POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Fully Non-stationary Spatial Functional Random Fields

Laurea Magistrale in Mathematical Engineering - Ingegneria Matematica

Author: Giacomo De Carlo

Advisor: Prof. Alessandra Menafoglio

Co-advisor: Dr. Riccardo Scimone

Academic year: 2021-2022

---

1. Introduction and Preliminaries

The thesis faces the problem of characterizing and analyzing non-stationary functional spatial random fields. By functional random field we mean a set of random variables

$$\{\mathcal{X}_{\mathbf{s}}, \in D \subseteq \mathbb{R}^d\} \tag{1}$$

taking values in a separable Hilbert space $\mathcal{H}$, indexed by a spatial location $\mathbf{s}$ varying in a continuous domain $D \subseteq \mathbb{R}^d$ (in our case $d = 2$). Such processes can be employed to model complex spatially distributed data objects, for instance curves, surfaces or images. Many spatial prediction methods have been developed so far to deal with complex spatial data (Menafoglio and Secchi [2017] for a structured review). However, most of them often lay their theoretical foundations upon global covariance models to capture variability and spatial dependence. The most common assumption regards the stationarity of the process generating the data: in this simplified setting, the covariance structure of the process only depends on the vector difference between locations, and its mean is spatially constant. While this assumption enormously simplifies the problems of parameter estimation and spatial prediction, it limits modeling capabilities and flexibility. Non-stationary covariance functions, on the other hand, allow to model processes whose variability properties change with location: this is critical, for instance, in applications as atmospherical and climatological studies.

1.1. Geostatistics for Hilbert data

In this work, we assume that process (1) is non-stationary, and that we can represent its elements $\mathcal{X}_s$, at a generic location $\boldsymbol{s} \in D$, as the sum of its mean $m_{\mathbf{s}} : D \to \mathcal{H}$, called drift, and a stochastic zero-mean residual $\delta_{\boldsymbol{s}}$, i.e.,

$$\mathcal{X}_{\boldsymbol{s}} = m_{\boldsymbol{s}} + \delta_{\boldsymbol{s}}.$$

The complete autocorrelation structure of the stochastic residual of process (1) is described by the family of cross-covariance operators, defined as

$$C_{\mathbf{s},\boldsymbol{\tau}} = \{C_{\mathbf{s},\boldsymbol{\tau}} : \mathcal{H} \to \mathcal{H}, \text{ s.t. } C_{\mathbf{s},\boldsymbol{\tau}} x =$$

$$\mathbb{E}\left[\langle \mathcal{X}_{\mathbf{s}} - m_{\mathbf{s}}, x\rangle (\mathcal{X}_{\boldsymbol{\tau}} - m_{\boldsymbol{\tau}})\right], \mathbf{s}, \boldsymbol{\tau} \in D\}.$$

By introducing an orthonormal basis $\{e_k\}_{k\in\mathbb{N}}$ for $\mathcal{H}$, we can obtain a discretized $K \times K$ matrix version of the family of cross-covariance operators as

$$C^K_{\mathbf{s},\boldsymbol{\tau}} = (\langle C_{\mathbf{s},\boldsymbol{\tau}} e_i, e_j\rangle)_{i,j=1,\dots,K},$$

by only considering the first K components of such basis. This discretized version of the cross-covariance operator corresponds to, and needs to fulfill the same properties of, the cross-covariance matrix function for multivariate data. An alternative is the global covariance viewpoint, corresponding to the classical scalar covariogram approach: it can be defined through the trace-covariogram function $C : D \times D \to \mathbb{R}$, such that

$$C(\mathbf{s}, \boldsymbol{\tau}) = Cov(\mathcal{X}_{\mathbf{s}}, \mathcal{X}_{\boldsymbol{\tau}}) = \mathbb{E}[\langle \mathcal{X}_{\mathbf{s}} - m_{\mathbf{s}}, \mathcal{X}_{\boldsymbol{\tau}} - m_{\boldsymbol{\tau}} \rangle].$$

While the complete structure is used for (discretized) operatorial kriging, the global viewpoint is sufficient for kriging with scalar weights (Menafoglio and Secchi [2017]). There exists an evident correspondence of the trace-covariogram function and the cross-covariance operator with, respectively, the covariogram for scalar data and the cross-covariance matrix function for multivariate data; this analogy makes it possible to tackle the problem of modeling and analysing non-stationary functional random fields by extending models and methods already present in the literature for simpler data, briefly reviewed in next Section.

