

Integrated Depths for Partially Observed Functional Data

Antonio Elías,
OASYS Group

Department of Applied Mathematics, Universidad de Málaga

Raúl Jiménez,

Department of Statistics, University Carlos III of Madrid

Anna M. Paganoni and Laura M. Sangalli

MOX Laboratory for Modeling and Scientific Computing

Dipartimento di Matematica, Politecnico di Milano

Abstract

Partially observed functional data are frequently encountered in applications and are the object of an increasing interest by the literature. We here address the problem of measuring the centrality of a datum in a partially observed functional sample. We propose an integrated functional depth for partially observed functional data, dealing with the very challenging case where partial observability can occur systematically on any observation of the functional dataset. In particular, differently from many techniques for partially observed functional data, we do not request that some functional datum is fully observed, nor we require that a common domain exist, where all of the functional data are recorded. Because of this, our proposal can also be used in those frequent situations where reconstructions methods and other techniques for partially observed functional data are inapplicable. By means of simulation studies, we demonstrate the very good performances the proposed depth on finite samples. Our proposal enables the use of benchmark methods based on depths, originally introduced for fully observed data, in the case of partially observed functional data. This include the functional boxplot, the outliergram and the depth vs depth classifiers. We illustrate our proposal on two case studies, the first concerning a problem of outlier detection in German electricity supply functions, the second regarding a classification problem with data obtained from medical imaging.

Keywords: Incomplete functional data, functional depth, robustness, functional outliers, functional boxplot, classification of partially observed functional data.

1 Introduction

In this paper we propose a depth measure for partially observed functional data.

Starting from the pioneering work of [Fraiman and Muniz \(2001\)](#), different definitions of depths for functional data have been proposed (see, e.g., [Fraiman and Muniz, 2001](#); [López-Pintado and Romo, 2009, 2011](#); [Ieva and Paganoni, 2013](#); [Claeskens et al., 2014](#)) and their theoretical properties have been thoroughly studied (see, e.g., [Nieto-Reyes and Battey, 2016](#); [Nagy et al., 2016](#); [Gijbels and Nagy, 2017](#)). Moreover, depths have shown to be remarkable tools for the visualization of functional data ([Hyndman and Shang, 2010](#); [Sun and Genton, 2011, 2012](#); [Genton et al., 2014](#); [Dai and Genton, 2018](#)), for functional outlier detection ([Arribas-Gil and Romo, 2014](#); [Nagy et al., 2017](#); [Ieva and Paganoni, 2017](#)) and for the classification of functional data ([Li et al., 2012](#); [Hubert et al., 2017](#)). On the other hand, the presence of partially observed data hampers the applicability of such methods.

For a random sample of functions from a compact interval $[a, b]$ to \mathbb{R} , partially observed data refers to the case where the records of the functions are available only on subsets of $[a, b]$, so that only fragments of each functional datum are available. This type of data (also referred to as incomplete or fragmented functional data, as well as functional snippets) are indeed common in real applications, and have been reported in several areas of research. In medical studies, for instance, typical sources of censoring of the data are patients missing medical visits or devices failing to record (see, e.g., [James et al., 2000](#); [James and Hastie, 2001](#); [Sangalli et al., 2009](#); [Delaigle and Hall, 2013](#); [Kraus, 2015](#); [Delaigle and Hall, 2016](#); [Lin et al., 2020](#)). In demography it is common that age-specific mortality rates for older ages are not completely observed due the too low number of survivors (see, e.g., [Human Mortality Database, 2019](#); [D'Amato et al., 2011](#)). In electricity markets, supply functions are incomplete because suppliers and buyers agree prices and quantities depending on the market conditions (see, e.g., [Kneip and Liebl, 2020](#); [Liebl and Rameseder, 2019](#)).

Likewise in the case of sparse functional data, where the datum is assumed to be recorded in a sparse way across the whole domain, also in the case of partially observed functional data, the missing data structure raises difficulties for the data analysis. This has stimulated a very active literature, and sparse or partial observability in functional data have been addressed in several respects, including functional principal component analysis (see, e.g., [James et al., 2000](#); [Yao et al.,](#)

2005; Di et al., 2014; Liu et al., 2017), mean and covariance estimation (see, e.g., Kraus, 2015; Liebl and Rameseder, 2019; Lin et al., 2020), estimation of the missing parts (see, e.g., Goldberg et al., 2014; Kraus, 2015; Delaigle and Hall, 2016; Kneip and Liebl, 2020) and supervised and unsupervised classification (see, e.g., James and Hastie, 2001; Delaigle and Hall, 2013; Stefanucci et al., 2018; Kraus and Stefanucci, 2019).

The introduction of a suitable notion of depth measure for partially observed functional data contributes to the literature in two ways. First of all, it expands the palette of tools that are available in such complex data setting, introducing a key concept of robust statistics. Moreover, it offers a crucial support to the other techniques for partially observed functional data cited before. In fact, such inferential techniques can be highly affected by the presence of outliers. Up to this moment, in partially observed functional data settings, authors had to rely on visual inspections or expert knowledge for detecting outliers and avoiding biased estimations (see, e.g., Kneip and Liebl, 2020; Liebl and Rameseder, 2019); even so, only obvious magnitude outliers can be visually detected (see, e.g., Hyndman and Shang, 2010; Sun and Genton, 2011; Arribas-Gil and Romo, 2014; Nagy et al., 2017). Depth measures are cornerstones for outliers detection and our proposal can thus provide an objective approach in this respect. In particular, the proposed depth measure enables the construction of the functional boxplot (Sun and Genton, 2011) and the outliergram (Arribas-Gil and Romo, 2014) in the context of partially observed functional data. This also offers crucial visualization tools for partially observed functional data. Moreover, the proposed depth permits the use of Depth vs Depth classifiers (Li et al., 2012; Cuesta-Albertos et al., 2017) for partially observed functional data, offering an important alternative to the classification methods so far proposed in this context.

It is important to point out that we here deal with the very challenging case where partial observability can occur systematically on any observation of the functional dataset. In particular, differently from many of the above cited techniques for partially observed functional data, we neither request that some functional datum is fully observed, nor we require that a common domain exist, where all of the functional data are recorded. Because of this, our proposal can also be used in those frequent situations where reconstructions methods and other techniques for partially observed functional data are inapplicable.

The class of depths we define belongs to the integrated depth measures family (Nagy et al.,

2016), that includes the Fraiman and Muniz Depth (Fraiman and Muniz, 2001) and the Modified Band Depth (López-Pintado and Romo, 2009). We account for the uncertainty related with the unobserved fragments by means of an appropriate weight function, that gives more weight in the computation of the depth to regions with high observational density and less weight to regions with low observational density. We name the proposed method Partially Observed Integrated Functional Depth (POIFD).

Sguera and López-Pintado (2020) have very recently proposed a depth measure for sparse functional data. Their proposal leverages on reconstruction of the sparsely observed functional data, using the approach in Goldsmith et al. (2013), and is in turn used by Qu (2022) to construct a boxplot for sparsely observed functional data.

Differently from Sguera and López-Pintado (2020), our proposed POIFD does not involve reconstruction of the data, using any of the reconstruction techniques for sparse data or partially observed functional data. Through extensive simulations studies, we show that the proposed POIFD works better than depth computed on reconstructed data, considering best-of-the-art reconstruction techniques for sparse functional data (such as Goldsmith et al., 2013) as well as for partially observed functional data (such as Kneip and Liebl, 2020; Liebl and Rameseder, 2019). These simulation studies are carried out under different missing data scenarios with different percentages of observability. Furthermore, our proposed POIFD can also be applied in all the situations where reconstruction methods are inefficient (for instance, in presence of outliers) or inapplicable.

In Section 2 we introduce the proposed depth for partially observed functional data and its sample version. For simplicity, we first describe the depth in the case of univariate functional data, and hence extend it to multivariate functional data.

In Section 3 we report simulation studies that highlight the very good performances of the proposed POIFD. In particular, we demonstrate the high agreement between the POIFD and the depth that could be computed if the functional data was fully observed. Moreover, considering the special case where the reconstruction techniques for partially observed functional data are applicable (i.e., when at least some of the curves are fully observed), we show that POIFD performs significantly better than the standard functional depths computed on the full domain after reconstruction of the data. Furthermore, we show that, in the special case where there is a common domain where all of the curves are recorded, the proposed POIFD performs significantly better

than the standard functional depths restricted to the common domain.

