



MOX-Report No. 07/2025

Functional-Ordinal Canonical Correlation Analysis With Application to Data from Optical Sensors

Patanè, G.; Nicolussi, F.; Krauth, A.; Gauglitz, G.; Colosimo, B. M.; Dede', L.;
Menafoglio, A.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it

<https://mox.polimi.it>

Functional-Ordinal Canonical Correlation Analysis With Application to Data from Optical Sensors

Giulia Patanè^{*,†}, Federica Nicolussi[†], Alexander Krauth[§],
Günter Gauglitz[§], Bianca Maria Colosimo[‡],
Luca Dede[†], and Alessandra Menafoglio[†]

* `giulia.patane@polimi.it`

† MOX, Department of Mathematics,
Politecnico di Milano,
Piazza Leonardo da Vinci 32, 20133 Milano, Italy.

§ Institute of Physical and Theoretical Chemistry (IPTC),
Eberhard Karls Universität Tübingen,
Auf der Morgenstelle 18, 72076 Tübingen, Germany

‡ Department of Mechanical Engineering,
Politecnico di Milano,
Piazza Leonardo da Vinci 32, 20133 Milano, Italy.

January 29, 2025

Abstract

We address the problem of predicting a target ordinal variable based on observable features consisting of functional profiles. This problem is crucial, especially in decision-making driven by sensor systems, when the goal is to assess a ordinal variable such as the degree of deterioration, quality level, or risk stage of a process, starting from functional data observed via sensors. We purposely introduce a novel approach called functional-ordinal Canonical Correlation Analysis (foCCA), which is based on a functional data analysis approach. FoCCA allows for dimensionality reduction of observable features while maximizing their ability to differentiate between consecutive levels of an ordinal target variable. Unlike existing methods for supervised learning from functional data, foCCA fully incorporates the ordinal nature of the target variable. This enables the model to capture and represent the relative dissimilarities between consecutive levels of the ordinal target, while also explaining these differences through the functional features. Extensive simulations demonstrate that foCCA outperforms current state-of-the-art methods in terms of prediction accuracy

in the reduced feature space. A case study involving the prediction of antigen concentration levels from optical biosensor signals further confirms the superior performance of foCCA, showcasing both improved predictive power and enhanced interpretability compared to competing approaches.

Keywords: Functional Data Analysis, Ordinal Data, Canonical Correlation Analysis, sensors

1 Introduction

In the current industrial landscape, quality features - such as product flaws, process yield, or degradation states—are increasingly linked to data acquired via sensors (e.g., signals, images, and video images). In many cases, the quality level is measured on an ordinal scale, while the observed signal data can be modeled as functional data. This problem naturally arises in the presence of modern sensing technologies that generate high-dimensional temporal sequences of data through video or image records. Examples can be found in manufacturing, where processes involving local thermal histories (e.g. welding, additive manufacturing, casting) often correlate final local flaws (e.g., porosity) or microstructure classifications to the spatio-temporal temperature patterns observed at the same location - see Figure image a) and b) (Bugatti & Colosimo, 2022; Yan, 2022). Similarly, in life-science and biotechnology applications, as the motivating example we will focus on, the reaction class observed at a given location can be linked to the local intensity patterns captured through video imaging (Figure 1, image c)).

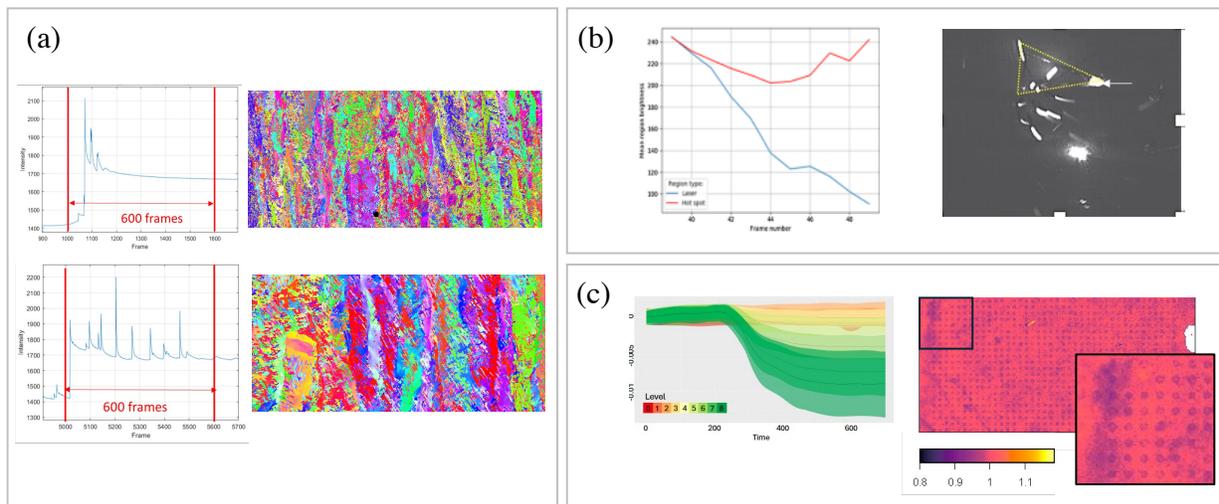


Figure 1: a) Functional data representing the cooling history observed during additive manufacturing in one location and the final microstructure (fine microstructure on the top and coarse microstructure on the bottom) b) Functional data representing two different cooling history observed during additive manufacturing with negligible or severe hot-spot phenomena (classified depending on severity) c) Functional data observed at different levels of the antibody-antigen interactions (case study described in Section 5).

Further examples of works that employ time-varying signals for ordinal prediction are typically found in the fields of psychometry and neuroscience (Li et al., 2023; Movahed et

al., 2021), medicine (Ruan et al., 2022), and again in manufacturing (Kundu et al., 2022; Zhou et al., 2011).

In this context, it is essential to define a method that leverages functional data as a predictive driver to estimate the final quality class on an ordinal scale. Although critical for the applications, to the best of our knowledge, no available methodology leverages high-dimensional or functional signals to predict an ordinal variable. In fact, state-of-the-art methods to address a similar problem typically include a preliminary dimensionality reduction step, followed by a correlation analysis. The preliminary step aims to capture the variability of the functional data without accounting for the need to explore their ability to predict the ordinal target variable. For example, within the framework of functional regression models, several authors have suggested using Functional Principal Component Analysis (fPCA, Ramsay and Silverman (2005)) as a dimensionality reduction step (Müller, 2005). Here, fPCA is applied to reduce the dimensionality of the data, and then the resulting scores are used in place of functional covariates (Akturk et. al, 2024), functional responses (Beretta et. al, 2020), or both (Hullait et. al, 2020). Unfortunately, when dealing with high-dimensional data, as in our case study, the global variability of the dataset can hide the ordinal patterns, significantly limiting the prediction power of the fPC scores. In the classification setting, functional Fisher’s Discriminant Analysis has been proposed (Hastie et. al, 1995; Ramsay and Silverman, 2005), as an alternative to fPCA. However, when applied to ordinal variables, Fisher’s approach may struggle to identify components that effectively discriminate between consecutive levels, especially if the ordinal levels are not evenly spaced.

Our novel contribution aims at bridging this gap by providing a statistical methodology for ordinal prediction based on a functional signal. Our novel approach jointly tackles the limitations of the current literature outlined above by proposing a novel method called functional-ordinal Canonical Correlation Analysis (foCCA). Our approach is inspired by functional Canonical Correlation Analysis (Ramsay and Silverman, 2005), but reinterprets this in an ordinal setting, by adapting the key concept of *optimal scoring* discussed by Hastie et. al (1995). In our newly developed solution, foCCA allows one to reduce the dimensionality of a functional dataset while maximizing correlation of the scores with the ordinal variable of interest, thus establishing interpretable differences between adjacent levels through an appropriate weight function. We shall show that foCCA enables one to (i) effectively reduce the data complexity, (ii) perform supervised classification of the ordinal levels in the reduced space, and (iii) obtain a set of functional and ordinal directions (which play the same role as the loadings in fPCA), which can be used to enhance the model interpretation.

In Section 2 we formulate the problem and we recall the dimensionality reduction methods typically employed in the functional setting. Section 3 introduces foCCA, providing some insights in the discretization of the problem, the interpretation of the canonical directions and the penalty parameter selection. In Section 4 we show an extensive Monte Carlo simulation used to analyse the performance of foCCA and state-of-the-art methods under varied conditions. This work, though widely applicable, is primarily motivated by a monitoring challenge from biosensor data, specifically focused on assessing antigen concentration levels within specific Regions of Interest (ROIs) based on video signals that track the temporal progression of an antibody-antigen reaction. In Section 5 we present the motivating case study and we show the performance of foCCA on its dataset. Section 6

presents the main conclusions of the work.

2 Problem formulation and state of the art

Let \mathcal{C} be an ordinal random variable with levels from 0 to $n_{\mathcal{C}}$.

Suppose that we are provided with n independent realizations of the functional variable X and \mathcal{C} , then our samples consists of pairs $\{x_i(t), c_i\}_{i=1}^n$. The main objective is to be able to predict the level \mathcal{C} , knowing the functional variable X . For any $c \in \{1, \dots, n_{\mathcal{C}}\}$, let D_c be defined as $D_c = \mathbb{1}_{\mathcal{C} \geq c}(\mathcal{C})$, and let $\{e_c\}_c$ be the canonical basis of $\mathbb{R}^{n_{\mathcal{C}}}$.

Only in foCCA, we encode \mathcal{C} into the random variable Y , defined as follow:

$$Y = \sum_{c=1}^{n_{\mathcal{C}}} D_c e_c \quad (1)$$

Notice that, by construction, the vector Y and the variable \mathcal{C} are in one-to-one correspondence, hence they contain exactly the same information.

