respect to $\boldsymbol{\eta}$, we obtain

$$\begin{aligned} \frac{\partial\phi}{\partial\boldsymbol{\eta}} &= -2 \int_T \mathbf{c}(t) dt + 2 \int_T \boldsymbol{\Sigma} \boldsymbol{\eta} dt - 2\mathbf{X}(s) \mathbf{v} \\ &= -2 \int_T \mathbf{c}(t) dt + 2 \left(\int_T \boldsymbol{\Sigma} dt \right) \boldsymbol{\eta} - 2\mathbf{X}(s) \mathbf{v}. \end{aligned}$$

Equating to zero the above expression, we have

$$\left(\int_T \boldsymbol{\Sigma} dt \right) \boldsymbol{\eta} + \mathbf{X}(s) \mathbf{v} = \int_T \mathbf{c}(t) dt. \tag{41}$$

Analogously, taking partial derivatives with respect to \mathbf{v} and equating them to zero, we reach the unbiasedness established as

$$\mathbf{X}^\top(s) \boldsymbol{\eta} + \mathbf{0} \mathbf{v} = \mathbf{X}(s_0). \tag{42}$$

From the formulations presented in (41) and (42), we reach

$$\begin{bmatrix} \int_T \boldsymbol{\Sigma} dt & \mathbf{X}(s) \\ \mathbf{X}^\top(s) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \int_T \mathbf{c}(t) dt \\ \mathbf{X}(s_0) \end{bmatrix}.$$

Given that the matrix presented on the left-hand side of the expression is positive definite, we can conclude that

$$\begin{bmatrix} \boldsymbol{\eta} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \int_T \boldsymbol{\Sigma} dt & \mathbf{X}(s) \\ \mathbf{X}^\top(s) & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \int_T \mathbf{c}(t) dt \\ \mathbf{X}(s_0) \end{bmatrix}. \tag{43}$$

From the formulation stated in (43), we attain at

$$\boldsymbol{\eta} = \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \left(\int_T \mathbf{c}(t) dt + \mathbf{X}(s) \mathbf{v} \right).$$

Then, we get

$$\begin{aligned} \mathbf{X}^\top(s_0) &= \left(\left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \left(\int_T \mathbf{c}(t) dt + \mathbf{X}(s) \mathbf{v} \right) \right)^\top \mathbf{X}(s), \\ \mathbf{X}^\top(s_0) &= \int_T \mathbf{c}^\top(t) dt \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \mathbf{X}(s) + \mathbf{v}^\top \mathbf{X}^\top(s) \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \mathbf{X}(s), \\ \mathbf{v}^\top &= \left(\mathbf{X}^\top(s_0) - \int_T \mathbf{c}^\top(t) dt \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \mathbf{X}(s) \right) \left(\mathbf{X}^\top(s) \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \mathbf{X}(s) \right)^{-1}, \\ \mathbf{v} &= \mathbf{W}^{-1}(\mathbf{X}(s_0) - \mathbf{Y}), \mathbf{W} = \mathbf{X}^\top(s) \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \mathbf{X}(s), \\ \mathbf{Y} &= \mathbf{X}^\top(s) \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \int_T \mathbf{c}(t) dt, \\ \boldsymbol{\eta} &= \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \left(\int_T \mathbf{c}(t) dt \right) + \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \mathbf{X}(s) \mathbf{W}^{-1}(\mathbf{X}(s_0) - \mathbf{Y}). \end{aligned}$$

Thus, the predictor given in (40) and the MSE are obtained as

$$\begin{aligned} \widehat{\zeta}_{s_0}(t) &= \left(\left(\int_T \mathbf{c}^\top(t) dt \right) \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} + (\mathbf{X}(s_0) - \mathbf{Y})^\top \mathbf{W}^{-1} \mathbf{X}^\top(s) \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \right) \boldsymbol{\zeta}(t), \\ \text{MSE}(s_0) &= \int \left(\boldsymbol{\eta}^\top \boldsymbol{\Sigma} \boldsymbol{\eta} \right) (t) dt - 2 \boldsymbol{\eta}^\top \int_T \mathbf{c}(t) dt + \int_T \sigma_0^2(t) dt. \end{aligned}$$