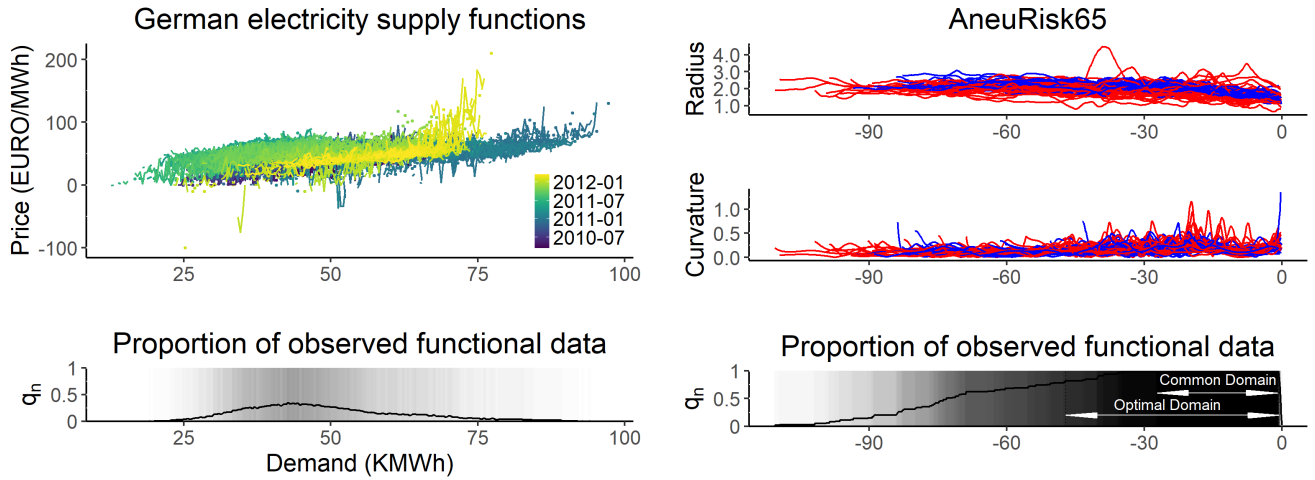


Figure 1: Left: German electricity supply functions. Top left: Daily electricity supply from March 15th, 2010 to March 14th, 2012; the y-axis represents the electricity price and the x-axis the level of the electricity demand. Bottom left: proportion of functional data that are observed for each given level of electricity demand. Right: AneuRisk65. Radius (top) and curvature (middle) of the internal carotid artery of 65 subjects; Upper group (blue in color printing / dark grey in black and white printing), lower group (red / light grey). The bottom panel shows the proportion of observed data, highlighting the presence of a common domain, where all data are observed.

In Section 4, we illustrate the use of the proposed POIFD in two challenging case studies. The first concerns outlier detection in German electricity supply functions (Kneip and Liebl, 2020; Liebl, 2019). The data are shown in the top left panel of Figure 1; the bottom left panel of the same figure displays the proportion of observed data at each point of the domain, highlighting that the missing portions are spread along all the domain. The second case study involves the AneuRisk65 dataset (see, e.g., Sangalli et al., 2009, 2014b) and has to do with a discrimination problem on data obtained from reconstruction of medical images. The data are shown in the top right panel of Figure 1. The bottom right panel of the same figure shows that for AneuRisk65 data there exists a small portion of the domain, named common domain, where all the functional data are observed.

The supplementary Material carries additional evidence of the good performances of the proposed POIFD. We there also consider the challenging scenario where none of the functional data has been recorded over the whole domain, in which case the reconstruction methods for partially observed functional data are inapplicable.

2 Depth for partially observed functional data

Following [Claeskens et al. \(2014\)](#), we define a functional depth starting from a finite dimensional (univariate) depth. Let \mathcal{P} denote the collection of all probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, where $\mathcal{B}(\mathbb{R})$ is the Borel σ -algebra on \mathbb{R} . Consider a function $D : \mathbb{R} \times \mathcal{P} \rightarrow [0, 1]$. We assume such function satisfies properties D_1 to D_7 from [Nagy et al. \(2016\)](#). Without loss of generality we consider functional data defined over the interval $[0, 1]$. In particular, hereinafter, $X : [0, 1] \rightarrow \mathbb{R}$ is a stochastic process with continuous trajectories, P is the law of X , and P_t is the marginal probability distribution of $X(t)$.

We recall the definition of integrated functional depth ([Claeskens et al., 2014](#); [Nagy et al., 2016](#)).

Definition 2.1. Integrated Functional Depth (IFD). Given a continuous function $x : [0, 1] \rightarrow \mathbb{R}$, a univariate depth measure D , and a weighting function $w : [0, 1] \rightarrow [0, \infty)$, with $\int_0^1 w(t)dt = 1$, the Integrated Functional Depth of x with respect to P is

$$\text{IFD}_w(x, P) = \int_0^1 D(x(t), P_t)w(t)dt.$$

2.1 Partially Observed Integrated Functional Depth

Let X_1, \dots, X_n be n independent realizations of X . We consider the case when the realizations X_1, \dots, X_n are only partially observed. To model such setting, similarly to [Delaigle and Hall \(2013\)](#), we consider a random observational mechanism Q , that generates compact subsets of $[0, 1]$, over which the functional data are observed. In particular, we assume Q generates compacts that consist of finite unions of closed intervals with strictly positive Lebesgue measure. Let \mathcal{O} be one such set generated by Q , and let $\mathcal{O}_1, \dots, \mathcal{O}_n$ be independent copies of \mathcal{O} . Then, for $1 \leq i \leq n$, the functional datum X_i , is only observed on \mathcal{O}_i . We assume that P and Q are independent, and that $(X_1, \mathcal{O}_1), \dots, (X_n, \mathcal{O}_n)$ are i.i.d. realizations from $P \times Q$. This assumption, termed Missing-Completely-at-Random, is standard in the literature of partially observed functional data (see, e.g., [Kraus, 2015](#); [Kneip and Liebl, 2020](#)).

Since X is only observed on a compact set \mathcal{O} , we define its depth by restricting an integrated functional depth to \mathcal{O} . To this aim, for $t \in [0, 1]$, let $Q(t) = \mathbb{P}(\mathcal{O} \ni t)$, i.e., the probability that the random set \mathcal{O} covers the point t . Without loss of generality, we assume that $Q(t) > 0$ for all

$t \in [0, 1]$. Moreover, we consider a bounded and a continuous function ϕ defined on $[0, 1]$ and such that $\int_{\mathcal{O}} \phi(Q(t))dt > 0$ almost surely; such function ϕ can for instance be the identity function on $[0, 1]$. We then define the following weighting function restricted to a compact set \mathcal{o} :

$$w_{\phi}(t|\mathcal{o}) = \frac{\phi(Q(t))}{\int_{\mathcal{O}} \phi(Q(t))dt}. \quad (1)$$

We now define the proposed Partially Observed Integrated Functional Depth for any continuous function $x : [0, 1] \rightarrow \mathbb{R}$ restricted to a compact set \mathcal{o} of $[0, 1]$.

Definition 2.2. Partially Observed Integrated Functional Depth (POIFD). The Partially Observed Integrated Functional Depth of (x, \mathcal{o}) with respect to $P \times Q$ is

$$\text{POIFD}((x, \mathcal{o}), P \times Q) = \int_{\mathcal{O}} D(x(t), P_t)w_{\phi}(t|\mathcal{o})dt,$$

where $w_{\phi}(t|\mathcal{o})$ is the weight function in [\(1\)](#).

If the data are fully observed, i.e., if $\mathbb{P}(\mathcal{O} \ni t) = 1$ for any $t \in [0, 1]$, then the proposed POIFD coincides with the IFD with constant weight function $w(t) = 1$ for all $t \in [0, 1]$.

2.2 Sample version

We now introduce the empirical version of POIFD, for its computation on finite samples. Denote by P_n the distribution that assigns mass $1/n$ to each sample curve X_1, \dots, X_n . Similarly, let Q_n be the distribution that assigns mass $1/n$ to each sample compact set $\mathcal{O}_1, \dots, \mathcal{O}_n$. Set $I(t) := \{1 \leq i \leq n : t \in \mathcal{O}_i\}$ and $q_n(t) = \#I(t)/n$. Finally, let $P_{t,n}$ denote the empirical distribution function of the univariate sample $\{X_i(t), i \in I(t)\}$.

We define the empirical version of POIFD using the plug-in approach. Namely,

$$\text{POIFD}((x, \mathcal{o}), P_n \times Q_n) = \int_{\mathcal{O}} D(x(t), P_{t,n})\phi(q_n(t))dt / \int_{\mathcal{O}} \phi(q_n(t))dt. \quad (2)$$

In practice, we only have access to discrete versions of the functional data X_i , on a discrete evaluation grid $\{t_1, \dots, t_T\}$, with $0 = t_1 < t_2 < \dots < t_T = 1$, which for simplicity we assume to

be common across data and equispaced¹. Note that for many t_ℓ , $\ell = 1, \dots, T$, such evaluation may indeed be missing, in conformity with the partially observed nature of the data considered in this work. We can then define the sample version of the POIFD by using a standard Riemann approximation.

Definition 2.3. Sample Partially Observed Integrated Functional Depth. Given the partially observed functional data $(X_1, \mathcal{O}_1), \dots, (X_n, \mathcal{O}_n)$, (partially) evaluated on the common grid $\{t_1, \dots, t_T\}$, with $0 = t_1 < t_2 < \dots < t_T = 1$, the sample version of the Partially Observed Integrated Functional Depth in Definition 2.2 is given by

$$\text{POIFD}_T((x, \mathcal{o}), P_n \times Q_n) = \sum_{t_\ell \in \mathcal{O}} D(x(t_\ell), P_{t_\ell, n}) \frac{\phi(q_n(t_\ell))}{\sum_{t_\ell \in \mathcal{O}} \phi(q_n(t_\ell))}. \quad (3)$$

Next section extends the proposed depth to multivariate functional data.