In Section 2.1 and 2.2, we provide an overview of classical methodologies that are widely employed for dimensionality reduction, aiming to capture essential features within high-dimensional datasets in a predictive context. Understanding these classical approaches is fundamental for contextualizing and contrasting with the innovative method introduced later in this study.

2.1 Functional Principal Component Analysis

PCA stands as a cornerstone in the field of dimensionality reduction, finding widespread applications across various domains, from machine learning to statistics. The PCA objective is to identify linear combinations of features that maximize the variance within a given dataset. While the introduction of PCA traces back to its roots in multivariate data analysis (Hotelling (1933)), its scope has considerably expanded in recent years to encompass diverse contexts and various data types. This adaptability has allowed PCA to be applied to entities ranging from points and functions to images and beyond. In this broad context, fPCA (and its smoothed versions) allows to understand and extract patterns from functional observations. In the following, we set the notation and briefly recall the smoothed version of the methodology; we refer to Ramsay and Silverman (2005), for further details. Let $(\Omega, \mathfrak{F}, \mathbb{P})$ be a probability space, and, for $\omega \in \Omega$, denote by $X(\omega)$ an element of the functional Hilbert space $L^2(I)$, where I is a closed interval in \mathbb{R} . Suppose we are provided with a functional dataset $\{x_i\}_{i=1}^n$, independent realizations of X , and we aim at reducing its dimensionality. Formally, the first functional principal component is the function $\beta^*(t) \in L^2(I)$ that maximizes the *penalized sample variance*:

$$\beta^*(t) = \operatorname{argmax}_{\beta} \frac{\widehat{\operatorname{Var}}(\langle \beta, X \rangle_{L^2(I)})}{\|\beta\|_{L^2(I)}^2 + \lambda \operatorname{Pen}(\beta)}, \quad (2)$$

where $\widehat{\operatorname{Var}}(\cdot)$ is the sample variance and, typically, $\operatorname{Pen}(\beta) = \|\beta''\|_{L^2(I)}^2$ and $\lambda > 0$ is the penalization parameter (Silverman (1996)). We can compute the subsequent principal components, with the additional constraint of being orthogonal to previously found

components, by solving the equivalent eigenvalue problem (Silverman (1996)). FPCA allows to gain valuable insights into the inherent structures and dynamics of the underlying processes. Nevertheless, when the objective is to utilize a functional dataset for predicting a new object, employing fPCA does not guarantee that the extracted components are the most appropriate for forecasting purposes, e.g. when the variance explained by any covariate is masked by a high global variance.

2.2 Functional Fisher's Discriminant Analysis

When the aim is supervised classification, a well-known dimensionality reduction technique is Fisher's Discriminant Analysis, by Fisher (1936). As PCA, this technique was extended to the functional framework in Ramsay and Silverman (2005), based on Hastie et. al (1995), in which Penalized Discriminant Analysis is discussed and the equivalence with Discriminant CCA and Optimal Scoring is demonstrated for categorical labels. Let X be a random function as in the previous Section (2.1). Let us suppose that our dataset consists of pairs $\{(x_i(t), c_i)\}_{i=1}^n$, where, for any i , x_i is a realization of X and $c_i \in \{c_0, c_1, c_2, \dots, c_{n_c}\}$ is the realization of a categorical variable. The first functional Fisher discriminant component is a function $\beta^* \in L^2(I)$ which maximizes the between-class sample variance, ensuring that the projected classes are well-separated, while simultaneously minimizing the within-class sample variance to maintain compactness within each class, formally:

$$\beta^*(t) = \operatorname{argmax}_{\beta} \frac{\widehat{\operatorname{Var}}_B(\langle \beta, X \rangle_{L^2(I)})}{\widehat{\operatorname{Var}}_W(\langle \beta, X \rangle_{L^2(I)}) + \lambda \operatorname{Pen}(\beta)}, \quad (3)$$

where Var_B is the between-class sample variance, and Var_W is the within-class sample variance, i.e.:

$$\left\{ \begin{array}{l} \widehat{\operatorname{Var}}_B(\langle \beta, X \rangle_{L^2(I)}) = \frac{1}{n_c} \sum_{c \in \{c_0, \dots, c_{n_c}\}} n^{(c)} (\langle \beta, \bar{x}_c - \bar{x} \rangle)^2 \\ \widehat{\operatorname{Var}}_W(\langle \beta, X \rangle_{L^2(I)}) = \frac{1}{n} \sum_{i \in \{1, \dots, n\}} (\langle \beta, x_i - \bar{x}_{c_i} \rangle)^2 \\ \bar{x}_c = \frac{1}{n^{(c)}} \sum_{i: c_i=c} x_i \quad \text{and} \quad \bar{x} = \frac{1}{n} \sum_{i \in \{1, \dots, n\}} x_i \\ n^{(c)} = |\{i : c_i = c\}| \quad \text{for any} \quad c \in \{c_0, c_1, \dots, c_{n_c}\} \end{array} \right. \quad (4)$$

This technique allows to extract the component that best discriminates among the classes. The Fisher's concept can be modified to be used with an ordinal label, rather than categorical, extending the idea of (McFee and Ellis (2005)) to the functional framework, i.e. substituting the classical sample between covariance Var_B with:

$$\widehat{\operatorname{Var}}_B^O(\langle \beta, X \rangle) = \sum_{c=0}^{n_c-1} n^{(c)} (\langle \beta, \bar{x}_c - \bar{x}_c^+ \rangle)^2 + n^{(c+1)} (\langle \beta, \bar{x}_{c+1} - \bar{x}_c^+ \rangle)^2 \quad (5)$$

where \bar{x}_c^+ is the mean vector of the x_i 's belonging to the union of the two consecutive levels c and $c+1$. With this modification, the discrimination is maximized between consecutive levels instead of among all the possible classes. In the following, we will refer to this method

with the name *functional-ordinal Fisher Discriminant* (foFD). When the underlying data distribution is approximately Gaussian and the class covariances are assumed to be equal, foFD provides an optimal linear transformation for discrimination. In view of this, foFD could be considered as a competitor method of that proposed in this work. However, we do not pursue this approach further as it is subject to limitations, e.g., when data do not come from a Gaussian process (Johnson and Wichern (2015)), when the variance within each class is maximized along the same direction as the variance between classes (since the same direction maximizes the numerator and the denominator of the fraction in Equation 3) or when the groups are separated by very different distances. The latter issue holds also net of the modification of the between variance definition for ordinal labels, proposed by McFee and Ellis (2005).

3 Functional-ordinal Canonical Correlation Analysis

We here present the functional-ordinal Canonical Correlation Analysis, which aims to provide a dimensionality reduction of a functional dataset, as to maximize the discrimination ability, in the reduced space, among the levels of an ordinal variable.

3.1 The maximization problem

Suppose that we are provided with n independent realizations of the functional variable X and the vector Y , which is defined in Equation 1 starting from an ordinal random variable \mathcal{C} . Thus, our samples consists of pairs $\{x_i(t), y_i\}_{i=1}^n \in L^2(I) \times \{0, 1\}^{n_c}$. The main objective is to be able to predict the Y , i.e. the level \mathcal{C} , knowing the functional variable X . We can reduce the dimensionality of the functional part of the dataset while extracting the features that best discriminate among different levels, adapting Canonical Correlation Analysis to the functional-ordinal context. More precisely, foCCA looks for $\beta^* \in L^2(I)$ and $\theta^* \in \mathbb{R}^{n_c}$ that maximize the sample correlation between the following random variables

$$\langle \beta, X \rangle_{L^2(I)} \quad \text{and} \quad \langle \theta, Y \rangle_{\mathbb{R}^{n_c}}$$

We call $\beta(t)$ the functional canonical component and θ the ordinal canonical component, while $\langle \beta, X \rangle_{L^2(I)}$ and $\langle \theta, Y \rangle_{\mathbb{R}^{n_c}}$ are, respectively, the functional and the ordinal scores. Formally, in analogy to Leurgans et al. (1993), where both datasets are functional, we want to find the functional object $\beta^*(t)$ and the vector θ^* that maximize the *penalized squared sample correlation*:

$$ccor_{sq_{\lambda_1, \lambda_2}}(\beta, \theta) = \frac{\widehat{Cov}(\langle \beta, X \rangle_{L^2(I)}, \langle \theta, Y \rangle_{\mathbb{R}^{n_c}})^2}{\left(\widehat{Var}(\langle \beta, X \rangle_{L^2(I)}) + \lambda_1 Pen(\beta) \right) \left(\widehat{Var}(\langle \theta, Y \rangle_{\mathbb{R}^{n_c}}) + \lambda_2 Pen(\theta) \right)} \quad (6)$$

where $\widehat{Cov}(\cdot)$ and $\widehat{Var}(\cdot)$ are, respectively, the sample covariance and variance, $Pen(\beta)$ is a smoothing penalization and $Pen(\theta)$ is an elastic net penalization, with the respective penalty parameters $\lambda_1 > 0$, $\lambda_2 > 0$. The need for penalization is deeply discussed in Silverman (1996). In our work, for the functional canonical component the penalization is $Pen(\beta) = \|\beta''\|_{L^2(I)}^2$, typically used in smoothing spline regression (Ramsay and

Silverman (2005)), and for the ordinal canonical component $Pen(\theta) = \|\theta\|^2$, i.e. the Ridge penalization (Hoerl and Kennard (1970)). Given that the described optimization problem is equivalent to a generalized eigenvalue problem, we can derive the first pair of canonical components $(\beta^{(1)}(t), \theta^{(1)})$ corresponding to the largest eigenvalue $\rho^{(1)}$, the second pair $(\beta^{(2)}(t), \theta^{(2)})$ corresponding to the second largest eigenvalue $\rho^{(2)}$, and so forth, up to a desired number of components m . In the next Section 3.2 we elucidate the computational tools employed to derive the pairs of canonical components, providing a detailed resource.