The value of the prediction variance is consequently reached as

$$\sigma_{s_0}^2 = \int_T \sigma_0^2(t) dt - \left(\int_T \mathbf{c}(t) dt \right)^\top \left(\int_T \boldsymbol{\Sigma} dt \right)^{-1} \left(\int_T \mathbf{c}(t) dt \right) + (\mathbf{X}(s_0) - \mathbf{Y})^\top \mathbf{W}^{-1} (\mathbf{X}(s_0) - \mathbf{Y}).$$

To solve the matrix equation system presented in (43), it is necessary to estimate the spatial autocorrelation matrix $\boldsymbol{\Sigma} = (\text{Cov}_{s_i s_j})$ with $\text{Cov}_{s_i s_j}(t) = \text{Cov}(\varepsilon_{s_i}(t), \varepsilon_{s_j}(t))$.

Utilizing the trace-variogram function, we get

$$\varrho_{s_i s_j}(t) = \varrho(\zeta_{s_i}(t), \zeta_{s_j}(t)) = \frac{1}{2} \text{Var}(\zeta_{s_i}(t) - \zeta_{s_j}(t)) = \sigma^2(t) - \text{Cov}_{s_i s_j}(t).$$

Incorporating this most recent relation from (41), we derive

$$\left(\int_T \boldsymbol{\varrho} dt \right) \boldsymbol{\eta} - \mathbf{X}(s) \mathbf{v} = \int_T \boldsymbol{\varrho}_{s_0} dt,$$

where $\boldsymbol{\varrho}_{s_0} = (\varrho_{s_i s_0})$. Therefore, the matrix system defined in (43) may be formulated as

$$\begin{bmatrix} \int_T \boldsymbol{\varrho} dt & \mathbf{X}(s) \\ \mathbf{X}^\top(s) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta} \\ -\mathbf{v} \end{bmatrix} = \begin{bmatrix} \int_T \boldsymbol{\varrho}_{s_0} dt \\ \mathbf{X}(s_0) \end{bmatrix}.$$

Consequently, solving this system yields the weights $\boldsymbol{\eta}$ and the Lagrange multipliers \mathbf{v} .

3.2. Residual Kriging and External Drift for Spatially Correlated Functional Data

Next, we present a review on residual kriging for functional data [40]. Assume that the residuals of a linear spatial trend are given by

$$e_{s_i}(t) = \zeta_{s_i}(t) - \widehat{\zeta}_{s_i}(t) = \zeta_{s_i}(t) - \left(\widehat{\alpha}(t) + \widehat{\beta}_1(t) x_i + \widehat{\beta}_2(t) y_i \right), \quad i \in \{1, \dots, n\}, \quad (44)$$

where (x_i, y_i) are the spatial coordinates.

While multiple regression can serve as an alternative, various models are available for detrending the mean of functional process.

Detrending plays a pivotal role in spatial analysis, particularly in identifying structures in the residuals. However, detrending may introduce bias in the experimental covariances or variograms of residuals [3]. It is essential for readers to tackle this step carefully, ensuring the advantages of detrending surpass potential drawbacks. Though this formulation highlights the 2D spatial coordinates x_i and y_i , it is mainly for illustrative purposes. The methodology is adaptable beyond 2D and can extend to higher-dimensional spaces, requiring additional spatial coordinates and associated drift terms. Building on the residuals outlined in (44), the estimation of a residual curve at unsampled locations can be achieved utilizing our discussion of Section 2. Specifically, a residual function can be predicted using

$$\widehat{e}_{s_0}(t) = \sum_{i=1}^n \eta_i e_{s_i}(t), \tag{45}$$

$$\widehat{e}_{s_0}(t) = \sum_{i=1}^n \eta_i(t) e_{s_i}(t), \tag{46}$$

$$\widehat{e}_{s_0}(t) = \sum_{i=1}^n \int_T \eta_i(t, v) e_{s_i}(t) dt. \tag{47}$$

When predicting the functional variable at a non-sampled site s_0 , we get

$$\widehat{\varsigma}_{s_0}(t) = \widehat{\alpha}(t) + \widehat{\beta}_1(t)x_0 + \widehat{\beta}_2(t)y_0 + \widehat{e}_{s_0}(t),$$

where $\widehat{e}_{s_0}(t)$ is defined in (45), (46), or (47).