2.3 Multivariate Partially Observed Integrated Functional Depth

The partially observed integrated functional depth in Definition 2.2 can be straightforwardly extended to multivariate functional data. In this case, we consider a K -dimensional stochastic process $\mathbf{X} = (X^{(1)}, \dots, X^{(K)})$ with law $\mathbf{P} = (P^{(1)}, \dots, P^{(K)})$, where each coordinate $X^{(k)}$, $1 \leq k \leq K$, is a continuous function on $[0, 1]$, and $P^{(k)}$ is its law. We moreover consider a multivariate observational process $\mathbf{Q} = (Q^{(1)}, \dots, Q^{(K)})$, that generates the compact sets $\mathbf{O} = (\mathcal{O}^{(1)}, \dots, \mathcal{O}^{(K)})$ of $[0, 1]^K$, allowing for different domains of observation along the different components. We assume that \mathbf{P} and \mathbf{Q} are independent and that $(\mathbf{X}_1, \mathbf{O}_1), \dots, (\mathbf{X}_n, \mathbf{O}_n)$ are i.i.d. realizations from $\mathbf{P} \times \mathbf{Q}$. A multivariate POIFD, inspired by the proposal by Ieva and Paganoni (2013), can then be defined as a weighted average of K univariate POIFDs:

$$\text{POIFD}((\mathbf{X}, \mathbf{O}), \mathbf{P} \times \mathbf{Q}) = \sum_{k=1}^K \alpha_k \text{POIFD}((X^{(k)}, \mathcal{O}^{(k)}), P^{(k)} \times Q^{(k)}) \quad (4)$$

¹When data do not come on a common grid, they can be referred to a common grid in a preprocessing step, using standard techniques such as binning. It is important to note that, for each functional datum X_i , the possible evaluation on a common grid is only done over those portions of the domain where the datum is observed (that is, without extrapolation on missing portions using reconstruction techniques for sparse or partially observed data).

where the non-negative weights $\alpha_1, \dots, \alpha_K$ sum to 1. The choice of these weights is problem driven. This definition is very flexible. As shown for instance in Section 4.2, the selection of the weights by cross-validation, in a classification problem, enables the construction of an accurate classifier, and offers a posteriori valuable information for the interpretation of the results.

It is also possible to consider the special case where the observational process is univariate, and select a compact subset of $[0, 1]$ where \mathbf{X} is observed, so that all K components $\{X^{(1)}, \dots, X^{(K)}\}$ of \mathbf{X} are observed over the same portion of the domain. We thus assume that $(\mathbf{X}_1, \mathcal{O}_1), \dots, (\mathbf{X}_n, \mathcal{O}_n)$ are i.i.d. realizations from $\mathbf{P} \times Q$, where \mathbf{P} and Q are independent. In this special case, alternatively to the definition above, following the approach by Claeskens et al. (2014), we can consider a depth measure D on \mathbb{R}^K , and define the multivariate POIFD

$$\text{POIFD}((\mathbf{x}, \mathcal{o}), \mathbf{P} \times Q) = \int_{\mathcal{O}} D(\mathbf{x}(t), \mathbf{P}_t) w_{\phi}(t|\mathcal{o}) dt,$$

where \mathbf{P}_t is the marginal law of $\mathbf{X}(t)$.

3 Simulation studies

In this section we illustrate the performances of the proposed POIFD, showing its superiority to any alternative, when the alternatives are applicable. In the definition of the weight of the POIFD, we use $\phi(q) = q$. We consider both the partially observed version of the Fraiman and Muniz depth (FM) (Fraiman and Muniz, 2001) and the partially observed version of the Modified Band Depth (MBD) (López-Pintado and Romo, 2009). We also use the partially observed version of the Modified Epigraph Index (MEI) (López-Pintado and Romo, 2011) in the outliergram. The MEI for partially observed functional data is defined as in Definition 2.2, with $D(x(t), P_{t,n})$ replaced by $1 - F_{t,n}(x(t))$, where $F_{t,n}$ is the empirical cumulative distribution function of the univariate sample $\{X_i(t), i \in I(t)\}$.

3.1 Data generation

We generate data as follows. We let P be the law of a Gaussian process. In particular, we sample from the model $X(t) = \mu(t) + \epsilon(t)$ where $\mu(t)$ is a centered Gaussian process with co-

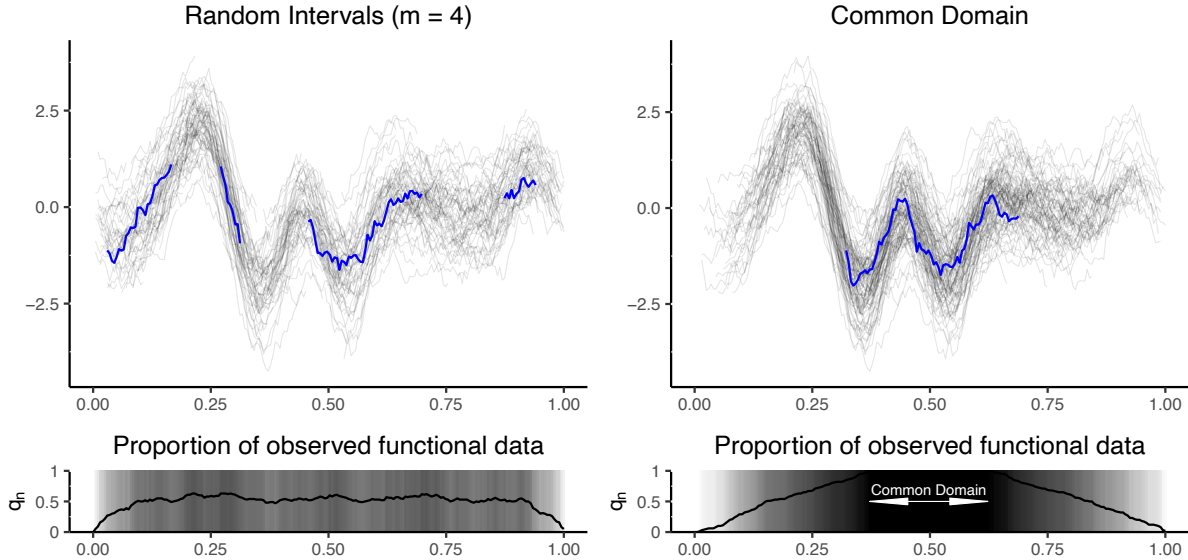


Figure 2: Simulated data. Data generated as detailed in Section [3.1](#), under two different observability patterns: the Random Intervals case (with $m = 4$ intervals) on the left; the Common Domain case on the right. Each functional datum is observed over an (expected) proportion of the full domain equal to $p = 50\%$. In each graph one random datum is plotted with a thick blue line (dark grey in black and white printing), to highlight the observability pattern. The bottom panels show the proportion of observed functional data at each time point of the domain.

variance $\rho_\mu(s, t) = \beta e^{-\gamma(\sin(\pi|s-t|)/\zeta)^2}$, for any $s, t \in [0, 1]$, and $\epsilon(t)$ is a centered Gaussian process with covariance $\rho_\epsilon(s, t) = \eta e^{-\lambda|s-t|}$. In particular, when performing N simulation repetitions, we generate N samples of dimension n , where the sample is $\{X_{j1}, \dots, X_{jn}\}$, for $j = 1, \dots, N$, and $X_{ji} = \mu_j(t) + \epsilon_{ji}(t)$. With this generating process, each random sample presents a different mean pattern μ_j , thus avoiding shape related biases. The set of parameters for all the results presented in the article is $\beta = 3$, $\gamma = 2$, $\zeta = 0.5$, $\eta = 0.5$, and $\lambda = 5$. We carried out simulations for other values of these parameters (non included here for sake of space), always reaching the same conclusions.

We consider an equispaced sampling grid $\{t_1, \dots, t_T\}$ of $T = 201$ points on $[0, 1]$.

As for the observation process Q , we consider the following patterns of observability, where p is the (expected) proportion of observability over the evaluation grid.

Sparse functional data: Each function is observed at $\lfloor p \times T \rfloor$ grid points randomly taken from $\{t_1, \dots, t_T\}$.

Partially observed functional data on random intervals: Each function is observed

on m disjoint intervals, spread along $[0, 1]$, with a total expected observed proportion p . To do so, for each function, we generate a random sample of size $\lfloor (m - p)/p \rfloor$ from a uniform distribution on $[0, 1]$. Then, we consider the intervals $[u(i - 1), u(i)]$, where $u(i)$ is the i -th order statistics of the sample. The function is hence observed on m of these intervals, chosen at random, but guaranteeing that they are disjoint.

Partially observed functional data with common domain: All functions are completely observed on a subset of the domain, named common domain, that is for simplicity centered in 0.5. This common domain is $\cap_{i=1}^n \mathcal{O}_i$. Specifically, each functional datum is observed on a domain that goes from a starting point to an ending point that are randomly generated. The starting point is sampled from a Uniform distribution on $[1/2 - p, 1/2)$, if $p \leq 1/2$, and from Uniform distribution on $[0, 1 - p)$, if $p > 1/2$. The ending point is sampled from a Uniform distribution on $(1/2, 1/2 + p]$, if $p \leq 1/2$, and from Uniform distribution on $(p, 1]$, if $p > 1/2$. This leads to a total (expected) proportion of observation p .

Figure 2 shows the same realizations of P , but partially observed under the two settings of Q , random interval and common domain, with $p = 50\%$. In both panels of the figure, one random functional datum is plotted in blue. The bottom panel displays the proportion q_n of observed functions.

3.2 Comparison with competing methods

We compare the proposed POIFD to standard depths computed on data that have been reconstructed using some of the most recent reconstruction methods for sparse and partially observed functional data, including Goldsmith et al. (2013), Kraus (2015) and Kneip and Liebl (2020). The method in Goldsmith et al. (2013) also constitutes the base for the depth for sparse functional data very recently proposed in Sguera and López-Pintado (2020), which we do not include here because of the unavailability of a publicly available code for this method. Our proposed POIFD is for partially observed functional data; nevertheless, its empirical version can also be computed on sparse functional data. For this reason, we also include the case of sparse functional data in our simulation studies.