3.2 Discretization of the maximization problem

We here enter into the details of the discretization of the dataset and of the optimization problem. Moreover, we provide the subsequent derivation of the generalized eigenvalue problem. These steps allows the computation of the sequence of m canonical correlation components. In the following, we assume both the functional and the ordinal data to be empirically centered to facilitate the computations. Net of the centering, the maximization problem (6) is equivalent to the following:

$$(\beta^*, \theta^*) = \operatorname{argmax}_{\beta, \theta} \frac{\frac{1}{n} \sum_{i=1}^n \langle \beta, X_i \rangle_{L^2(I)} \langle \theta, y_i \rangle_{\mathbb{R}^{n_c}}}{\left(\frac{1}{n} \sum_{i=1}^n \langle \beta, X_i \rangle_{L^2(I)}^2 + \lambda_1 \langle \beta, \beta \rangle_{L^2(I)} \right) \left(\frac{1}{n} \sum_{i=1}^n \langle \theta, y_i \rangle_{\mathbb{R}^{n_c}}^2 + \lambda_2 \langle \theta, \theta \rangle_{\mathbb{R}^{n_c}} \right)} \quad (7)$$

Let us focus on the discretization of the functional dataset, as the ordinal dataset is already finite dimensional. Suppose that we can represent, possibly after a smoothing regression, each functional datum $x_i(t)$, where $1 \leq i \leq n$, through a finite basis expansion, i.e.:

$$x_i(t) = \sum_{j=1}^{n_J} a_{ij} \phi_j(t)$$

where a_{ij} is the coefficient associated with the basis function $\phi_j(t) \in L^2(I)$. If we define the following matrices:

- $\mathcal{A} = \{a_{ij}\}_{ij}$, where a_{ij} is the coefficient of the i -th unit associated to the j -th spline, in centered dataset $\{x_i(t)\}_i$.
- $\mathcal{D} = \{d_{il}\}_{il}$ where d_{il} is the coefficient of the i -th unit associated to the l -th basis vector, in the centered $\{y_i\}_i$.
- $V_{12} = n^{-1} \mathcal{A}^T \mathcal{D}$, $V_{11} = n^{-1} \mathcal{A}^T \mathcal{A}$, $V_{22} = n^{-1} \mathcal{D}^T \mathcal{D}$.
- $K = \left\{ \int_I \frac{\partial^2 \phi_j}{dt^2} \frac{\partial^2 \phi_k}{dt^2} \right\}_{j,k}$
- $J = \left\{ \int_I \phi_j \phi_k \right\}_{j,k}$

By setting $\beta(t) = \sum_j b_j \phi_j(t)$, the maximization problem in Equation 7 can be discretized as follows:

$$(b^*, \theta^*) = \operatorname{argmax}_{b, \theta} \frac{(b^T J V_{12} \theta)^2}{(b^T J V_{11} J b + \lambda_1 b^T K b) (\theta^T V_{22} \theta + \lambda_2 \theta^T \theta)} \quad (8)$$

which is in turn equivalent to maximize the value $(b^T JV_{12}\theta)$ imposing $(b^T JV_{11}Jb + \lambda_1 b^T Kb) = (\theta^T V_{22}\theta + \lambda_2 \theta^T \theta) = 1$. Hence, the optimization problem in (6) is equivalent to the following generalized eigenvalue problem (Ramsay and Silverman (2005)):

$$\begin{bmatrix} 0 & JV_{12} \\ V_{21}J & 0 \end{bmatrix} \begin{bmatrix} b \\ \theta \end{bmatrix} = \rho \begin{bmatrix} JV_{11}J + \lambda_1 K_1 & 0 \\ 0 & V_{22} + \lambda_2 I \end{bmatrix} \begin{bmatrix} b \\ \theta \end{bmatrix} \quad (9)$$

The eigenvector $[b^{*T}, \theta^{*T}]$ with highest eigenvalue ρ allows to reconstruct the first canonical component as $\beta^*(t) = \sum_j b_j^* \phi_j(t)$, while the second block θ^* directly provides the first ordinal canonical component.

3.3 Interpretation of canonical components

In this section, we enter into the interpretation of the canonical components derived from the maximization problem introduced above. By elucidating the interpretation of these components, we seek to unravel valuable insights into the relationships between the functional and the ordinal datasets. Recall that foCCA computes the function $\beta^* \in L^2(I)$ and the vector $\theta^* \in \mathbb{R}^{n_C}$ that maximize the correlation between the functional scores $\langle \beta, X \rangle_{L^2(I)} = \int_I \beta(t)X(t)dt$ and the ordinal scores $\langle \theta, Y \rangle = \sum_{j=1, \dots, n_C} \theta_j Y_j$. The functional canonical component $\beta^*(t)$ is the weight given to the functional datum $X(t)$ across time, in the functional score computation. This function makes clear which are temporal intervals that are valuable for discrimination and highlights how they influence the functional score that optimizes the prediction of the ordinal variable. For what concerns the ordinal canonical component, notice that the suitable representation of \mathcal{C} through Y allows to make flexible the distance between consecutive levels in the correlation analysis, exploiting the ordinal nature of the variable \mathcal{C} . Indeed, θ^* contains the components representing the step between two consecutive ordinal variables. For instance, if we consider two realizations of Y , such that $c_1 = \hat{c} - 1$ and $c_2 = \hat{c}$ ($\hat{c} \in \{1, \dots, n_C\}$), i.e.

$$y_1 = \sum_{c=1}^{\hat{c}-1} e_c \quad \text{and} \quad y_2 = \sum_{c=1}^{\hat{c}} e_c ,$$

then the difference between the two ordinal scores is $(\theta^{*T} y_2 - \theta^{*T} y_1) = \theta^{*T} (y_2 - y_1) = \theta^{*T} e_{\hat{c}} = \theta_{\hat{c}}$. It is important to note that even after centering the dataset Y , the disparity between two ordinal observations remains constant. Consequently, the interpretation of the coefficients in θ^* remains unaffected by this centering process.

3.4 A note on penalty parameter selection

Since our final goal is to optimize the dimensionality reduction in terms of classification power, we select the parameters with a cross-validation method: we fix the K-fold Cross Validated Mean Absolute Error (MAE_{CV}) as the loss function, i.e.:

$$MAE_{CV} = \sum_{k=1}^K \sum_{i \in F_k} \frac{|c_i - \hat{c}_i^{(-F_k)}|}{n_{F_k}} , \quad (10)$$

where $\hat{c}_i^{(-F_k)}$ is the level associated with the nearest centroid (calculated by excluding F_k) within the space of the first m components, where m is fixed. Here, we identify the penalty parameters corresponding to the minimum of the smoothed loss curve (or surface). As revealed by sensitivity analyses, smoothing is essential as the narrow selection of the penalty parameter based solely

on the raw loss curve (or surface) may be influenced by the subdivision into K-folds. To elaborate further, the loss is smoothed through smoothing cubic splines regression (with smoothing parameter selected via GCV, Wood (2004)); since in foCCA we have to select two parameters, we consider smoothing regression with tensor product of cubic splines (Wood et. al (2015)).

4 Simulation Study

In this section, we present a simulation study conducted to assess the performance of the dimensionality reduction method foCCA, and comparing it with the performance of fPCA and foFD, under specific conditions, via Monte Carlo simulations. The foCCA method, by design, is expected to outperform competitors in identifying functional components that effectively differentiate between various levels of an ordinal variable, even in cases where the intervals between consecutive levels are not uniformly distanced. Consequently, we conduct comparative analyses across two distinct simulated scenarios: (a) A scenario where the variance attributed to the ordinal variable is lower than the overall dataset variance, which, in particular, poses challenges for fPCA; (b) A scenario where the successive levels of the ordinal variable exhibit non-uniformity in distance, which, in principle, poses challenges to all the methods. Codes and data to reproduce the simulated experiments are available at <https://github.com/gPatstat/foCCA>.

This structured approach allows us to systematically evaluate and compare how each method responds to challenges specific to the underlying data structures, providing information into their respective strengths and limitations in practical applications.

4.1 Simulated scenarios

The simulated functional dataset consists of $n = 1000$ functional data built on a basis of $J = 10$ cubic B-splines, with equispaced knots and range $I = [0, 100]$, where the coefficients of the splines are independently randomly sampled from a distribution that depends on the scenario we aim at creating. The number of levels is fixed equal to $(n_c + 1) = 9$, as in the case study in Section 5, and for any unit $i \in \{1, \dots, n\}$, the level c_i is randomly sampled from the discrete uniform distribution over $\{0, 1, 2, \dots, 8\}$.