1.2. Non-stationary scalar and multivariate fields: state-of-the-art

Several approaches already exist to handle non-stationary data in scalar and multivariate fields. In particular, Paciorek and Schervish [2006] introduced a very well-known family of univariate non-stationary Matérn correlation functions that allow for spatially varying anisotropy.

Theorem 1.1. *(Paciorek and Schervish [2006]) If an isotropic stationary correlation function $R_S(\cdot)$ is positive definite in $\mathbb{R}^d$ for every $d = 1, 2, ...,$ then the function $R_{NS} : D \times D \to [-1, +1]$ defined by*

$$R_{NS}(\mathbf{s}_i, \mathbf{s}_j) = \frac{2^{\frac{d}{2}} |\Sigma_{\mathbf{s}_i}|^{\frac{1}{4}} |\Sigma_{\mathbf{s}_j}|^{\frac{1}{4}}}{|\Sigma_{\mathbf{s}_i} + \Sigma_{\mathbf{s}_j}|^{\frac{1}{2}}} R_S\left(\sqrt{Q_{ij}}\right) \quad (2)$$

is positive definite in $\mathbb{R}^d$ for $d = 1, 2, ...,$ and is a non-stationary correlation function.

For any $\mathbf{s} \in D$, $\Sigma_{\mathbf{s}}$ Is a real-valued symmetric positive definite d-dimensional square matrix that can represent local geometric anisotropy. Indeed, for each $\mathbf{s}_i, \mathbf{s}_j \in D$

$$Q_{ij} = \mathbf{h}^T \left(\frac{\Sigma_{\mathbf{s}_i} + \Sigma_{\mathbf{s}_j}}{2}\right)^{-1} \mathbf{h}, \ \mathbf{h} = \mathbf{s}_i - \mathbf{s}_j, \quad (3)$$

is the quadratic form that we use to take into account geometric anisotropy by rotating and stretching the plane according to the eigenvalues and eigenvectors of the anisotropy matrices. Thus, the non-stationary covariogram function C for scalar data can effortlessly be obtained from (2), for each $\mathbf{s}_i, \mathbf{s}_j \in D$, as:

$$C(\mathbf{s}_i, \mathbf{s}_j) = \sigma(\mathbf{s}_i)\sigma(\mathbf{s}_j) R_{NS}(\mathbf{s}_i, \mathbf{s}_j), \quad (4)$$

where $\sigma : D \to \mathbb{R}^+$ indicates the standard deviation of the process in $\mathbf{s}$. Fouedjio et al. [2016] developed a computational method that allows one to produce local estimates of the parameters in (4) and eventually smooth them out to obtain a spatial map of mean, anisotropy and variance. Fouedjio et al. [2018] later extended the covariance model of Paciorek and Schervish [2006] and the estimation method of Fouedjio et al. [2016] for multivariate data by considering a linear model of coregionalization where coregionalization matrices are spatially varying. If $1 \leq r \leq p$ is an integer and $u = 1, ..., r$, R^u_{NS} are valid one-dimensional non-stationary correlation functions as defined in (2) and $A^u(\mathbf{s})$ are $p \times p$ spatially varying coregionalization matrices, the cross-covariance matrix function takes the following form

$$\mathbf{C}_{\mathbf{s},\boldsymbol{\tau}} = \sum_{u=1}^{r} A^u(\mathbf{s}) A^u(\boldsymbol{\tau})^T R^u_{NS}(\mathbf{s}, \boldsymbol{\tau}), \ \forall \mathbf{s}, \boldsymbol{\tau} \in D$$

and $i, j = 1, ..., p$.

Our aim is to develop a consistent theoretical framework for modeling non-stationary covariance structures for Hilbert data as an infinite-dimensional analogue of the work in Paciorek and Schervish [2006] and Fouedjio et al. [2018], respectively, for the trace-covariogram and the family of cross-covariance operators. We propose a fully non-stationary functional linear model of coregionalization (FNF-LMC) and estimate its spatially varying parameters by extending the methods of Fouedjio et al. [2016, 2018] to the functional context.

2. Fully Non-Stationary Functional Linear Model of Coregionalization (FNF-LMC)

We here propose an original extension of the LMC to the infinite-dimensional case, which allows to retrieve a closed form for any of the variability measures of the modeled process and to use estimation methods analogue to those in Fouedjio et al. [2016, 2018] due to the similarity of the novel model to its finite dimensional counterparts.