It is important to point out that the methods for partially observed data, such as Kraus (2015)

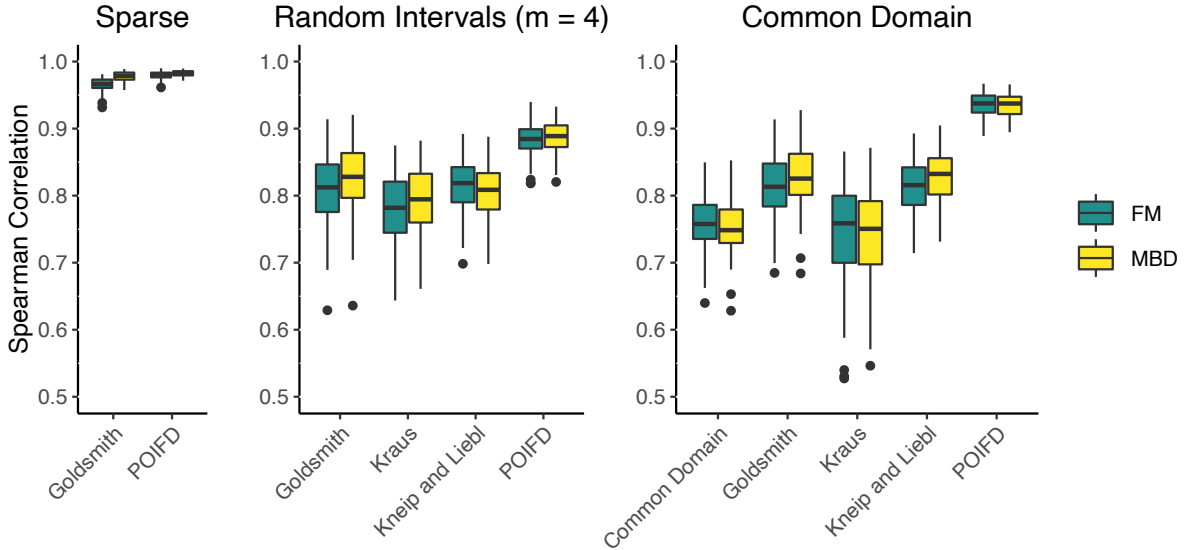


Figure 3: Spearman correlation between depths on completely observed sample and various alternatives to compute depths on partially observed data. The alternatives include: computation of depths on the common domain (in the Common Domain case); computation of depths after reconstruction with Goldsmith et al. (2013), with Kneip and Liebl (2020) and with Kraus (2015); the proposed POIFD. The results are obtained considering the Fraiman and Muniz depth (FM) and the Modified Band Depth (MBD). In each setting, to enable comparison with depth computed on reconstructed data, 25% of functional data in the sample are completely observed, while for the remaining 75% there is an average observation proportion $p = 50\%$ for each functional datum.

and Kneip and Liebl (2020), can estimate consistently the covariance only when there exist functions completely observed over the total domain $[0, 1]$. For this reason, we here let some of the data be observed on the whole evaluation grid. In Section 7.2 of the Supplementary Material, we instead show that our proposed POIFD has very good performances also in the more challenging scenario where no datum is fully observed.

Specifically, we generate $N = 100$ samples, each consisting of $n = 100$ functional data X_{j1}, \dots, X_{jn} , obtained from the process P described above. For each sample, we let 25% randomly chosen functional data to be observed over the whole evaluation grid. The remaining 75% of the functional data are only partially observed, sampling the observation patterns from one of the observational processes Q described above.

We then compute the following alternative depths:

- only in the Common Domain case, the depths restricted to the common domain;

- the depth computed after reconstruction of the missing portions with Goldsmith et al. (2013) with Kraus (2015) and with Kneip and Liebl (2020);
- the proposed POIFD.

To evaluate the goodness of these alternatives, we also compute the depths for the completely observed functional data by IFD with constant weight function $w(t) = 1$ (that coincides with the POIFD for fully observed data). We then measure the agreement between the depths on completely observed data and the depths on the partially observed data, by Spearman correlation. Figure 3 shows the boxplots of Spearman correlations across the $N = 100$ replicates, under the various observation processes, with $p = 50\%$. Under all scenarios, the proposed POIFD attains higher correlations and lower dispersion than any other option considered. Table 1 in the Supplementary Material reports the mean correlations under all settings considered, with $p = 50\%, 25\%, 10\%$ and $m = 1, 4$ for the random interval case. The proposed POIFD is superior to any alternative in all cases. The same conclusions are reached when the agreement is measured by Pearson correlation and by Willmott index (Duveiller et al., 2016), that returns the maximum value of 1 only if the relationship between the two vectors is not simply linear but is exactly the identity. The proposed POIFD is also computationally way more efficient than any of the other considered methods. For instance, for all the simulations in Figure 4, the mean computing time in seconds is: Goldsmith (2013): 12.0068; Kraus (2015): 8.7625; Kneip and Liebl (2020): 37.3021; POIFD: 0.1164.

3.3 Outliers, boxplot and outliergram for partially observed functional data

The presence of outliers may affect any statistical analysis. This is especially the case when the data are high dimensional such as for functional data (Locantore et al., 1999). In this context, fundamental statistics like the mean and the covariance are very sensitive to outliers and become unreliable, and this in turns may negatively affect analyses such as principal component analysis (see, e.g., Hubert et al., 2005). A great number of outlier detection methods for functional data have been introduced, and many of them leverage on depth measures (see, e.g., Febrero-Bande et al., 2008; Hyndman and Shang, 2010; Sun and Genton, 2011; Arribas-Gil and Romo, 2014; Narisetty and Nair, 2015; Nagy et al., 2017). Thanks to the proposed POIFD, these techniques can now

be applied to partially observed functional data. In particular, we here consider an extension of the functional boxplot (Sun and Genton, 2011) and the outliergram (Arribas-Gil and Romo, 2014) to partially observed functional data; these two methods enable the detection of magnitude and shape outliers and also constitutes two insightful visualization tools.

To illustrate the proposal, we contaminate the Gaussian samples in Section 3.1 with shape and magnitude outliers (Dai et al., 2020, see). Following the literature, magnitude outliers are obtained from $X_{ji}^{Mag}(t) = \mu_j(t) + u_{ji} + \epsilon_{ji}(t)$ where $u_{ji} \sim Unif[2.5, 4]$ with probability 1/2, and $u_{ji} \sim Unif[-4, -2.5]$ otherwise. Shape outliers are generated from $X_{ji}^{Shape}(t) = \mu_j(t) + \tilde{\mu}_j(t) + \epsilon_{ji}(t)$ where $\tilde{\mu}$ has covariance $\rho_\mu(s, t)$ with parameters $\beta = 2$ and $\zeta = 0.05$; thus, $X^{Shape}(t)$ is a function with the same mean as $X(t)$, but rougher.

Figure 4 illustrates both outlier detection tools (Common Domain case, $p = 50\%$). For visualization reasons, we here include only two magnitude outliers, in yellow (light grey in black and white printing) and two shape outliers in green (dark grey in black and white printing). The left panel shows the functional boxplot for partially observed data. We here enrich the functional boxplot (Sun and Genton, 2011) by colouring its elements by a grey scale that depends on the proportion $q_n(t)$ of observed functional data at each time point; the median is highlighted in white. The central region contains the fifty percent of the data with the highest POIFDs; the whiskers are constructed considering the functional inter-quantile range inflated by 1.5. Notice that the two magnitude outliers fall partly outside of the whiskers and, thus, they are correctly detected as magnitude outliers. Shape outliers may instead be hidden in the middle of the data, and might thus be missed by the functional boxplots.

The outliergram for partially observed functional data is shown in the right panel of Figure 4. This is based on the quadratic relationship between the partially observed version of the Modified Band Depth and the partially observed version of the Modified Epigraph Index; such quadratic relationship is shown in the graph as a solid parabola. The dotted parabola delimits the shape outlier region, and is built by inflating by 1.5 the solid parabola. The shape outliers are clearly unmasked.