In the scenario (a), a singular Monte Carlo simulation consists of a dataset of realizations $(x_i(t), c_i)_{i=1}^n$. For any unit $i \in \{1, \dots, n\}$, $x_i(t) = \sum_{j=1}^J a_{ij} \phi_j(t)$, where a_{ij} is the coefficient of the i -th unit associated with the j -th spline of the basis. The coefficients $\{a_{ij}\}_{i=1}^n$ are independently sampled from the normal distribution $\mathcal{N}(\mu_j + h_{ij}, \sigma_j)$, where, for any spline j , μ_j is the general mean, h_{ij} is a level-specific increment and σ_j is the standard deviation. For any Monte Carlo simulation, $\{\mu_j\}_{j=1}^J$ are i.i.d sampled from $U([-10, 10])$, and $\{\sigma_j\}_{j=1}^J$ are i.i.d sampled from $U([0, 10])$. The level-specific increment h_{ij} is equal to $\sum_{k=0}^{c_i} \gamma(k)$ if $j = s$, and it is equal 0 otherwise. For any Monte Carlo simulation the spline $s = \operatorname{argmin}_{j \in \{1, \dots, J\}} \{\sigma_j\}$ is the spline

along which we generate the ordinal pattern, defined by the increment $\gamma(k) \sim \mathcal{N}(q_a \cdot 10, 1)$ for any $k \in \{1, \dots, n_c\}$ and $\gamma(0) = 0$. In particular, q_a denotes a parameter that controls the severity of the perturbations across levels applied in the generated scenario. The reason why we generate different $\{\mu_j\}_{j=1}^J$, $\{\sigma_j\}_{j=1}^J$ and γ for each Monte Carlo simulation, is that we want to be sure that all the methods are challenged: indeed, for any simulated dataset we change the mean function, the direction along which the ordinal pattern is built, the direction along which the global variance is the highest and the mutual distances between the consecutive levels. Hence, the methods are compared on an exhaustive and wide scenario. Notice that, when $q_a = 0$ the increment is

null for each spline j , i.e. the variance explained by the ordinal variable is zero along the entire domain. On the contrary, when $q_a > 0$, along the spline s the coefficients have a mean increment $(q_a \cdot 10)c_i$, i.e. the larger q_a , the higher the expected ratio between the variance explained by the ordinal variable and the global variance. See Figure 2 to see some examples of scenarios of type (a).

In the scenario (b) the successive levels of the ordinal variable exhibit a non-uniform separation, both in terms of distance and in terms of component along which the distance is maximized. Each Monte Carlo simulation consists of a dataset of realizations $(x_i(t), c_i)_{i=1}^n$, with $x_i(t) = \sum_{j=1}^J a_{ij} \phi_j(t)$, where a_{ij} are independently sampled from $\mathcal{N}(\mu_{ij}, 1)$. In particular, $\mu_{ij} = \mu_j^{(0)} + q_b \mathbb{1}_{\{c_i \geq 1\}} (\sum_{1 \leq c \leq \min(4, c_i)} \mu_j^{(1)}) + \mathbb{1}_{\{c_i \geq 5\}} \sum_{5 \leq c \leq c_i} \mu_j^{(2)}$, where for any Monte Carlo simulation $\mu_j^{(k)}$ are i.i.d sampled from $\mathcal{U}(-1, 1)$. Hence, the mean μ_{ij} is characterized by a global mean term $\mu_j^{(0)}$ and by two additional terms that depend on the level c_i associated with the unit. More precisely, $\mu_j^{(1)}$ is the increment applied to μ_{ij} for each level $k \leq c_i$ with $k = 1, \dots, 4$, while $\mu_j^{(2)}$ is the analogous increment among contiguous levels but for the highest levels $k \geq c_i$ with $k = 5, \dots, 8$. In practice, let us consider to start with $\mu_{ij} = \mu_j^{(0)}$. The value q_b is the ratio between the increments of the highest levels and the one of the lowest levels, hence it represents the severity of the Scenario (b). A lower value of q_b indicates less uniform separation between consecutive levels, illustrating greater disparity in the increments for higher levels compared to lower levels.

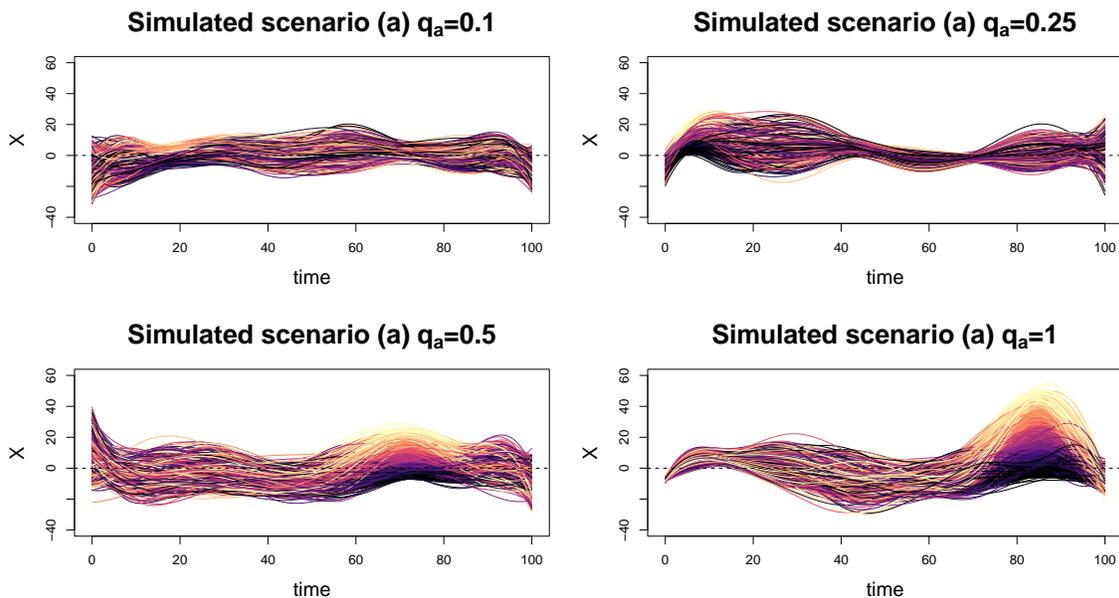


Figure 2: Examples of simulated scenarios of type (a), depending on q_a

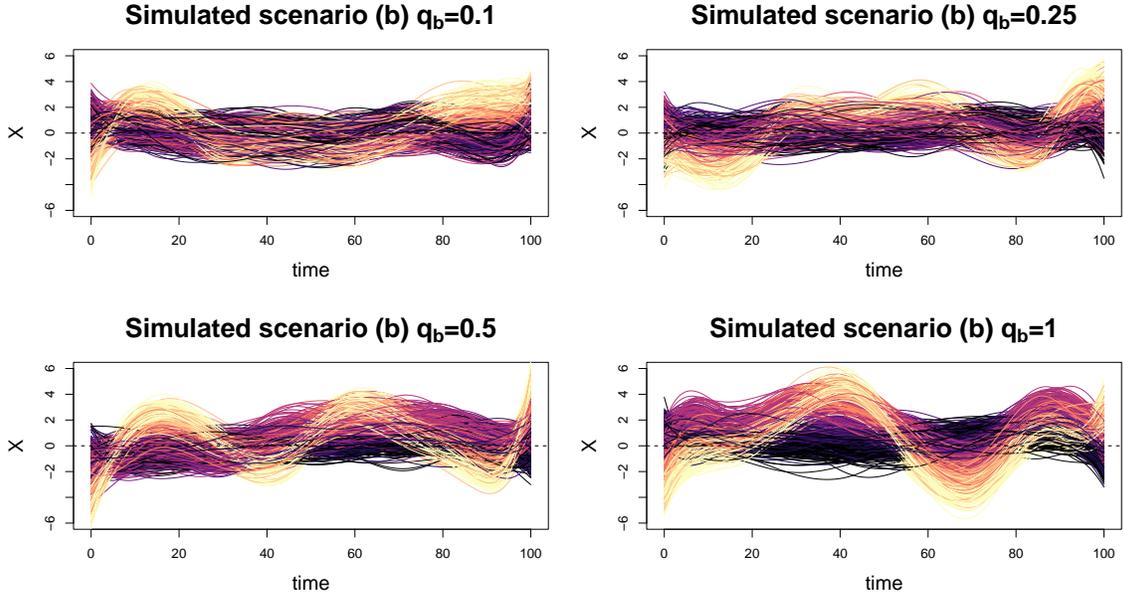


Figure 3: Examples of simulated scenarios of type (b), depending on q_b

4.2 Results

In this Section, we aim to estimate the performance of foCCA via Monte Carlo simulations, built by varying the ratio q_a , i.e. generating scenarios of type (a), and q_b , i.e. generating scenarios of type (b). We compare the results with the ones obtained through fPCA and foFD. We run $M = 500$ simulations for each scenario, letting q_a and q_b vary in the range $\{0, 0.1, 0.2, \dots, 1\}$, and we evaluate the K-fold MAE, defined in Section 3.4, with $K = 5$. By doing this, we can see how the severity of the scenario affects the models.

Results are depicted in Figure 4. In the scenario (a) we can observe that the performances of all the three methods improve as q_a increases. This is due to the fact that, the higher q_a , the clearer the separation of the levels (see also Figure 2). Indeed, a high ordinal variability, with respect to the global variability, drive all the methods towards the identification of components that intrinsically separate the levels. However, foCCA proves consistently better than the competitors across all the severity levels. Notice that, when q_a is high, foFD is the worst method. A possible reason for this result relies in the fact that the direction that maximizes the ordinal variance is, by construction, the same as the one that maximizes the global variance. Contrarily, when q_a is small, the performance of fPCA is the poorest. This phenomenon can be explained by the fact that, in this scenario, the ordinal variance is negligible with respect to the global variance, hindering the identification of discriminant directions through fPCA.

Let us now focus on scenario (b). In this scenario, whatever the value of q_b , the levels are not equally separated. However, as q_b grows, the recognition of the lowest levels should, in principle, be possible for each method. We can observe from the Monte Carlo simulations, in Figure 4, that as q_b increases, the error of all the methods decreases. However, for each $q_b \in [0, 1]$, while fPCA and foFD tends to overlap in terms of MAE, the difference between these two methods and foCCA is always high in mean. In conclusion, foCCA is able to keep a significantly lower error than competitors in both scenarios proposed. The methods behave in a similar way only when there is an ordinal variance comparable to the global one, with levels equispaced along the same components.