Another approach involves using explanatory variables distinct from the spatial coordinates to estimate the spatial trend. In that case, consider the model formulated as

$$\varsigma_{s_i}(t) = \alpha(t) + \sum_{q=1}^Q \varrho_q(t)c_{q,i} + \sum_{r=1}^R \beta_r(t)x_{r,i}(t) + \varepsilon_{s_i}(t), \quad i \in \{1, \dots, n\}. \tag{48}$$

In this model, $\alpha(t)$ acts as a functional intercept, $c_{p,i}$ is the value of a scalar explanatory variable, and $x_{q,i}(t)$ is the value of an explanatory variable at functional location s_i . The terms $\varrho_p(t)$ and $\beta_q(t)$ represent their respective regression parameters. The drift, as detailed in (48), incorporates both scalar and functional explanatory variables. Moreover, the parameters $\varrho_p(t)$ and $\beta_q(t)$ are in functional form, allowing estimation of the nonlinear effects in the explanatory variables. Thus, the resulting residuals can be expressed as

$$e_{s_i}(t) = \varsigma_{s_i}(t) - \widehat{\varsigma}_{s_i}(t) = \varsigma_{s_i}(t) - \left(\widehat{\alpha}(t) + \sum_{q=1}^Q \widehat{\varrho}_q(t)c_{q,i} + \sum_{r=1}^R \widehat{\beta}_r(t)x_{r,i}(t) \right), \quad i \in \{1, \dots, n\}.$$

Using the relations described in (45), (46), or (47), a prediction for the error $\widehat{e}_{s_0}(t)$ at a non-sampled location s_0 can be derived. Then, the prediction of the functional variable is given as

$$\widehat{\varsigma}_{s_0}(t) = \widehat{\alpha}(t) + \sum_{q=1}^Q \widehat{\varrho}_q(t)c_{q,0} + \sum_{r=1}^R \widehat{\beta}_r(t)x_{r,0}(t) + \widehat{e}_{s_0}(t).$$

3.3. R Software and Packages of Spatial Statistics

In the geostatistical framework, particularly when addressing kriging and cokriging for stationary and non-stationary functional random fields, the computational tools we employ are of paramount importance. The R software, an unparalleled open-source environment for statistical computing, stands as a cornerstone in this domain, primarily due to its rich and ever-growing ecosystem of packages dedicated to spatial statistics [42].

Among the myriad of packages available, `sp` serves as a linchpin, offering classes and methods dedicated to spatial data. This package essentially acts as the bedrock upon which many subsequent spatial packages have been developed [43]. The `gstat` package enriches this foundation, extending capabilities specifically for geostatistical analysis. Its offerings, especially concerning variogram estimation and various kriging methods, have been elucidated earlier in this section and in Section 2. These methods are deeply rooted in seminal works, particularly those presented in [23,27,28,39], among other pioneers in the field [40,41]. Importantly, `gstat` acts as a bridge, allowing the seamless translation of theoretical underpinnings into tangible computational solutions [44].

For those at the intersection of FDA and spatial statistics, the `fda` package is highly relevant. It provides tools for FDA, complementing our discussion on functionally-coherent spatial data and continuous time-varying kriging [45]. Beyond these foundational tools, `spdep` addresses the need for complex spatial regression, encapsulating both linear and nonlinear models. With capabilities such as defining spatial weights and analyzing spatial autocorrelation, it proves indispensable, especially when navigating the intricacies of multifaceted spatial structures [46]. The recent expansion in R geostatistical tools further underscores its stature. For instance, the `EpiGeostats` package enhances visualization capacities, making interpretation of geostatistical disease risk maps intuitive [47].

Additionally, the `GeoSim` package fills a niche by focusing on pluri-Gaussian simulation, bridging gaps between categorical and continuous variables [48]. For researchers in soil science, the insights from [49] provide invaluable guidance on merging linear mixed models with geostatistics [49]. The `covatest` package simplifies the selection of space-time covariance functions, a task pivotal to many geostatistical analyses [50]. And for those working with compositional data, the work presented in [51] serves as a comprehensive guide, detailing geostatistical methods in R language.