Finally, Figure 5 is like Figure 3, but we have here perturbed samples with outliers. In particular, each of the $N = 100$ samples is composed by a total of $n = 112$ functional data, of which 100 data are generated as X in Section 3.1 while 12 data are outliers, generated as X^{Mag} or X^{Shape}

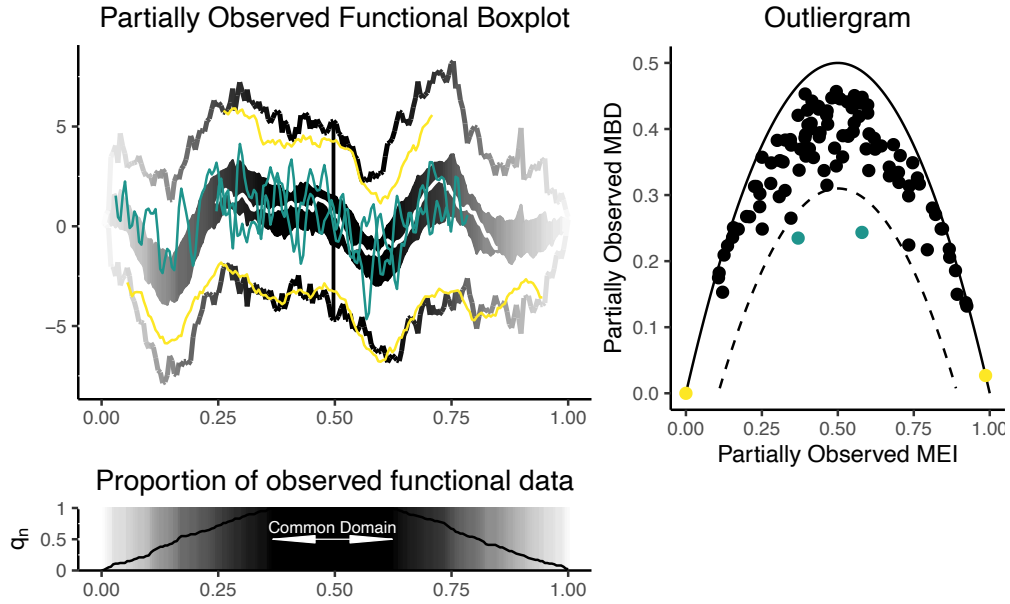


Figure 4: Functional boxplot and outliergram for partially observed functional data. Common Domain case ($p = 50\%$), with addition of two magnitude outliers (yellow in color printing / light grey in black and white printing) and two shape outliers (green in color printing / dark grey in black and white printing). Left panel: functional boxplot for partially observed functional data; its elements are colored in a grey scale that depends on the proportion q_n of observed data; the median is highlighted as a thick white curve. Right panel: outliergram for partially observed functional data.

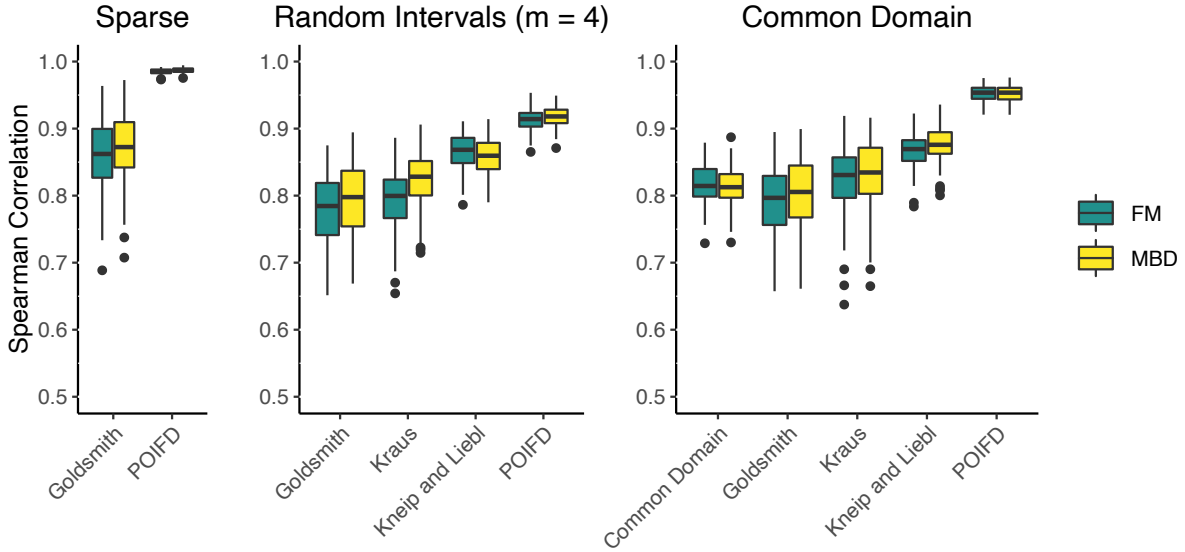


Figure 5: Same as Figure 3, but with outliers included in the dataset.

above, with probability 0.5 each. Likewise in Section 3.2, to enable comparison of the proposed POIFD with depths computed on reconstructed data, we let 25% of the data be observed over the whole evaluation grid, while the remaining 75% are only partially observed. Figure 6 displays the proportion of outliers correctly identified by functional boxplot and outliergram, when computing the depth for sparse and partially observed functional data using the proposed POIFD or using standard depth after reconstruction with Goldsmith et al. (2013), with Kneip and Liebl (2020) and with Kraus (2015). The proportion of correctly identified outliers is significantly higher when using POIFD rather than any alternative.

We can also measure the improvement in the reconstruction methods proposed in Kraus (2015) and Kneip and Liebl (2020), when the outliers detected on the basis of the proposed POIFD are removed. In particular, we consider the mean square error on each functional datum, comparing the reconstructed functional datum with the completely observed datum (excluding from the computation of the mean square error the 25% of the functional data that are fully observed). In the Random Interval case, with $m = 4$, $p = 50\%$, removal of the outliers detected by the proposed POIFD improves the estimation (in average over the $N = 100$ repetitions) by 29.68% when using the technique proposed by Goldsmith et al. (2013), by 48.84% when using the technique by Kraus (2015) and by 63.67% when using the technique by Kneip and Liebl (2020). Analogous results are

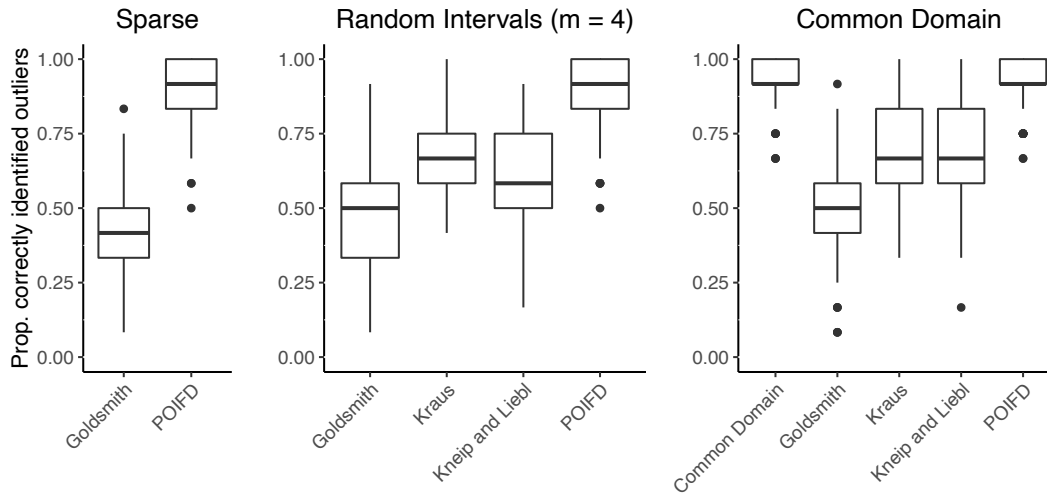


Figure 6: Outlier detection. Proportion of correctly identified outliers by functional boxplot and outliergram, when computing the depth for sparse and partially observed functional data using the proposed POIFD or using standard depth after reconstruction with Goldsmith et al. (2013), with Kneip and Liebl (2020) and with Kraus (2015).

obtained in all other cases, highlighting the importance of proposed POIFD as a crucial support to other techniques for partially observed data.

4 Case studies

4.1 Outlier detection in German electricity supply functions

The range of observability of the electricity price and its supplied quantity changes from one day to another, so that the prices as functions of the quantity are partially observed functional data. We consider here the dataset analyzed in Kneip and Liebl (2020) and Liebl (2019); available at Supplemental Materials A of Liebl (2019). Kneip and Liebl (2020) and Liebl (2019) deal with the problem of reconstructing the unobserved parts and testing the difference in prices before and after the Germany’s nuclear phaseout. To this aim, the authors discard functions corresponding to weekends and holidays as being potential outliers. In addition, they propose to delete any function with prices or supplied quantity greater than a given threshold.

Here we consider the proposed POIFD, and we use the functional boxplot for partially observed

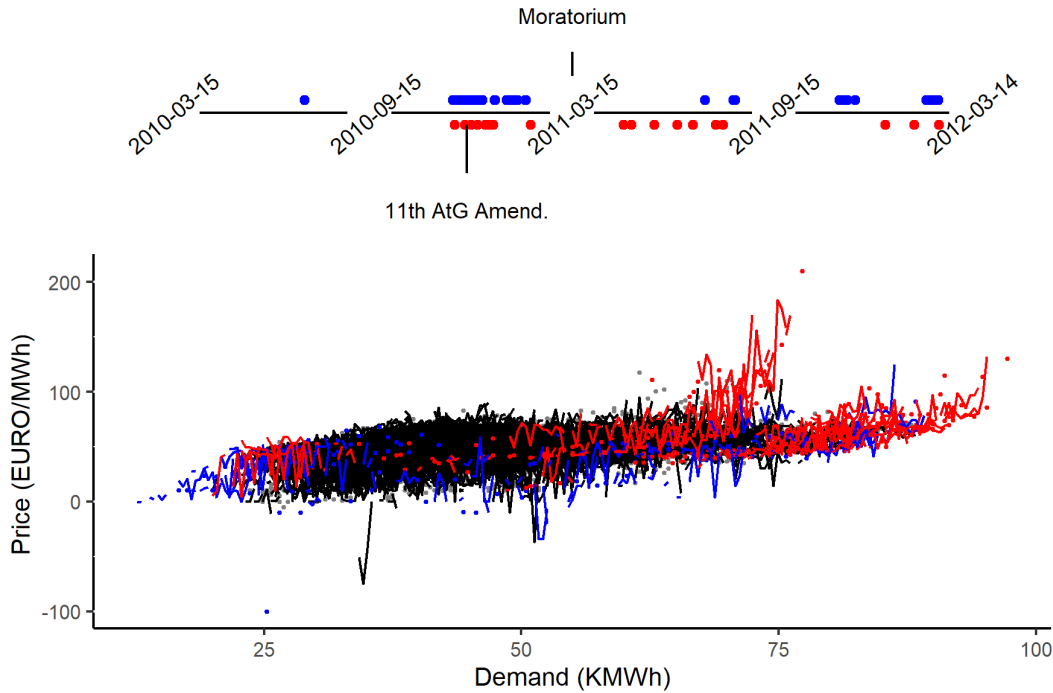


Figure 7: German electricity supply functions, from March 15th, 2010 to March 14th, 2012. Outliers corresponding to working days are plotted in red (light grey in black and white printing) and those corresponding to non-working days in blue (dark grey in black and white printing). The top panel shows the chronological position of the outliers.

functional data and the outliergram based on the the proposed partially observed modified epigraph index, to detect magnitude and shape outliers.