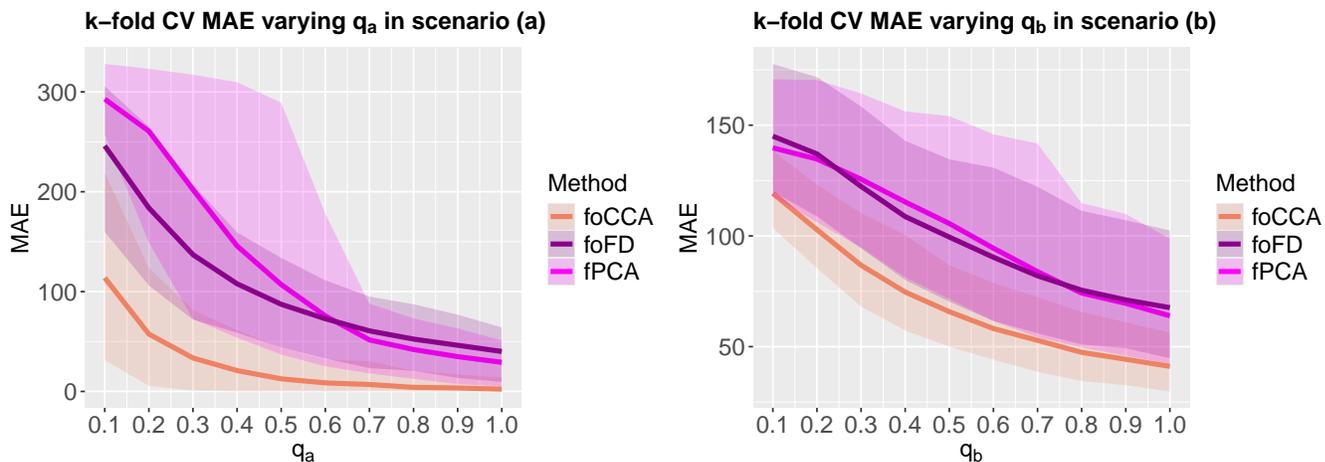


Figure 4: Monte Carlo simulations: the line represents the mean values, while the bands' boundaries are the pointwise quantiles 5% and 95%

5 A case study: analyzing biological interactions from biosensor video signals

5.1 Optical biosensors data: background

In recent decades, optical biosensors have become increasingly important in a wide variety of areas such as biomedical diagnostics (Sharma et al. (2021)), drug discovery (Cooper (2006)), food safety (Narsaiah et al. (2012)), or environmental monitoring (Sharma & Sharma (2023); Long, Zhu & Shi (2013)). To enhance the time-efficiency and sustainability of optical biosensors, imaging techniques have been developed to measure multiple interactions between the receptor (such as an enzyme, nucleic acid, an antibody, tissue, organelles or even whole cell) and the analyte (e.g. an antigen of interest) in parallel. Due to these challenges and progressions in complexity and design (Narlawar et al. (2024); Gauglitz (2010)), the generated datasets are becoming high-dimensional, yielding time-consuming and memory-intensive data processing.

In this study, a reflectometric imaging sensor is used to track the progression of the interaction between an antibody and an antigen immobilized onto a transducer, based on a video of the light intensity. The primary statistical challenge lies in analyzing temporal profiles of light intensity within Regions of Interest (ROIs), where the antibody is present with a nominal concentration level. In particular, our interest lies in analyzing the light intensity evolution depending on the level of antigen concentration, and therefore, in predicting the level of antigen concentration, based the optical signal, in a new ROI.

5.2 Preprocessing

The available dataset consists of a video signal, recording one image per second, captured through a reflectometric sensor. Each frame represents the light intensity of the reflected light, measured at a time instants. Here, 1035 spots—each containing a certain concentration of the antigen t-BSA—are arranged in a regular grid on the sensor surface, identifying 1035 regions of interest (ROIs). In practice, some spots may not exist or contain minimal reagent concentration. Note that visually detecting the presence of a spot is straightforward, as it appears darker than the background in the reflectometric image. Figure 5 reports a schematic representation of the dataset.

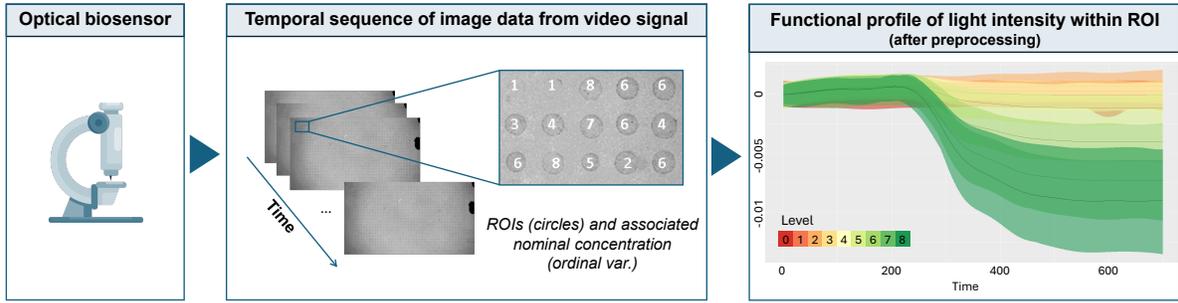


Figure 5: Biosensor data, a zoom on a subset of ROIs labeled with a nominal concentration level. On the right, the functional boxplot of the smoothed log-ratio, stratified by concentration level. Coloring of the boxplots is consistent with that of the associated levels (on the right). Note that levels 0 denotes the control group, only levels $c_i > 1$ display a spot on the image, and only levels $c_i > 2$ presents a positive antigen concentration.

Preprocessing of the video signal consists of light cleaning and subsequent construction of the ordinal-functional dataset, where the statistical unit is the ROI and the target variables are (i) a functional data point that represents the temporal evolution of the light intensity and (ii) an ordinal variable that represents the level of concentration of reagent.

For each ROI, the nominal concentration level is associated with an ordinal label. In order to detect magnitude outliers among the ROIs, we eliminate exogenous sources of light from the reflectometric images, such as external lights and irregularities in the surface coating. In our case study, it is realistic to assume that the exogenous sources of light are constant over time. We model the logarithm of light intensity at time 0 as a Generalized Additive Model (GAM, Hastie and Tibshirani (1986)), using as predictor the pixels's spatial coordinates (x, y) , with a nonlinear contribution, and the categorical information of whether or not the pixel belongs to a ROI. By dividing the light intensity by the model-based expected light disturbance (or by subtracting in the case of their logarithms) on the reflectometric images, we effectively minimize variations in light that do not stem from the presence of the antigen spot. Net of the light cleaning, we can easily check if a pixel belongs to random spots or determine if a randomly sampled pixel is located within or outside an ROI.

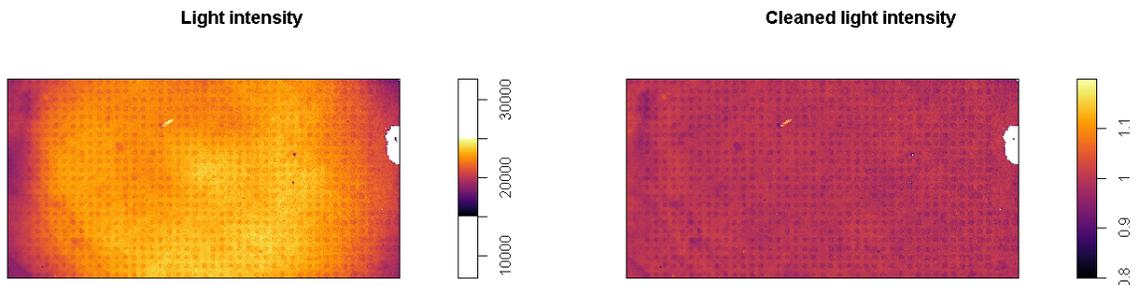


Figure 6: Light intensity vs cleaned light intensity at time 0

The light intensity, ranging from 0 to $2^{16}-1$, exhibits a compositional nature. To conduct standard statistical analyses like dimensionality reduction, modeling, or clustering, it is crucial to work with absolute values, rather than the relative information associated with the compositional nature of the data (Pawlowsky-Glahn et. al (2015)). Various transformations for compositional data have been proposed in the literature to ensure meaningful and interpretable results. Among these, the log-ratio approach, initially introduced by Aitchison (1982), stands out as one of the most popular methodologies. Following these principles, we construct a functional datum representing the light intensity of the i -th ROI as follows. We subtract the GAM model from the logarithm of light intensity, obtaining a cleaned, time-varying log-light intensity, $f_j(t)$, for each pixel j in the image D . For each ROI i , we compute the 5%-trimmed mean of the cleaned signals $f_j(t)$ corresponding to $N = 100$ randomly extracted pixels within the ROI, resulting in $\bar{f}_i^{ROI}(t)$. The trimming is based on the Modified Band depth, introduced by Lopez-Pintado and Romo (2009). This ensures robustness in estimating the mean function within the ROI, which is then taken as representative of the signal for that ROI. Additionally, we sample $N = 100$ pixels in the neighborhood of each ROI i and compute the 5%-trimmed mean $\bar{f}_i^{BG}(t)$ to robustly estimate the mean in the background area adjacent to ROI i . After conducting a sensitivity analysis regarding the quantity of points sampled within and outside the region of interest, it became apparent that selecting more than 100 points does not result in a noticeable enhancement. To account for the compositional nature of the signals, we compute the log-ratio of the light intensity:

$$x_i(t) = (\bar{f}_i^{ROI}(t) - \bar{f}_i^{ROI}(0)) - (\bar{f}_i^{BG}(t) - \bar{f}_i^{BG}(0)) . \quad (11)$$

We then smooth the log-ratio (11) using smoothing splines regression (Reinsch (1967)), employing 20 cubic splines with equally distributed knots and a penalty parameter of 10^4 selected via GCV to obtain a meaningful functional profile. In conclusion, each ROI i is associated with a functional datum $x_i(t)$ and a concentration level c_i of the reagent. Figure 5 displays (on the right) the functional boxplot of the log-ratio of the light intensity, colored according to the associated level of antigen concentration. The final dataset consists of the logratio profiles dataset deprived of the level-specific functional outliers, recognised through the Modified Band Depth (Lopez-Pintado and Romo (2009)).