Confronted with the complexities that riddle our study, especially in deciphering kriging predictions for stationary and non-stationary functional random fields, a multifaceted approach becomes imperative. This entails harnessing methods from the aforementioned packages, often in synergy, and crafting custom scripts that address distinct challenges. This adaptability underscores the R environment's unmatched versatility in tackling advanced spatial statistical problems. For those poised on the brink of deeper exploration, package vignettes remain an invaluable primer. Additionally, the breadth and depth of tutorials and applications available serve as a testament to the boundless possibilities that R offers for spatial data analysis [52]. Another noteworthy contribution to the R spatial statistics toolkit is a comparative case study on R packages designed for analyzing areal data.

This study offers insights into the optimal use-cases and advantages of various spatial packages, assisting practitioners in making informed decisions [53]. In the ever-evolving domain of mixed-effects models with spatial considerations, the `sdmTMB` package stands out. This package provides swift and versatile solutions, particularly suited for integrated linear models with random effects that address both geospatial and space-time variations. Such capabilities underscore its importance for detailed spatial analyses [54].

4. Discussion and Conclusions

The analysis of spatially correlated functional data has gained prominence in contemporary data analysis, notably in the big data era. Functional geostatistics, as a result, serves as a linchpin in deepening our grasp and modeling of intricate spatial processes. A pivotal factor in this analysis is the assessment of stationarity. This property, which underpins the stability of statistical characteristics over the spatial domain, guides the choice of the suitable kriging method for functional data. Table 1 summarizes the main results and contributions of our work.

Table 1. Summary of main results and contributions.

Aspect/Concept	Description/Details	Significance	Challenges Addressed
Stationarity evaluation	It assesses stability across the spatial domain	It is essential for the right kriging method selection	Accurate assessment of functional processes’ stationarity
Spatial prediction methodologies	It gives a comprehensive review of methods under both stationary and non-stationary conditions	It provides integrated view tailored to stationarity conditions	Challenges posed by different stationarity conditions
Applications	It ranges from environmental monitoring to biomedical research	It facilitates informed decisions in various domains	Capturing accurate spatial dependence and functional variability
Practical implications	It facilitates informed decisions in environmental management and resource allocation	It enhances practical applications	Promoting the understanding of space-time-functional variability interplay
Further advancements	It stimulates innovative techniques and approaches	It broadens the scope of functional geostatistics	Complex spatial patterns and functional variations
Contribution to spatial statistics	It gives broader implications beyond functional geostatistics	It filters understanding of spatial statistics	Interplay between space, time, and functionality

Delving into this topic, the present article offered a detailed review of methodologies tailored for the spatial prediction of functional geostatistical data, encompassing both stationarity and non-stationarity scenarios. Through an exploration of various scholarly proposals, this study shed light on the prevailing challenges and breakthroughs in this realm. The significance of our findings spans both theoretical and applied dimensions.

The real-world implications of our exploration are expansive, influencing domains like environmental monitoring, geosciences, and even biomedical research. By fostering an accurate modeling paradigm for spatial dependence and functional variability, the groundwork is laid for experts to make insightful decisions in these arenas, particularly in spheres like environmental stewardship, resource allocation, and risk appraisal. Moreover, this article acts as a catalyst for innovative strides in functional geostatistics, spotlighting the hurdles brought about stationarity contexts. This beckons the genesis of avant-garde techniques to analyze and predict spatially intertwined functional data—a contribution benefiting not just functional geostatistics but also enriching spatial statistics at large.

In summation, the contributions herein fortify the foundational aspects of analyzing, modeling, and predicting spatially dependent functional data. By weaving in considerations of diverse stationarity contexts, we amplify our prowess in deciphering intricate spatial patterns and functional idiosyncrasies typical of real-world datasets. Harnessing these methodologies promises heightened prediction accuracy, fostering judicious decision-making. This not only augments realms like environmental governance and resource distribution but also heralds a new dawn of research and novelty in the domain, broadening the horizons of spatial statistics and its multifaceted applications.

In the setting of spatial data exhibiting non-symmetric distributions, several scholarly investigations have been undertaken [55–58], as well as in quantile regression with spatial data [59], and diagnostic analysis in regression models [21,60]. However, much investigation on this field must be explored still, even in traditional spatial models [61].