The upper left panel of Figure 1 shows the rainbow plot (Hyndman and Shang, 2010) of the daily electricity supply functions, from March 15th, 2010, to March 14th, 2012. Each daily function is observed at 24 discretization points, one per each hour of the day. Liebl (2019) analyzes only peak hours (9am-8pm) of the working days from March 15, 2012 to March 14, 2013.

It should be pointed out that, from 1998, Germany has driven the phase out of the nuclear energy. However, in September 2010 the country reached a new agreement for extending the operating lives of the German nuclear power plants. This new agreement ended up with the 11th Law Amending the Atomic Energy Act of December 8th, 2010. This period corresponds to the blue-green functions with demand values higher than 75 KMWh. In March 11th, 2011, the Fukushima Daiichi nuclear disaster made Germany react with a moratorium (March 15th, 2011) and shut down forty per cent of its nuclear power plants (period corresponding to green-yellow

functions). This policy has affected both electricity prices and supply. The bottom left panel of Figure 1 plots the proportion q_n of observed data, showing low observability along all supply values (with maximum observability 0.31).

The bottom panel of Figure 7 shows the outliers detected by the functional boxplot and outliergram, based on the proposed POIFD. The outliers are 59 out of the 729 data. Among the 59 outliers, 40 correspond to working days and 19 to non-working days (weekends and holidays); these are plotted in red (light grey in black and white printing) and blue (dark grey in black and white printing), respectively. The top panel of Figure 7 shows the chronology of the outliers. Remarkably, more than a half of the outliers are located between November 2010 and February of 2011, coinciding with a stance shift of the Government which regards German nuclear policy. This suggests that during the days preceding and following the 11th Atomic Energy Act Amendment (blue-green spectrum functions) the electricity market reacted with a shift of the supply curve to the right due to an expected increase in the quantity produced in the following years. Our approach also automatically identifies as outliers the functional data with peaks above the threshold of 200 Euro/MWh, that are discarded by Liebl (2019). In addition to the functional data that exceed this threshold, other functions close in time are also flagged as outliers.

4.2 Classification of AneuRisk65 data

The AneuRisk65 dataset collects the morphology of the inner carotid arteries of 65 subjects. The data are obtained from the reconstruction of three-dimensional angiographic images (see, e.g., Sangalli et al., 2009, 2014b) and are publicly available at <https://statistics.mox.polimi.it/aneurisk/>. The data have been collected and analyzed with the aim of evaluating the role of vascular geometry and haemodynamics on the pathogenesis of cerebral aneurysms. One statistical challenge concerning the analysis of these data is the discrimination of subjects depending on the presence and location of the cerebral aneurysms. In particular, Sangalli et al. (2010) and subsequent papers distinguish the subjects in two groups: the so-called Upper group, composed by those patients having the most dangerous aneurysms (that certainly are within the skull), and the so-called Lower group, composed by subjects without any visible aneurysm during the angiography or having less dangerous aneurysms (because possibly outside of the skull). Likewise previous studies, we focus on two geometrical features of the inner carotid artery, that highly

influence the haemodynamics: its radius profile and its curvature profile (proxied by the profile of the curvature of the carotid centreline). The right panels of Figure 1 displays the data. These functional data are only partially observed. In fact, only a portion of the inner carotid artery of each patient is included in the angiography, with longer portions being available for certain subjects, and shorter for others, depending on where the angiographic image has been centered. In particular the mean proportion of observed data is 62.19%. The missing portions are located only at the extremes of the domain, similarly to the simulation setting with the Common Domain case in Section 3. Specifically, a common domain is available, where the data have been recorded for all statistical units; this corresponds to the portion of inner carotid artery that has been acquired in the angiographic image of all the 65 subjects. This common domain is indicated in the bottom right panel of Figure 1, and constitutes about 23% of the full domain.

The first studies on the AneuRisk65 data, that discriminate Upper group and Lower group subjects, are based on principal component scores and are restricted to the common domain where all data are recorded; see, e.g., Sangalli et al. (2009, 2014a). Recently, Stefanucci et al. (2018) shows that, by using a principal component method for partially observed functional data, it is possible to obtain a better discrimination than the one given in Sangalli et al. (2009, 2014a). In particular, Stefanucci et al. (2018) considers progressive enlargements from the common domain to the full domain, and shows that the best discrimination result is obtained considering a domain that is larger than the common domain but smaller than the full domain. A similar idea is investigated in Kraus and Stefanucci (2019), where the authors though restrict their attention to the radius profiles only.

We here aim to discriminate between the two groups of subjects by Depth vs Depth (DD) classifiers (Cuesta-Albertos et al., 2017), based on the proposed POIFD. Since these functional data are bivariate we consider a bivariate POIFD. In particular, we use the weighted average of the univariate POIFDs for radius and curvature, according to the definition in Equation 4. This enables us to explore different weights $(\alpha, 1 - \alpha)$ given to the two components, ranging from considering only the curvature ($\alpha = 0$) to considering only the radius ($\alpha = 1$). Moreover, we set

$$\phi(q) = \begin{cases} 0 & \text{when } q < q^* \\ q - q^* & \text{when } q \geq q^* \end{cases}$$

where $0 \leq q^* \leq 1$ and q^* is the minimum proportion of observed curves for which the weight function is non-null. This allows us to explore different portions of the domain through different values of q^* , ranging from the full domain ($q^* = 0$) to the common domain ($q^* = 1$). In particular, we select the optimal weight given to radius and curvature (i.e., the value of α) and the optimal domain (i.e., the value of q^*) by minimizing the missclassification error. This approach is very flexible and also highly informative. In fact, on one hand it offers insights on the relative importance of each component of the data; on the other hand, it permits to identify the portion of the domain where the discrimination between the two groups is more relevant.

Specifically, we consider the Leave-1-out missclassification Error Rate (L1ER) for quadratic discriminant analysis on the DD-plot based on the proposed POIFD considering a modified band depth. The left panel of Figure 8 shows the L1ER corresponding to different weights on radius (α) and different portions of the domain (q^*). The miss-classification error increases sharply for low values of the weight on radius, pointing out that the radius is determinant in the discrimination of the two groups of subjects. Moreover, this plot highlights that neither focusing the analysis on the full domain ($q^* = 0$) nor restricting the analysis to the common domain ($q^* = 1$) are optimal for discrimination purposes, supporting the findings in Stefanucci et al. (2018). In particular, the minimum classification error is achieved for an equally weighted bivariate POIFD of radius and curvature, and for $q^* = 0.804$, corresponding to the optimal domain displayed in the bottom right panel of Figure 1, where at least 80.4% of the functional data are observed.

The right panel of Figure 8 shows the DD-plot in correspondence of the optimal values of α and q^* , together with the classification regions determined by quadratic discriminant analysis. The L1ER is 13.8%, tying with the best result obtained in Stefanucci et al. (2018) with the domain selection procedure (compared to 21.54% in Sangalli et al., 2009), with an apparent error rate of 10.76% (compared to 15.38% in Sangalli et al., 2009). Table 3 in Section 8.2 in the Supplementary Materials reports the absolute, relative, and conditional confusion matrices according to L1ER and to APER: the missclassification errors are all lower than those found in the previous discrimination analyses performed on the AneuRisk65 data; see Sangalli et al. (2009), as well as Stefanucci et al. (2018) and Kraus and Stefanucci (2019), that consider the domain selection.