5.3 Dimensionality reduction

In this section, we show the results obtained by applying the method foCCA, presented in Section 3, to the preprocessed dataset introduced in Section 5.2. From a biophysical standpoint, levels 0 and 1 do not have intrinsic distinctions, and our interest does not lie in differentiating between them. However, in general, the difference between consecutive levels is not known a priori. For this reason, at the training phase (including the training steps in the iterations of each K-fold CV), we did not merge these two classes and thus challenged the methods to select the best components for ordinal prediction when the differences in levels are extremely uneven. We merge them into a single level, denoted as 1, only to display and evaluate the confusion matrices. We select the penalty parameters, among a grid of proposed parameters, via K-fold Cross Validation, as detailed in Section 3.4, with $K = 5$. The selected couple of parameters for foCCA are $(\lambda_1, \lambda_2) = (100, 1000)$. For the analysis here presented, the dataset cleaned from outliers was used; hence, hereafter we do not consider robust estimators for the mean and the covariance operators of the quantity involved. Alternatively, one may use one of the robust estimation procedures present in the literature (see, e.g., Bali et. al (2011)).

The first row of Figure 7 displays the scatterplot of the first two functional scores of foCCA, alongside the marginal densities of these scores. Notably, the marginal densities of the first score

distinctly separate the highest concentration levels, while the lowest four levels remain largely overlapped. The second component effectively distinguishes levels $\{0, 1\}$ from levels $\{2, 3, 4\}$, i.e. ROIs without spots from those with spots and a low antigen concentration.

The method foCCA also provides a clear understanding of the difference, with respect to the most relevant features of the functional profiles, among the levels. Indeed, by observing the first ordinal components (Figure 8, right) combined with their respective functional components (Figure 8, left), we can understand, for every couple of consecutive levels, how much they are separated by the corresponding functional score. In addition, we can determine whether moving to the next level results in an increase or a decline of the functional score. In Figure 9 it is clear that the first functional component gives a high weight to the final phase of the reaction (time $t \geq 400$). The vector θ_1 , i.e. the first ordinal component takes high values in the last components, in particular between levels 5 and 6 (see the right plot in Figure 8), where θ_1 has the maximum value. All these considerations are consistent with the fact that the highest levels are well separated in the 2D functional scores' space (Figure 7, top) and the levels 5 and 6 are here the most distant. On the contrary, the second score seems to separate the lowest levels (spot vs. non-spot); this observation is coherent with the information given by the right plot in Figure 8, where the highest component of θ_2 is 1:2, i.e. the increment on the ordinal score passing from level 1 to level 2 (see Section 3.3 to interpret the ordinal component). Lastly, in Figure 8, one may notice that the second functional component gives a negative weight in the initial slot (approximately $t < 60$): this first phase, in practice, is fundamental to discriminate between levels 1 and 2, since the initial light intensity variation is due to the presence of the spot and the lower the concentration of antigen (null at level 2) the higher the power of this variation. In conclusion, foCCA is able to provide further analysis of the levels paired with the functional profiles and, in our case study, it is consistent with the biological knowledge.

5.4 Validation

We now validate our method by comparing the dimensionality reduction obtained through foCCA with the ones obtained with fPCA, foFD, focusing on the ability to separate the consecutive levels of concentration. Moreover, we will compare the nearest centroid classifiers built on the first two components identified by foCCA, fPCA and foFD. For the sake of completeness, we shall also compare the results with a scalar competitor, built upon considering as score the last value of the log-ratio. The latter is a heuristic alternative to the first three methods, since, at a first sight, it can seem sufficient to predict the level of concentration. In order to have a fair comparison among foCCA, fPCA and foFD, we select the penalty parameters at the best of their ordinal classification power (analogously as in Section 3.4). The parameter selected for fPCA is $\lambda = 10^6$, while for foFD is $\lambda = 10^8$. Figure 7 shows that, as foCCA, fPCA and foFD provide a first score that distinctly separate the highest concentration levels. However, the second component of fPCA, contrarily to foCCA, does not exhibit clear separation between any pair of successive levels, while the second component of foFD separates the consecutive levels 1:2, but in a much weaker way than foCCA. By observing the confusion matrices in Figure 12, it is clear that only foCCA has a high accuracy in the low levels of concentration. Moreover, fPCA, foFD and, even more, the heuristic method, show frequent misclassifications of 2 or more levels, especially when the target level is low, while in foCCA it happens much more rarely.

A possible interpretation of these results can be given. Indeed, the functional direction along which the level 1 (i.e. 0 and 1) and the levels higher than 1 are most separated do not present a high global variance. For this reason, fPCA does not identify this direction as relevant. Moreover, the levels are not intrinsically equidistant, hence foFD tends to find the functional directions that maximize the separation among the most different successive levels. Contrarily, foCCA, through the use of

vector θ , naturally handles the varying distances of consecutive levels, approaching or moving away from levels to maximize correlation with the functional component. Lastly, the heuristic method, which would provide a fast and simple alternative to the functional methodologies, suggests that a scalar index does not contain enough information to distinguish all the levels. In conclusion, foCCA outperforms all the tested competitors in predicting the level of concentration of antigen, thanks to the flexibility gained with the ordinal canonical component θ , and the targeting of the ordinal variable in its objective functional.

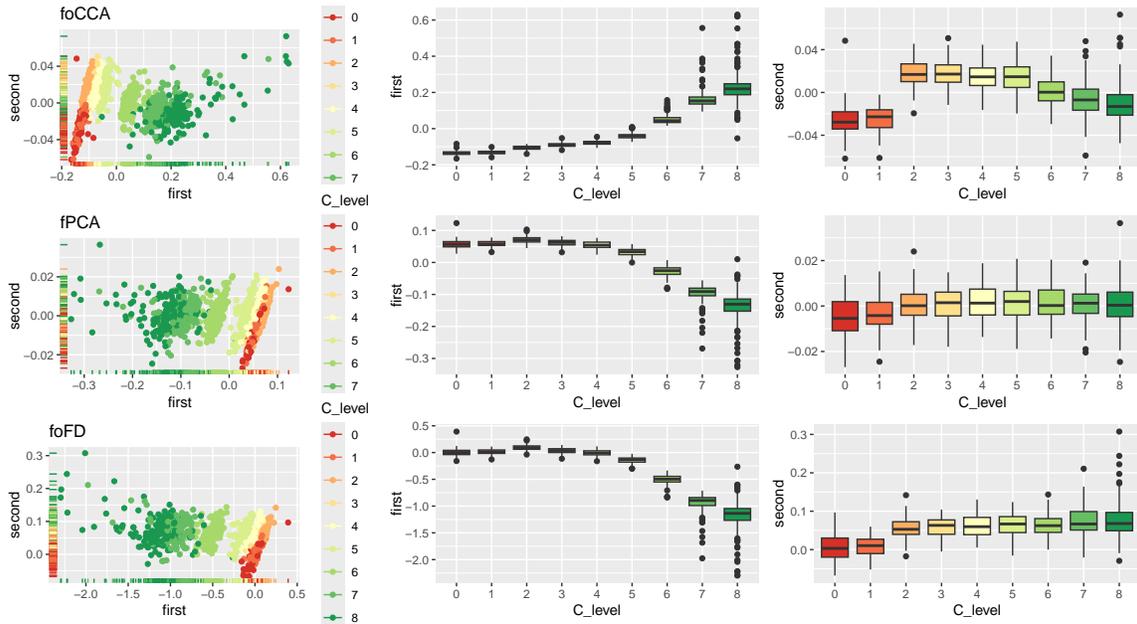


Figure 7: A comparison of the three methods. First column: the scatterplots of the first two scores. Second column: boxplots of the first score. Third column: boxplots of the second score.