Let $\{e_k\}_{k\in\mathbb{N}}$ be an orthonormal basis for the separable Hilbert space $\mathcal{H}$. We propose the following fully non-stationary functional linear model of coregionalization (FNF-LMC):

$$\mathcal{X}_{\boldsymbol{s}} = m_{\boldsymbol{s}} + \sum_{u=1}^{r} A^u(\boldsymbol{s}) w^u_{\boldsymbol{s}}, \tag{5}$$

where each $A^u : D \to \mathcal{B}(\mathcal{H})$ is a spatially varying bounded linear operator in D, and each $w^u_{\boldsymbol{s}}$ is a zero-mean and square-integrable $\mathcal{H}$-valued spatial random process, whose covariance model is further specified below. The idea of this model is to start from an infinite sequence of scalar non-stationary processes as defined in (2), each one with the same distribution, and to construct a functional random field by treating each scalar process as the projection of the functional process on a component of the chosen orthonormal basis. The non-stationarity of the global variance and cross-covariance comes from the fact that the linear bounded operators $A^u(\cdot)$ are allowed to vary in space. We also impose on all the scalar projections the non-stationary anisotropic correlation of model (2), i.e. $R(w^u_{\mathbf{s},k}, w^u_{\boldsymbol{\tau},k}) = R^u_{NS}(\mathbf{s},\boldsymbol{\tau})$, where, from now on, $w^u_{\mathbf{s},k} = \langle w^u_{\mathbf{s}}, e_k\rangle$. The resulting functional random field is fully non-stationary, since the mean $m_{\mathbf{s}}$ is not spatially constant, the non-stationarity for anisotropy comes from the form of the non-stationary autocorrelation functions of the projections of the processes $w^u_{\boldsymbol{s}}, k$, and we reasonably expect that, since $A^u(\cdot)$ are functions of the space variable, the family of cross-covariance operators would also be non-stationary with respect to space. The residual of process (5) $\delta_{\boldsymbol{s}} = \sum_{u=1}^{r} A^u(\boldsymbol{s}) w^u_{\boldsymbol{s}}$, is modelled as a sum of r independent terms in order to capture different layers of local anisotropy or variability. We can now obtain the full correlation structure of process (5), in what is our main contribution.

Theorem 2.1. *For any $\mathbf{s} \in D$, let $\Omega^2_{\mathbf{s}}$ be the covariance operator in $\mathcal{H}$ s.t. $\langle \Omega^2 e_n, e_m\rangle = \mathbb{E}(w_{\mathbf{s},n} w_{\mathbf{s},m})$ for any $n, m \in \mathbb{N}$, where $\{e_k\}_{k\in\mathbb{N}}$ is the orthonormal basis with respect to which process (5) is defined. The following hold:*

- *The operator $\Omega^2_{\mathbf{s}}$ is spatially constant in D and it is trace-class, i.e. $\Omega^2_{\mathbf{s}} = \Omega^2 \in \mathcal{B}_1(\mathcal{H})$, for any $\mathbf{s} \in D$. Consequently, the square-root of $\Omega^2_{\mathbf{s}}$ is Hilbert-Schmidt, i.e. $\Omega = \sqrt{\Omega^2} \in \mathcal{B}_2(\mathcal{H})$.*
- *The family of cross-covariance operators for the model (5) can be identified, for each $\mathbf{s}, \boldsymbol{\tau} \in D$, as*

$$C_{\mathbf{s},\boldsymbol{\tau}} = \sum_{u=0}^{r} A^u(\boldsymbol{\tau}) \Omega^2 A^u(\mathbf{s})^* R^u_{NS}(\mathbf{s},\boldsymbol{\tau}), \tag{6}$$

which represents a valid model.

Based on Theorem 2.1, we can now model functional processes of arbitrary complexity starting from scalar correlation functions. Moreover, since we know the full variability structure, we can easily retrieve the explicit expression for the trace-covariogram, directly extending the model of Paciorek and Schervish [2006].

Proposition 2.1. *The global covariance and trace-covariogram functions for the model (5), when r=1, can be computed as follows:*

$$C(\mathbf{s},\boldsymbol{\tau}) = \sigma(\mathbf{s})\sigma(\boldsymbol{\tau}) R^*_{NS}(\mathbf{s},\boldsymbol{\tau}) \tag{7}$$

with

$$\sigma^2(\mathbf{s}) = \|\Omega A^*(\mathbf{s})\|^2_{HS}, \tag{8}$$

and

$$R^*_{NS}(\mathbf{s},\boldsymbol{\tau}) = \frac{\langle \Omega A^*(\mathbf{s}), \Omega A^*(\boldsymbol{\tau})\rangle_{HS}}{\|\Omega A^*(\mathbf{s})\|_{HS} \|\Omega A^*(\boldsymbol{\tau})\|_{HS}} R_{NS}(\mathbf{s},\boldsymbol{\tau}) \tag{9}$$

In expression (9), $\langle A, B\rangle_{HS} := Tr(A^*B)$ with A,B Hilbert-Schmidt operators is the Hilbert-Schmidt inner product, and $\|A\|_{HS} := \sqrt{\langle A, A\rangle_{HS}}$ is the Hilbert-Schmidt norm.