Even if we restrict the analysis to the radius only, the discrimination results are better than those found so far by any of the previous analyses. In particular, the DD-classifier based on linear

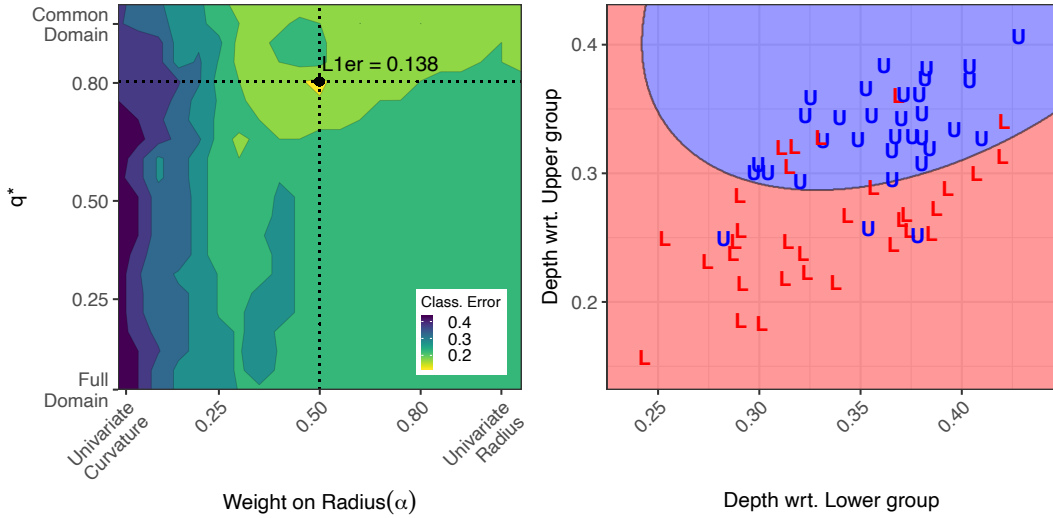


Figure 8: AneuRisk65 data. Left panel: Leave-one-out classification error for different values of the weight on radius and for different values of q^* , the minimum proportion of observed functional data for which the weight function in the POIFD is non-null. Right panel: DD-plot for the optimal value of the weight on radius and of q^* ; upper group in blue (dark grey in black and white printing) and Lower group in red (light grey in black and white printing).

discriminant analysis of univariate POIFD with the Modified Band Depth provides a L1ER of 16.92% (compared to 23.20% in [Kraus and Stefanucci \(2019\)](#)). The optimal value of q^* is here 0.853, corresponding to almost the same domain found in the multivariate case, corroborating the importance of this region for classification purposes.

It should also be pointed out that our analysis, differently from previous analyses, does not require data alignment.

5 Discussion

This work has introduced the first notion of depth measure for partially observed functional data. The investigation of its theoretical properties will be the object of a future dedicated work. Through extensive simulation studies, we have shown the very good performances of the proposed POIFD and its superiority to alternative approaches already available in the literature, whenever the latter are applicable (i.e., in the simplified settings where at least some of the curves are fully observed, or where there is a common domain where all data are recorded). The functional

boxplot and the outliergram based on the proposed POIFD are able to unmask both shape and amplitude outliers. This offers a fundamental support to the available techniques for partially observed functional data, which can be highly affected by outliers. Moreover, the proposed POIFD can efficiently be used in classification problems, as illustrated with the application to Aneurisk65 data, where a DD-classifier based on the proposed depth leads to a missclassification error lower than any of the other discrimination techniques ever applied to these challenging data.

We are certain that the proposed depth will prove highly valuable in a number of applications where partially observed functional data are encountered.

SUPPLEMENTARY MATERIAL

Additional results: Extensive simulation results (Section 7). Additional output concerning the analysis of German electricity supply curves and AneuRisk65 dataset (Section 8). (AdditionalResults.pdf)

R-package: R-package `fdaPOIFD` available at CRAN ([Elías et al., 2021](#)) includes the functions to compute the depth, plot the boxplot and the outliergram for partially observed functional data, as well as a vignette to reproduce the simulations.

References

- Arribas-Gil, A. and Romo, J. (2014). Shape outlier detection and visualization for functional data: the outliergram. *Biostatistics*, 15(4):603–619.
- Claeskens, G., Hubert, M., Slaets, L., and Vakili, K. (2014). Multivariate functional halfspace depth. *Journal of the American Statistical Association*, 109(505):411–423.
- Cuesta-Albertos, J. A., Febrero-Bande, M., and Oviedo de la Fuente, M. (2017). The DD^g -classifier in the functional setting. *TEST*, 26(1):119–142.
- Dai, W. and Genton, M. G. (2018). Functional boxplots for multivariate curves. *Stat*, 7:e190, 10.
- Dai, W., Mrkvička, T., Sun, Y., and Genton, M. G. (2020). Functional outlier detection and taxonomy by sequential transformations. *Computational Statistics & Data Analysis*, 149:106960.

- D'Amato, V., Piscopo, G., and Russolillo, M. (2011). The mortality of the italian population: Smoothing techniques on the lee-carter model. *Annals of Applied Statistics*, 5(2A):705–724.
- Delaigle, A. and Hall, P. (2013). Classification using censored functional data. *Journal of the American Statistical Association*, 108(504):1269–1283.
- Delaigle, A. and Hall, P. (2016). Approximating fragmented functional data by segments of markov chains. *Biometrika*, 103(4):779–799.
- Di, C., Crainiceanu, C. M., and Jank, W. S. (2014). Multilevel sparse functional principal component analysis. *Stat*, 3(24872597):126–143.
- Duveiller, G., Fasbender, D., and Meroni, M. (2016). Revisiting the concept of a symmetric index of agreement for continuous datasets. *Scientific Reports*, 6:19401.
- Elías, A., Jiménez, R., Paganoni, A., and Sangalli, L. (2021). *fdaPOIFD: Partially Observed Integrated Functional Depth*. R package version 1.0.2.
- Febrero-Bande, M., Galeano, P., and González-Manteiga, W. (2008). Outlier detection in functional data by depth measures, with application to identify abnormal nox levels. *Environmetrics*, 19(4):331–345.
- Fraiman, R. and Muniz, G. (2001). Trimmed means for functional data. *Test*, 10(2):419–440.
- Genton, M. G., Johnson, C., Potter, K., Stenchikov, G., and Sun, Y. (2014). Surface boxplots. *Stat*, 3:1–11.
- Gijbels, I. and Nagy, S. (2017). On a general definition of depth for functional data. *Statistical Science*, 32(4):630–639.
- Goldberg, Y., Ritov, Y., and Mandelbaum, A. (2014). Predicting the continuation of a function with applications to call center data. *Journal of Statistical Planning and Inference*, 147:53–65.
- Goldsmith, J., Greven, S., and Crainiceanu, C. (2013). Corrected confidence bands for functional data using principal components. *Biometrics*, 69(1):41–51.
- Hubert, M., Rousseeuw, P., and Segaeert, P. (2017). Multivariate and functional classification using depth and distance. *Advances in Data Analysis and Classification*, 11(3):445–466.

- Hubert, M., Rousseeuw, P. J., and Vanden Branden, K. (2005). Robpca: A new approach to robust principal component analysis. *Technometrics*, 47(1):64–79.
- Human Mortality Database (2019). University of california, berkeley (usa) and max planck institute for demographic research (germany). Available at www.mortality.org.
- Hyndman, R. J. and Shang, H. L. (2010). Rainbow plots, bagplots, and boxplots for functional data. *Journal of Computational & Graphical Statistics*, 19(1):29–45.
- Ieva, F. and Paganoni, A. M. (2013). Depth measures for multivariate functional data. *Communications in Statistics - Theory and Methods*, 42(7):1265–1276.
- Ieva, F. and Paganoni, A. M. (2017). Component-wise outlier detection methods for robustifying multivariate functional samples. *Statistical Papers*, pages 1–20.
- James, G., Hastie, T., and Sugar, C. (2000). Principal component models for sparse functional data. *Biometrika*, 87(3):587–602.
- James, G. M. and Hastie, T. J. (2001). Functional linear discriminant analysis for irregularly sampled curves. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):533–550.
- Kneip, A. and Liebl, D. (2020). On the optimal reconstruction of partially observed functional data. *The Annals of Statistics*, 48(3):1692–1717.
- Kraus, D. (2015). Components and completion of partially observed functional data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(4):777–801.
- Kraus, D. and Stefanucci, M. (2019). Classification of functional fragments by regularized linear classifiers with domain selection. *Biometrika*, 106(1):161–180.
- Li, J., Cuesta-Albertos, J. A., and Liu, R. Y. (2012). Dd-classifier: Nonparametric classification procedure based on dd-plot. *Journal of the American Statistical Association*, 107(498):737–753.
- Liebl, D. (2019). Nonparametric testing for differences in electricity prices: The case of the fukushima nuclear accident. *The Annals of Applied Statistics*, 2(13):1128–1146.
- Liebl, D. and Rameseder, S. (2019). Partially observed functional data: The case of systematically missing parts. *Computational Statistics & Data Analysis*, 131:104 – 115. High-dimensional and functional data analysis.

- Lin, Z., Wang, J.-L., and Zhong, Q. (2020). Basis expansions for functional snippets.
- Liu, C., Ray, S., and Hooker, G. (2017). Functional principal component analysis of spatially correlated data. *Statistics and Computing*, 27(6):1639–1654.
- Locantore, N., Marron, J. S., Simpson, D. G., Tripoli, N., Zhang, J. T., Cohen, K. L., Boente, G., Fraiman, R., Brumback, B., Croux, C., Fan, J., Kneip, A., Marden, J. I., Peña, D., Prieto, J., Ramsay, J. O., Valderrama, M. J., Aguilera, A. M., Locantore, N., Marron, J. S., Simpson, D. G., Tripoli, N., Zhang, J. T., and Cohen, K. L. (1999). Robust principal component analysis for functional data. *Test*, 8(1):1–73.
- López-Pintado, S. and Romo, J. (2009). On the concept of depth for functional data. *Journal of the American Statistical Association*, 104(486):718–734.
- López-Pintado, S. and Romo, J. (2011). A half-region depth for functional data. *Computational Statistics & Data Analysis*, 55(4):1679–1695.
- Nagy, S., Gijbels, I., and Hlubinka, D. (2017). Depth-based recognition of shape outlying functions. *Journal of Computational & Graphical Statistics*, 26(4):883–893.
- Nagy, S., Gijbels, I., Omelka, M., and Hlubinka, D. (2016). Integrated depth for functional data: statistical properties and consistency. *ESAIM. Probability and Statistics*, 20.
- Narisetty, N. N. and Nair, V. N. (2015). Extremal depth for functional data and applications. *Journal of the American Statistical Association*, pages 1–38.
- Nieto-Reyes, A. and Battey, H. (2016). A topologically valid definition of depth for functional data. *Statistical Science*, 31(1):61–79.
- Qu, Zhuo and, G. M. G. (2022). Sparse functional boxplots for multivariate curves.
- Sangalli, L. M., Secchi, P., and Vantini, S. (2014a). Analysis of aneurisk65 data: k -mean alignment. *Electronic Journal of Statistics*, 8(2):1891–1904.
- Sangalli, L. M., Secchi, P., and Vantini, S. (2014b). Aneurisk65: A dataset of three-dimensional cerebral vascular geometries. *Electronic Journal of Statistics*, 8(2):1879–1890.
- Sangalli, L. M., Secchi, P., Vantini, S., and Veneziani, A. (2009). A case study in exploratory functional data analysis: Geometrical features of the internal carotid artery. *Journal of the American Statistical Association*, 104(485):37–48.