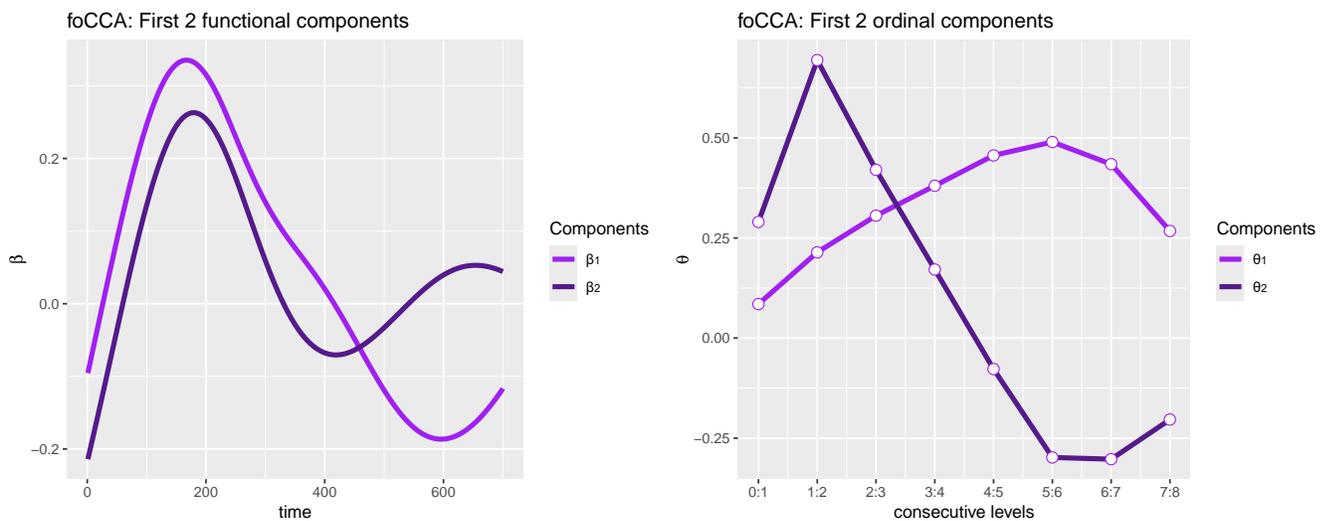


Figure 8: The first two components of foCCA

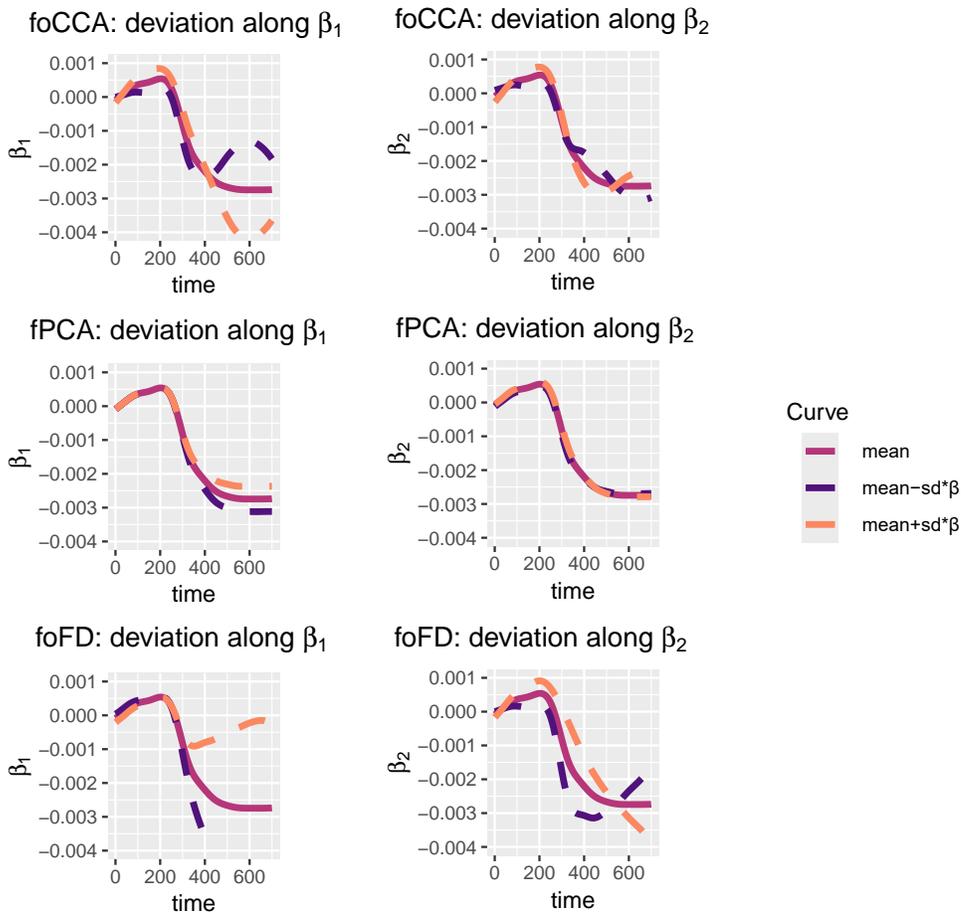


Figure 9: Deviations along the first two functional components

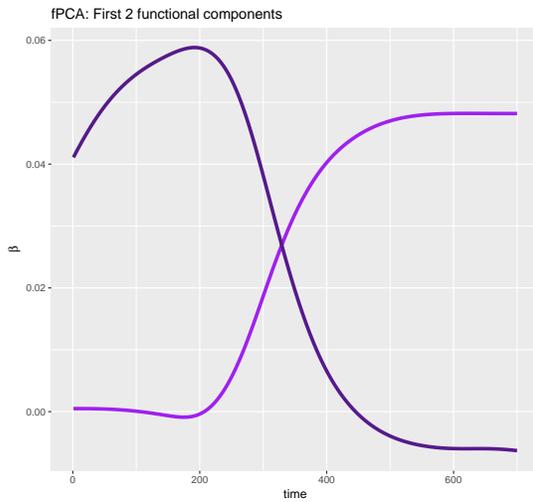


Figure 10: The first two components of fPCA

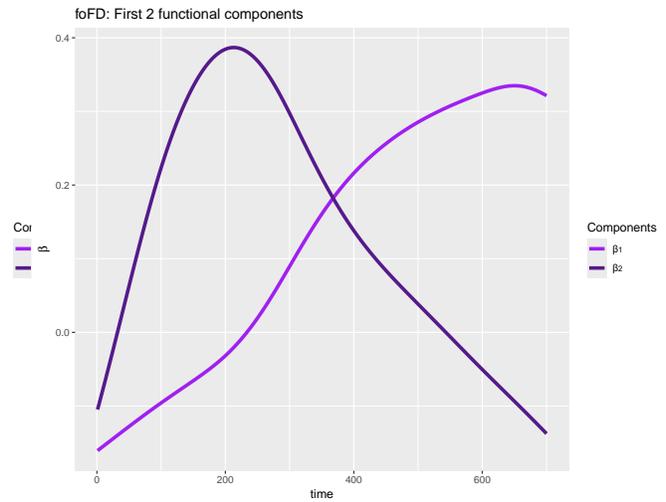


Figure 11: The first two components of foFD

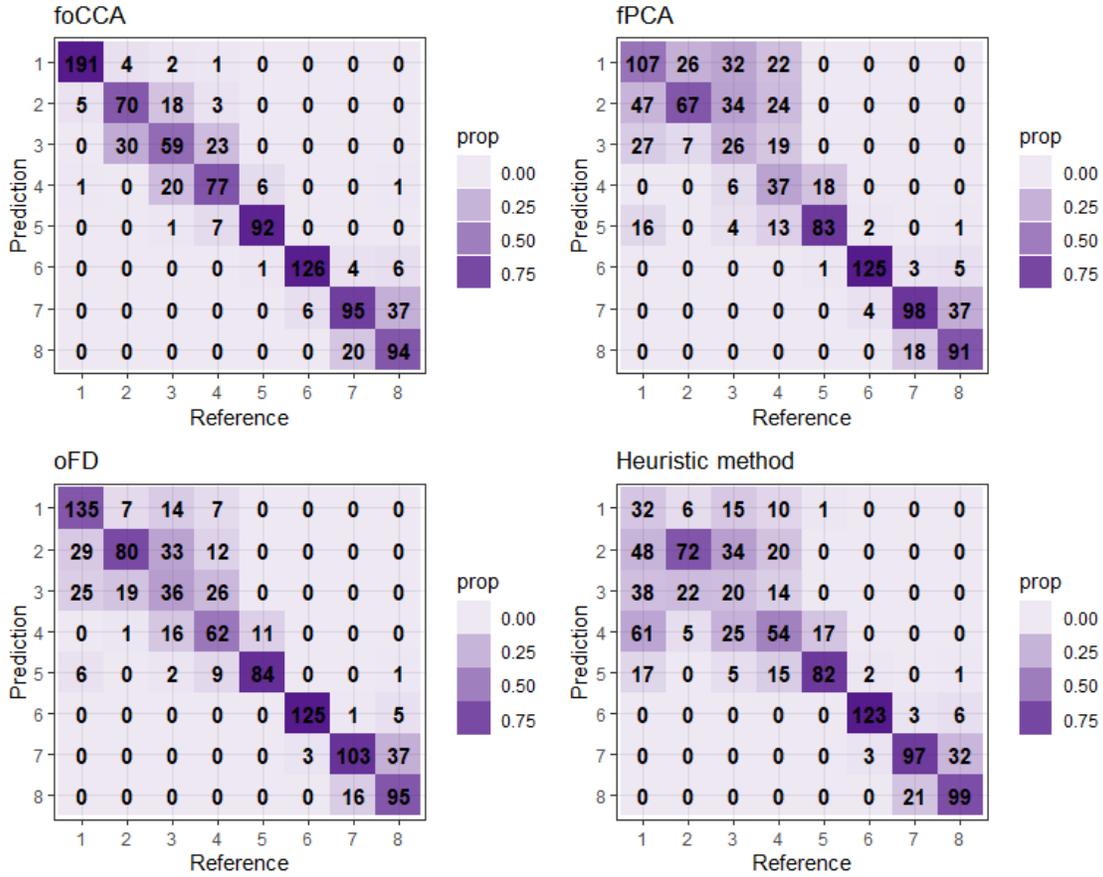


Figure 12: Comparison of the K-fold CV confusion matrices, with $K = 5$



Figure 13: Comparison of sensitivity and specificity

6 Conclusions

The work presented is inspired by a case study, when we want to maximize the ability to predict the antigen concentration level of each ROI, based on an optical signal. At this aim, we proposed a novel methodology to predict ordinal variables based functional data, foCCA. Through an extensive simulation study, we showed that foCCA performs significantly better, from an ordinal predictive point of view, than the competitors (i.e. fPCA and foFD) in two realistic scenarios for any level of severity of both. In the case study, foCCA gives better ordinal predictive results than competitors, both functional (fPCA and foFD) and scalar (heuristic method) and, additionally, it provides signal-based distances between the consecutive levels, which in principle are not homogeneous. Moreover, the results are consistent with the physical knowledge about the ROIs. Our approach is particularly suitable when the final aim is ordinal prediction based on high dimensional data or when the signal-based differences between consecutive levels are considered of particular interest. As future improvement, the foCCA can be made able to predict in real-time the ordinal level. Indeed, in process control an early predicted level can be fundamental to take real-time actions, e.g. with the aim in view of building a digital twin.