Author Contributions: Conceptualization, R.G. and V.L.; methodology, R.G., V.L. and C.C.; formal analysis, R.G., V.L. and C.C.; investigation, R.G.; writing—original draft preparation, R.G.; writing—review and editing, V.L. and C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by FONDECYT, grant number 1200525 (V.L.), from the National Agency for Research and Development (ANID) of the Chilean government under the Ministry of Science, Technology, Knowledge, and Innovation; and by Portuguese funds through the CMAT—Research Centre of Mathematics of University of Minho—within projects UIDB/00013/2020 and UIDP/00013/2020 (C.C.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would also like to thank the Editors and Reviewers for their constructive comments which led to improve the presentation of the manuscript.

Conflicts of Interest: There are no conflict of interest declared by the authors.

References

- Ramsay, J.; Silverman, B. *Functional Data Analysis*; Springer: New York, NY, USA, 2005.
- Christakos, G. *Modern Spatiotemporal Geostatistics*; Oxford University Press: Oxford, UK, 2000.
- Chilès, J.P.; Delfiner, P. *Geostatistics: Modeling Spatial Uncertainty*; Wiley: Hoboken, NJ, USA, 2009.
- Ripley, B.D. *Spatial Statistics*; Wiley: New York, NY, USA, 2005.
- Cressie, N. *Statistics for Spatial Data*; Wiley: New York, NY, USA, 2015.
- Goulard, M.; Voltz, M. Geostatistical interpolation of curves: A case study in soil science. In *Geostatistics Tróia'92*; Springer: Dordrecht, The Netherlands, 1993; Volume 1, pp. 805–816.
- Ignaccolo, R.; Mateu, J.; Giraldo, R. Kriging with external drift for functional data for air quality monitoring. *Stoch. Environ. Res. Risk Assess.* **2014**, *28*, 1171–1186. [[CrossRef](#)]
- Menafoglio, A.; Secchi, P.; Dalla Rosa, M. A universal kriging predictor for spatially dependent functional data of a Hilbert space. *Electron. J. Stat.* **2013**, *7*, 2209–2240. [[CrossRef](#)]
- Menafoglio, A.; Grujic, O.; Caers, J. Universal kriging of functional data: Trace-variography vs cross-variography? Application to gas forecasting in unconventional shales. *Spat. Stat.* **2016**, *15*, 39–55. [[CrossRef](#)]
- Mateu, J.; Giraldo, R. (Eds.) *Geostatistical Functional Data Analysis*; Wiley: Hoboken, NJ, USA, 2022.
- Christakos, G. *Modern Spatiotemporal Geostatistics*; Oxford University Press: New York, NY, USA, 2000.
- Christakos, G.; Olea, R.A.; Serre, M.L.; Yu, H.-L.; Wang, L.-L. *Interdisciplinary Public Health Reasoning and Epidemic Modelling: The Case of Black Death*; Springer: New York, NY, USA, 2002.
- Christakos, G.; Hristopulos, D.T.; Bogaert, P. Spatiotemporal Environmental Health Modelling: A Tractatus Stochasticus. *Stoch. Environ. Res. Risk Assess.* **2000**, *14*, 245–262.
- Christakos, G. *Physical Geography, Geosystems and Spatiotemporal Geostatistics*; Elsevier: Amsterdam, The Netherlands, 2012.
- Christakos, G. *Environmental Health Modelling: An Introductory Manual*; Springer: Cham, Switzerland, 2013.
- Christakos, G. Analysis of Environmental Data using Neural Networks. *Environ. Sci. Technol.* **1970**, *4*, 110–117.
- Christakos, G.; Serre, M.L.; Demyanov, V. Neural Network Residual Analysis of the Spatial Estimation of Radioactivity. *Stoch. Environ. Res. Risk Assess.* **2001**, *15*, 209–231.
- Wu, J.; He, X.; Christakos, G. Quantitative Integration of Spatiotemporal Data Sources in Modern Geosciences. *Adv. Geosci.* **2021**, *56*, 183–201.
- Strandberg, J.; de Luna, S.; Mateu, J. *A Comparison of Spatiotemporal and Functional Kriging Approaches*; Springer: New York, NY, USA, 2022.
- Cortes, D.L.; Camacho-Tamayo, J.H.; Giraldo, R. Spatial prediction of soil infiltration using functional geostatistics. *AUC Geographica* **2018**, *53*, 149–155. [[CrossRef](#)]
- Garcia-Papani, F.; Uribe-Opazo, M.A.; Leiva, V.; Aykroyd, R.G. Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. *Stoch. Environ. Res. Risk Assess.* **2017**, *31*, 105–124. [[CrossRef](#)]
- Dabo-Niang, S.; Ferraty, F.; Monestiez, P.; Nerini, D. A cokriging method for spatial functional data with applications in oceanology. In *Functional and Operatorial Statistics*; Physica-Verlag: Heidelberg, Germany, 2008; pp. 237–242.
- Giraldo, R.; Herrera, L.; Leiva, V. Cokriging prediction using as a secondary variable a functional random field with application in environmental pollution. *Mathematics* **2020**, *8*, 1305. [[CrossRef](#)]
- Azevedo, L. Model reduction in geostatistical seismic inversion with functional data analysis in seismic inversion. *Geophysics* **2022**, *87*, M1–M11. [[CrossRef](#)]
- Dabo-Niang, S.; Ternynck, C.; Thiam, B.; Yao, A.F. Nonparametric statistical analysis of spatially distributed functional data. In *Geostatistical Functional Data Analysis*; Mateu, J., Giraldo, R., Eds.; Springer: New York, NY, USA, 2022; pp. 175–210.
- Montero, J.M.; Fernández-Avilés, G. Functional kriging prediction of pollution series: The geostatistical alternative for spatially-fixed data. *Estud. Econ. Apl.* **2015**, *33*, 145–174.