In particular, $R^*_{NS}(\mathbf{s},\boldsymbol{\tau})$ is a valid non-stationary correlation function.

We finally introduce an assumption yielding a simplified version of model (5), specifically needed to deal with the trace-covariogram estimation and functional kriging.

Assumption 2.1. *There exists a spatial scalar function* $\alpha(\cdot) : D \to \mathbb{R}$ *and a linear bounded operator* $A \in B(\mathcal{H})$ *such that* $A(\mathbf{s}) = \alpha(\mathbf{s})A$, $\forall \mathbf{s} \in D$.

This can be interpreted as a proportionality assumption; the operators $A(\mathbf{s})$ are still allowed to vary spatially, so the global variance (8) function is non-stationary as well. Note that Assumption 2.1 is crucial to avoid the explicit dependence of the trace-covariogram on the coregionalization matrices, as it is equivalent to require that $\frac{\langle \Omega A^*(\mathbf{s}), \Omega A^*(\boldsymbol{\tau})\rangle_{HS}}{\|\Omega A^*(\mathbf{s})\|_{HS}\|\Omega A^*(\boldsymbol{\tau})\|_{HS}} = 1, \forall \mathbf{s}, \boldsymbol{\tau} \in D$. Under Assumption 2.1 the trace-covariogram becomes:

$$C(\mathbf{s}, \boldsymbol{\tau}) = Cov(\mathcal{X}_{\mathbf{s}}, \mathcal{X}_{\boldsymbol{\tau}}) = \sigma(\mathbf{s})\sigma(\boldsymbol{\tau})R_{NS}(\mathbf{s}, \boldsymbol{\tau}) \tag{10}$$

which is the exact functional equivalent of (4). Having established a consistent theoretical framework for non-stationary functional random processes, we shall now focus on the development of suited estimation procedures.

3. Covariance estimation and spatial prediction

The FNF-LMC (5) allows us to develop two independent non-stationary covariance estimation methods, Op-NS for the estimation of the family of cross-covariance operators (6) and Trace-NS for the estimation of the trace-covariogram function (10). The former method can be used to perform spatial prediction by means of discretized operatorial kriging, while the latter through trace-kriging with scalar weights. The estimation methods are a direct adaptation of the work in Fouedjio et al. [2016, 2018]. In these works, estimations are carried out in the framework of local stationarity, which assumes stationarity to be a viable assumption locally but not globally. For a functional process local stationarity can be interpreted as a slow spatial variability of its mean and of the spatial parameters of the family of cross-covariance operators. In particular, the complete correlation structure of process (5) can be characterized by the spatial family of vectors of parameters $\boldsymbol{\theta}_{\mathbf{s}_0} := \{A(\mathbf{s}_0)^u, \lambda^u_{1\mathbf{s}_0}, \lambda^u_{2\mathbf{s}_0}, \phi^u_{\mathbf{s}_0}\}_{u=1}^r$ at any location $\mathbf{s}_0 \in D$, where each $\lambda_{1\mathbf{s}_0}, \lambda_{2\mathbf{s}_0}, \phi_{\mathbf{s}_0}$ stand for the eigenvalues of the anisotropy matrix $\Sigma^u_{\mathbf{s}_0}$ of model (2) and the corresponding azimuth angle. For both Op-NS and Trace-NS the idea of the methods is to perform a set of local analyses in a pre-defined grid of locations, called *anchor points*, to obtain a discrete set of vectors of parameters in such locations; these are eventually smoothed to obtain a continuous map of spatially varying parameters and, consequently, a complete model of (cross-)covariance. So, the method involves, in the following order: (i) a local stationary trace-variogram and cross-variogram kernel estimator, (ii) a local weighted least squares parametric fitting of valid stationary models and (iii) a spatial kernel smoothing of the local estimates. The local stationarity assumption allows us to estimate the covariance and the mean separately, because in each local analysis we restrict the considered pairs to spatially close locations so they have approximately the same mean $m_{\mathbf{s}_i} \approx m_{\mathbf{s}_j}$ and we can resort to a slight variation of the usual stationary trace-variogram and cross-variogram estimators while reasonably considering them unbiased (e.g. for the trace-variogram $\mathbb{E}[\|\mathcal{X}_{\mathbf{s}_i} - \mathcal{X}_{\mathbf{s}_j}\|^2] = 2\gamma(\mathbf{s}_i - \mathbf{s}_j) + \|m_{\mathbf{s}_i} - m_{\mathbf{s}_j}\|^2$). In each anchor point, the empirical (cross-)variogram is different because, coherently with the local stationarity assumption, we weigh the importance of each pair of locations through a spatial kernel, in a way that pairs that are closer to the anchor points are more relevant to its estimate.