Acknowledgements

The authors gratefully thank Johanna Hutterer (University of Tübingen) for the interesting discussions about the biophysical consistency of the results and her contribution in the editing of the article. This research has received funding by the European Commission under the “HORIZON-CL4-2021-DIGITALEMERGING-01 project BioProS - Biointelligent Production Sensor to Measure Viral Activity” (grant agreement no. 101070120), 2022-2026”. GP, FN, LD, and AM acknowledge the initiative “Dipartimento di Eccellenza 2023–2027”, MUR, Italy, Dipartimento di Matematica, Politecnico di Milano. LD acknowledges his membership to INdAM GNCS - Gruppo Nazionale per il Calcolo Scientifico (National Group for Scientific Computing), Italy.

References

- Aitchison J. (1982). The Statistical Analysis of Compositional Data. *Monographs on Statistics and Applied Probability*. Springer.
- Akturk B., Beyaztas U., Shang H. L., Mandal A. (2024). Robust functional linear regression. *Advances in Data Analysis and Classification*.
- Bali J. L., Boente G., Tyler D. E., Wang J. L.(2011). Robust Functional Principal Components: a projection-pursuit approach. *The Annals of Statistics*, 39(6), 2852—2882.
- Beretta D., Grillo S., Pigoli D., Bionda E., Bossi C., Tornelli C. Hullait H., Leslie D. S., Pvalidis N. G.and King S. (2020). Functional principal component analysis as a versatile technique to understand and predict the electric consumption patterns. *Sustainable Energy, Grids and Networks*,21, 100308.
- Bugatti, M. & Colosimo B. M. (2022) Towards real-time in-situ monitoring of hot-spot defects in L-PBF: A new classification-based method for fast video-imaging data analysis. *Journal of Intelligent Manufacturing* , 33(1), 293—309.
- Cooper M. A. (2006). Optical biosensors: Where next and how soon? *Drug Discovery Today* 11(23–24), 1061–1067.

- Fisher R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics* 7, 179–188.
- Gauglitz G. (2010). Direct optical detection in bioanalysis: an update. *Anal Bioanal Chem* 389, 2363–2372.
- Hastie T., Buja A., Tibshirani R. (1995). Penalized discriminant analysis. *Annals of Statistics* 23, 73–102.
- Hastie T. and Tibshirani R. (1986). Generalized Additive Models. *Statistical Science* 1 (3), 297–310.
- Hoerl A. E., Kennard R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* 12(1), 55–67.
- Hotelling H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24, 417–441.
- Hullait H., Leslie D. S., Pvalidis N. G. and King S. (2020). Robust Function-on-Function Regression. *Technometrics*, 63(3), 396–409.
- Johnson R. A. and Wichern D. W. (2015). Applied Multivariate Statistical Analysis, 6th Edition. *Pearson, India*
- Kundu P., Darpe A. K., Kulkarni M. S., and Zuo M. (2022). Online damage severity level classification in gears under natural damage progression. *The International Journal of Advanced Manufacturing Technology*
- Leurgans S. E., Moyeed R. A., Silverman B. W. (1993). Canonical Correlation Analysis when the Data are Curves. *Journal of the Royal Statistical Society. Series B (Methodological)* 55(3), 725–740.
- Li Y., Li B., Ding J., Feng Y., Ma M., Han Z., Xu Y., and Xia L. (2023). A Novel Framework for Forecasting Mental Stress Levels Based on Physiological Signals. *Neural Information Processing*, 2023, 123–138.
- Long F., Zhu A., Shi H. (2013). Recent advances in optical biosensors for environmental monitoring and early warning. *Sensors* 13(10), 13928–13948.
- Lopez-Pintado S. and Romo J. (2009). On the Concept of Depth for Functional Data. *Journal of the American Statistical Association*, 104(486), 718–734.
- McFee B., Ellis D. P. W. (2014). Learning to segment songs with ordinal linear discriminant analysis *International conference on acoustics, speech and signal processing (ICASSP)*.
- Movahed R. A., Jahromi G. P., Shahyad S., and Meftahi G. H. (2021). A Novel Framework for Classifying Major Depressive Disorder Using EEG Signals. *Journal of Neuroscience Methods*, 2021.
- Müller H. G. (2005). Functional Modelling and Classification of Longitudinal Data *Scandinavian Journal of Statistics*, 32(2), 223–240.
- Narlawar S. S., Sharma R., Sahoo J., Kaushik A., Gandhi S. (2024). Navigating the landscape of optical biosensors. *Chemical Engineering Journal* 490, 151661.

- Narsaiah K., Jha S. N., Bhardwaj R., et al. (2012). Optical biosensors for food quality and safety assurance—a review. *Journal of Food Science and Technology* 49, 383–406.
- Pawlowsky-Glahn V., Egozcue J. J., Tolosana-Delgado R.(2015). Modeling and Analysis of Compositional Data. *Statistics in Practice, John Wiley & Sons*.
- Ramsay J. O., Silverman B. W. (2005). Functional Data Analysis. *Springer, New York*
- Ruan H., Dai X., Chen S., and Qiu X. (2022). Arrhythmia Classification and Diagnosis Based on ECG Signal: A Multi-Domain Collaborative Analysis and Decision Approach. *Electronics*, 11(19):3251.
- Reinsch C.H. (1967). Smoothing by spline functions. *Numerical Mathematics* 10, 177–183.
- Sharma A., Mishra R. K., Goud K. Y., Mohamed M. A., Kummari S., Tiwari S., Li Z., Narayan R., Stanciu L. A., Marty J. L. (2021). Optical biosensors for diagnostics of infectious viral disease: A recent update. *Diagnostics* 11(11), 2083.
- Sharma K., Sharma M. (2023). Optical biosensors for environmental monitoring: Recent advances and future perspectives in bacterial detection. *Environmental Research* 236, Part 2, 116826.
- Silverman B. W. (1996). Smoothed functional principal components analysis by choice of norm. *Annals of Statistics* 24, 1–24.
- Wood S. N. (2004). Stable and Efficient Multiple Smoothing Parameter Estimation for Generalized Additive Models. *Journal of the American Statistical Association* 99, 673–686.
- Wood S. N., Goude Y., Shaw S. (2015). Generalized additive models for large datasets. *Journal of the Royal Statistical Society, Series C* 64(1), 139–155.
- Yan, H., Grasso, M., Paynabar, K., & Colosimo, B. M. (2022). Real-time detection of clustered events in video-imaging data with applications to additive manufacturing. *IISE Transactions*, 54(5), 464-480.
- Zhou R., Serban N., and Gebraeel N. (2011). Degradation Modeling Applied to Residual Lifetime Prediction Using Functional Data Analysis. *Annals of Applied Statistics*, 5(2B), 1586–1610.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

- 06/2025** Torzoni, M.; Manzoni, A.; Mariani, A.
Enhancing Bayesian model updating in structural health monitoring via learnable mappings
- 05/2025** Buchwald, S.; Ciaramella, G.; Verani, M.
Greedy reconstruction algorithms for function approximation
- 04/2025** Andrini, D.; Magri, M.; Ciarletta, P.
Nonlinear morphoelastic theory of biological shallow shells with initial stress
- 02/2025** Corda, A.; Pagani, S.; Vergara, C.
Influence of patient-specific acute myocardial ischemia maps on arrhythmogenesis: a computational study
- 01/2025** Dede', L.; Parolini, N.; Quarteroni, A.; Villani, G.; Ziarelli, G.
SEIHRDV: a multi-age multi-group epidemiological model and its validation on the COVID-19 epidemics in Italy
- 109/2024** Liverotti, L.; Ferro, N.; Matteucci, M.; Perotto, S.
A PCA and mesh adaptation-based format for high compression of Earth Observation optical data with applications in agriculture
- 110/2024** Pederzoli, V.; Corti, M.; Riccobelli, D.; Antonietti, P.F.
A coupled mathematical and numerical model for protein spreading and tissue atrophy, applied to Alzheimer's disease
- 108/2024** Arostica, R.; Nolte, D.; Brown, A.; Gebauer, A.; Karabelas, E.; Jilberto, J.; Salvador, M.; Bucelli, M.; Piersanti, R.; Osouli, K.; Augustin, C.; Finsberg, H.; Shi, L.; Hirschvogel, M.; Pfaller, M.; Africa, P.C.; Gsell, M.; Marsden, A.; Nordsletten, D.; Regazzoni, F.; Plank, G.; Sundnes, J.; Dede', L.; Peirlinck, M.; Vedula, V.; Wall, W.; Bertoglio, C.
A software benchmark for cardiac elastodynamics
- 107/2024** Chen, J.; Ballini, E.; Micheletti, S.
Active Flow Control for Bluff Body under High Reynolds Number Turbulent Flow Conditions Using Deep Reinforcement Learning
- 106/2024** Brunati, S.; Bucelli, M.; Piersanti, R.; Dede', L.; Vergara, C.
Coupled Eikonal problems to model cardiac reentries in Purkinje network and myocardium