27. Giraldo, R.; Delicado, P.; Mateu, J. Ordinary kriging for function-valued spatial data. *Environ. Ecol. Stat.* **2011**, *18*, 411–426. [[CrossRef](#)]
28. Giraldo, R.; Delicado, P.; Mateu, J. Continuous time-varying kriging for spatial prediction of functional data: An environmental application. *J. Agric. Biol. Environ. Stat.* **2010**, *15*, 66–82. [[CrossRef](#)]
29. Fernández-Pascual, R.M.; Espejo, R.; Ruiz-Medina, M.D. Moment and Bayesian wavelet regression from spatially correlated functional data. *Stoch. Environ. Res. Risk Assess.* **2016**, *30*, 523–557. [[CrossRef](#)]
30. Bouzebda, S.; Soukariah, I. Non-parametric conditional U-processes for locally stationary functional random fields under stochastic sampling design. *Mathematics* **2023**, *11*, 16. [[CrossRef](#)]
31. Adler, R.J.; Taylor, J.E. *Random Fields and Geometry*; Springer: New York, NY, USA, 2007.
32. García-Pérez, A. On robustness for spatio-temporal data. *Mathematics* **2023**, *10*, 1785. [[CrossRef](#)]
33. Griffith, D.A. Spatial autocorrelation and unorthodox random variables: The uniform distribution. *Chil. J. Stat.* **2022**, *13*, 133–153. [[CrossRef](#)]
34. Jelsema, C.M.; Paul, R.; McKean, J.W. Robust estimation of reduced rank models to large spatial datasets. *REVSTAT Stat. J.* **2020**, *18*, 203–221.
35. McCord, M.; Lo, D.; Davis, P.; McCord, J.; Hermans, L.; Bidanset, P. Applying the geostatistical eigenvector spatial filter approach into regularized regression for improving prediction accuracy for mass appraisal. *Appl. Sci.* **2022**, *12*, 10660. [[CrossRef](#)]
36. Alvo, M.; Mu, J. COVID-19 data analysis using Bayesian models and nonparametric geostatistical models. *Mathematics* **2023**, *11*, 1359. [[CrossRef](#)]
37. Wackernagel, H. Cokriging versus kriging in regionalized multivariate data analysis. *Geoderma* **1994**, *62*, 83–92. [[CrossRef](#)]
38. Nerini, D.; Monestiez, P.; Manté, C. Cokriging for spatial functional data. *J. Multivar. Anal.* **2010**, *101*, 409–418. [[CrossRef](#)]
39. Giraldo, R.; Delicado, P.; Mateu, J. Spatial prediction of a scalar variable based on data of a functional random field. *Commun. Estadística* **2017**, *10*, 315–344. [[CrossRef](#)]
40. Franco-Viloria, M.; Ignacollo, R. *Universal, Residual and External Drift Functional Kriging*. In *Geostatistical Functional Data Analysis*; Mateu, J., Giraldo, R., Eds.; Springer: New York, NY, USA, 2015; pp. 55–72.
41. Caballero, W.; Giraldo, R.; Mateu, J. A universal kriging approach for spatial functional data. *Stoch. Environ. Res. Risk Assess.* **2013**, *27*, 1553–1563. [[CrossRef](#)]
42. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021. Available online: <https://www.R-project.org/> (accessed on 2 August 2023).
43. Pebesma, E.J.; Bivand, R.S. Classes and methods for spatial data in R. *R News* **2005**, *5*, 9–13.
44. Pebesma, E.J. Multivariable geostatistics in S: The gstat R package. *Comput. Geosci.* **2004**, *30*, 683–691. [[CrossRef](#)]