The mean is estimated in each anchor point $\mathbf{s}_l$ by the best linear unbiased estimator (BLUE) through ordinary kriging only considering points in a locally stationary neighbourhood of $\mathbf{s}_l$. The full drift $m_{\mathbf{s}}$ is retrieved by kernel smoothing of the local anchor estimates in the whole spatial domain D as it is done for any other parameter in $\boldsymbol{\theta}_{\mathbf{s}}$.

Eventually, spatial prediction can be performed through simple kriging, by approximating the actual mean of process (5) as the estimated smoothed mean. For Op-NS we estimate the coregionalization matrices and can use the discretized version of (6) for operatorial kriging. On the other hand, the structure of the method Trace-NS does not allow to estimate coregionalization matrices but only the global spatial standard deviation computed

as $\sigma_{\mathbf{s}} = \sqrt{tr(A^K_{\mathbf{s}} A^{KT}_{\mathbf{s}})}$. Because of this reason, when using Trace-NS, we do not have the possibility to compute the exact expression for the non-stationary trace-covariogram (7), so we always need to resort to the approximation (10) under Assumption 2.1, even when this could not be reasonable.

4. Results and conclusions

We explored the strengths and potential drawbacks of the novel methodologies through an extensive Monte Carlo study on simulated data. In particular, coherently with theoretical results, we found out that, under a spatial proportionality assumption about cross-covariances, the simplicity of the method Trace-NS makes it preferable to Cross-NS. However, when this assumption is not viable, Trace-NS is not flexible enough to retrieve a correct covariance structure, so Cross-NS proves to be a better choice for spatial prediction. Note that, even when parameters are correctly estimated for Trace-NS, if Assumption 2.1 is not viable, the global trace-covariogram model (7) should not be used. Indeed, simulations showed that spatially varying anisotropy was always adequately estimated by Trace-NS, even in non-proportional scenarios; so, if the estimation of local anisotropy is the specific goal of the analysis, we believe Trace-NS would be a better option compared with Op-NS. In completely stationary contexts, the novel non-stationary methods proved themselves robust enough to, at least, be a first exploratory approach in most contexts. If the estimated covariance structures clearly do not show non-stationarities, classic scalar and operatorial methods can still be a better option. Nevertheless, we recognize that, in contexts where Trace-NS is not sufficient to characterize the full variability structure of a spatial process, the numerical solution of the optimization problem associated with a complex cross-covariance operator can quickly become a bottleneck for the performance of the method.

We finally used the introduced methods to analyse rainfall data in the continental USA in order to identify areas where the probability of natural hazards would be highest. Conveniently with the specific aim of the problem, we model the process through probability density functions in the Bayes Hilbert space. Trace-NS delivered the best cross-validation performance and was employed to locate critical regions in the conclusive outcomes.

In conclusion, we think that the most significant contribution of this work is the interpretability and descriptive potential of the introduced non-stationary covariance structures. Even in contexts where the numerosity of data makes it computationally prohibitive to perform kriging, we can still estimate non-stationarities (e.g. local anisotropy, spatial variations in autocorrelation) to gain a deeper comprehension of the statistical properties of processes generating data.

5. Implementation

The methods in Fouedjio et al. [2016, 2018] and the novel methods for functional data (Trace-NS, Op-NS) were implemented in *LocallyStationaryModels* an R package entirely based on C++ code (De Carlo et al. [2022]).

References

G. De Carlo, L. Crippa, and R. Scimone. LSM, an R package. 2022. URL `https://github.com/giacomodecarlo/FunctionalLSM`.

F. Fouedjio, N. Desassis, and J. Rivoirard. A generalized convolution model and estimation for non-stationary random functions. *Spatial Statistics*, page 16:35–52, 2016.

F. Fouedjio, N. Desassis, and J. Rivoirard. A fully non-stationary linear coregionalization model for multivariate random fields. *Stochastic Environmental Research and Risk Assessment*, pages 32(6):1699–1721, 2018.

A. Menafoglio and P. Secchi. Statistical analysis of complex and spatially dependent data: a review of object oriented spatial statistics. *European Journal of Operational Research*, page 258(2):401–410, 2017.

C.J. Paciorek and M.J. Schervish. Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, pages 17:483–506, 2006.