45. Ramsay, J.O.; Hooker, G.; Graves, S. *Functional Data Analysis with R and MATLAB*; Springer: New York, NY, USA, 2009.
46. Bivand, R.S.; Pebesma, E.; Gómez-Rubio, V. *Applied Spatial Data Analysis with R*; Springer: New York, NY, USA, 2013.
47. Ribeiro, M.; Azevedo, L.; Pereira, M.J. EpiGeostats: An R Package to Facilitate Visualization of Geostatistical Disease Risk Maps. *Math. Geosci.* **2023**, 1–17. [[CrossRef](#)]
48. Valakas, G.; Modis, K. GeoSim: An R-package for pluri-Gaussian simulation and co-simulation between categorical and continuous variables. *Appl. Comput. Geosci.* **2023**, *19*, 100130. [[CrossRef](#)]
49. Slaets, J.I.; Boeddinghaus, R.S.; Piepho, H.P. Linear mixed models and geostatistics for designed experiments in soil science: Two entirely different methods or two sides of the same coin? *Eur. J. Soil Sci.* **2021**, *72*, 47–68. [[CrossRef](#)]
50. Cappello, C.; De Iaco, S.; Posa, D. covatest: An R package for selecting a class of space-time covariance functions. *J. Stat. Softw.* **2020**, *94*, 1–42. [[CrossRef](#)]
51. Tolosana-Delgado, R.; Mueller, U. *Geostatistics for compositional data with R*; Springer: New York, NY, USA, 2021.
52. Bivand, R.S.; Pebesma, E.; Gómez-Rubio, V. *Applied Spatial Data Analysis with R*; Springer: New York, NY, USA, 2008.
53. Bivand, R. R packages for analyzing spatial data: A comparative case study with areal data. *Geogr. Anal.* **2022**, *54*, 488–518. [[CrossRef](#)]
54. Anderson, S.C.; Ward, E.J.; English, P.A.; Barnett, L.A. sdmTMB: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. *bioRxiv* **2022**. [[CrossRef](#)]
55. Martínez, S.; Giraldo, R.; Leiva, V. Birnbaum-Saunders functional regression models for spatial data. *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 1765–1780. [[CrossRef](#)]
56. Garcia-Papani, F.; Leiva, V.; Ruggeri, F.; Uribe-Opazo, M.A. Kriging with external drift in a Birnbaum-Saunders geostatistical model. *Stoch. Environ. Res. Risk Assess.* **2018**, *32*, 1517–1530. [[CrossRef](#)]
57. Sanchez, L.; Leiva, V.; Galea, M.; Saulo, H. Birnbaum-Saunders quantile regression models with application to spatial data. *Mathematics* **2020**, *8*, 1000. [[CrossRef](#)]
58. Kotz, S.; Leiva, V.; Sanhueza, A. Two new mixture models related to the inverse Gaussian distribution. *Methodol. Comput. Appl. Probab.* **2010**, *12*, 199–212. [[CrossRef](#)]
59. Leiva, V.; Sanchez, L.; Galea, M.; Saulo, H. Global and local diagnostic analytics for a geostatistical model based on a new approach to quantile regression. *Stoch. Environ. Res. Risk Assess.* **2020**, *34*, 1457–1471. [[CrossRef](#)]

60. Leiva, V.; Rojas, E.; Galea, M.; Sanhueza, A. Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data. *Appl. Stoch. Models Bus. Ind.* **2014**, *30*, 114–128. [[CrossRef](#)]
61. Giraldo, R.; Leiva, V.; Christakos, G. Leverage and Cook distance in regression with geostatistical data: Methodology, simulation, and applications related to geographical information. *Int. J. Geogr. Inf. Sci.* **2021**, *37*, 607–633. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.