- Sangalli, L. M., Secchi, P., Vantini, S., and Vitelli, V. (2010). k-mean alignment for curve clustering. *Computational Statistics & Data Analysis*, 54(5):1219–1233.
- Sguera, C. and López-Pintado, S. (2020). A notion of depth for sparse functional data. *TEST*.
- Stefanucci, M., Sangalli, L. M., and Brutti, P. (2018). Pca-based discrimination of partially observed functional data, with an application to aneurisk65 data set. *Statistica Neerlandica*, 72(3):246–264.
- Sun, Y. and Genton, M. G. (2011). Functional boxplots. *Journal of Computational & Graphical Statistics*, 20(2):316–334.
- Sun, Y. and Genton, M. G. (2012). Adjusted functional boxplots for spatio-temporal data visualization and outlier detection. *Environmetrics*, 23(1):54–64.
- Yao, F., Muller, H.-G., and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*, 100(470):577–590.

Supplementary material for Integrated Depths for Partially Observed Functional Data

Antonio Elías,
OASYS Group
Department of Applied Mathematics, Universidad de Málaga
Raúl Jiménez,
Department of Statistics, University Carlos III of Madrid
Anna M. Paganoni and Laura M. Sangalli
MOX Laboratory for Modeling and Scientific Computing
Dipartimento di Matematica, Politecnico di Milano

7 Simulation studies: additional results

7.1 Comparison with alternative methods

Table [1](#) summarizes the results obtained for the simulations described in Section [3.2](#) of the main article, under various levels of observability, $p = 50\%$, 25% , 10% , and numbers of intervals $m = 1, 4$, in the random interval case. But for one case (sparse data, $p = 0.25\%$) where the MBD on data reconstructed using [Goldsmith et al. \(2013\)](#) is comparable to the proposed POIFD (has higher Spearman and lower Wilcott), in all other cases the proposed POIFD is significantly superior to all other alternatives.

7.2 Comparison with alternative weights

We now consider the more challenging scenario where none of the functional data is fully observed: the data are as in Section [3.2](#) of the main text, but all the data are partially observed according to one of the considered observational processes, whilst none is fully observed. Table [2](#) shows that also in this more challenging case the proposed POIFD has very high agreement with the depth on the fully observed data. The same table also shows the results that would be obtained for different choices of the weight with respect to the one in equation [\(1\)](#): the first corresponds to Inverse Probability Weighting (IPW), where the weight is $1/\mathbb{P}(\mathcal{O} \ni t)$, while the second corresponds to a constant weight $w(t) = 1$ for all $t \in [0, 1]$. In the case of sparse data, the proposed POIFD with the weight in [\(1\)](#) and with a constant weight return comparable results, while IPW returns inferior (or at most equal) agreements with the true depths. In the case of partially observed functional data, the proposed POIFD with the weight in equation [\(1\)](#) is always

Agreement	Fraiman and Muniz (FM)						Modified Band Depth (MBD)					
	Spearman			Willmott			Spearman			Willmott		
	50	25	10	50	25	10	50	25	10	50	25	10
Observability (%)												
Sparse												
Goldsmith (2013)	0.965	0.936	0.855	0.829	0.777	0.670	0.977	0.952	0.874	0.858	0.808	0.700
POIFD	0.981	0.942	0.863	0.916	0.844	0.747	0.983	0.947	0.874	0.920	0.852	0.761
Random Intervals												
m = 1												
Goldsmith (2013)	0.737	0.523	0.361	0.631	0.447	0.241	0.752	0.546	0.375	0.654	0.482	0.290
Kraus (2015)	0.736	0.467	0.312	0.600	0.268	0.038	0.761	0.473	0.311	0.650	0.333	0.097
Kneip and Liebl (2020)	0.809	0.509	0.329	0.699	0.415	0.244	0.821	0.539	0.349	0.721	0.455	0.291
POIFD	0.886	0.668	0.498	0.749	0.489	0.250	0.886	0.667	0.495	0.756	0.507	0.270
m = 4												
Goldsmith (2013)	0.813	0.674	0.632	0.694	0.576	0.488	0.827	0.697	0.657	0.716	0.608	0.527
Kraus (2015)	0.790	0.585	0.535	0.646	0.422	0.341	0.801	0.591	0.539	0.677	0.459	0.373
Kneip and Liebl (2020)	0.820	0.559	0.439	0.729	0.511	0.412	0.811	0.540	0.426	0.727	0.504	0.407
POIFD	0.879	0.743	0.679	0.765	0.611	0.527	0.884	0.750	0.693	0.774	0.627	0.551
Common Domain												
Depth common domain	0.760	0.648	0.570	0.491	0.277	0.078	0.755	0.638	0.561	0.507	0.310	0.133
Goldsmith (2013)	0.816	0.763	0.740	0.703	0.653	0.633	0.830	0.773	0.752	0.722	0.673	0.657
Kraus (2015)	0.826	0.755	0.717	0.686	0.598	0.55	0.830	0.756	0.722	0.716	0.639	0.595
Kneip and Liebl (2020)	0.909	0.858	0.827	0.816	0.763	0.732	0.913	0.863	0.835	0.829	0.778	0.749
POIFD	0.937	0.904	0.885	0.832	0.786	0.761	0.936	0.901	0.884	0.837	0.792	0.768

Table 1: Agreement between depths on completely observed sample and various alternatives to compute depths on partially observed data. The alternatives include: computation of depths on the common domain (in the Common Domain case); computation of depths after reconstruction following Goldsmith et al. (2013), after reconstruction following Kneip and Liebl (2020), and after reconstruction following Kraus (2015); the proposed POIFD.

significantly superior to any alternatives. The poor results attained by IPW are due to the fact that IPW gives more weight to portions of the domain where the missingness is more severe, and thus the depth of the real data is more difficult to ascertain. Differently from IPW, the weight we use in Eq. (1) gives more weight in the computation of the depth to regions with high observational density and less weight to regions with low observational density.

Agreement	Fraiman and Muniz (FM)						Modified Band Depth (MBD)					
	Spearman			Willmott			Spearman			Willmott		
Observability (%)	50	25	10	50	25	10	50	25	10	50	25	10
Sparse Data												
IPW	0.975	0.918	0.675	0.902	0.815	0.604	0.978	0.923	0.661	0.908	0.820	0.588
Constant	0.975	0.921	0.721	0.903	0.819	0.635	0.978	0.927	0.713	0.909	0.825	0.627
POIFD	0.975	0.919	0.733	0.903	0.817	0.646	0.978	0.925	0.730	0.909	0.823	0.665
Random Intervals												
m = 1												
IPW	0.324	0.297	0.282	0.408	0.269	0.021	0.333	0.304	0.284	0.410	0.276	0.035
Constant	0.757	0.557	0.374	0.587	0.320	0.053	0.759	0.555	0.368	0.596	0.341	0.083
POIFD	0.886	0.602	0.394	0.745	0.403	0.114	0.884	0.600	0.382	0.752	0.419	0.140
m = 4												
IPW	0.688	0.590	0.478	0.590	0.464	0.372	0.701	0.600	0.476	0.682	0.561	0.452
Constant	0.831	0.669	0.515	0.689	0.523	0.397	0.839	0.680	0.512	0.701	0.541	0.402
POIFD	0.853	0.690	0.550	0.717	0.542	0.414	0.860	0.697	0.553	0.729	0.558	0.425
Common Domain												
IPW	0.369	0.312	0.324	0.458	0.421	0.405	0.385	0.325	0.337	0.464	0.429	0.408
Constant	0.827	0.779	0.748	0.678	0.615	0.585	0.827	0.781	0.748	0.909	0.825	0.627
POIFD	0.944	0.917	0.900	0.838	0.791	0.765	0.942	0.915	0.895	0.842	0.797	0.771

Table 2: Comparison with alternative weighting functions. Agreement between depths on completely observed sample and the depth on partially observed data, considering different specifications of the weight. The alternatives include: Inverse Probability Weighting (IPW); constant weight; the weight in eq. (1) for the proposed POIFD.

8 Case studies: additional results

8.1 Outlier detection in German electricity supply functions

Figure 10 shows the functional boxplot and the outliergram for the German electricity supply functions, highlighting the detected outliers. It should be pointed out that the boxplot is able to point out some outliers that are not the typical magnitude outliers nor shape outliers. These outliers are functional data that are observed over a portion of the domain where the central 50% are not instead typically observed. These novel type of outliers are pointed out in this case because in this applied problem the assumption of Missing-Completely-at-Random does not hold.

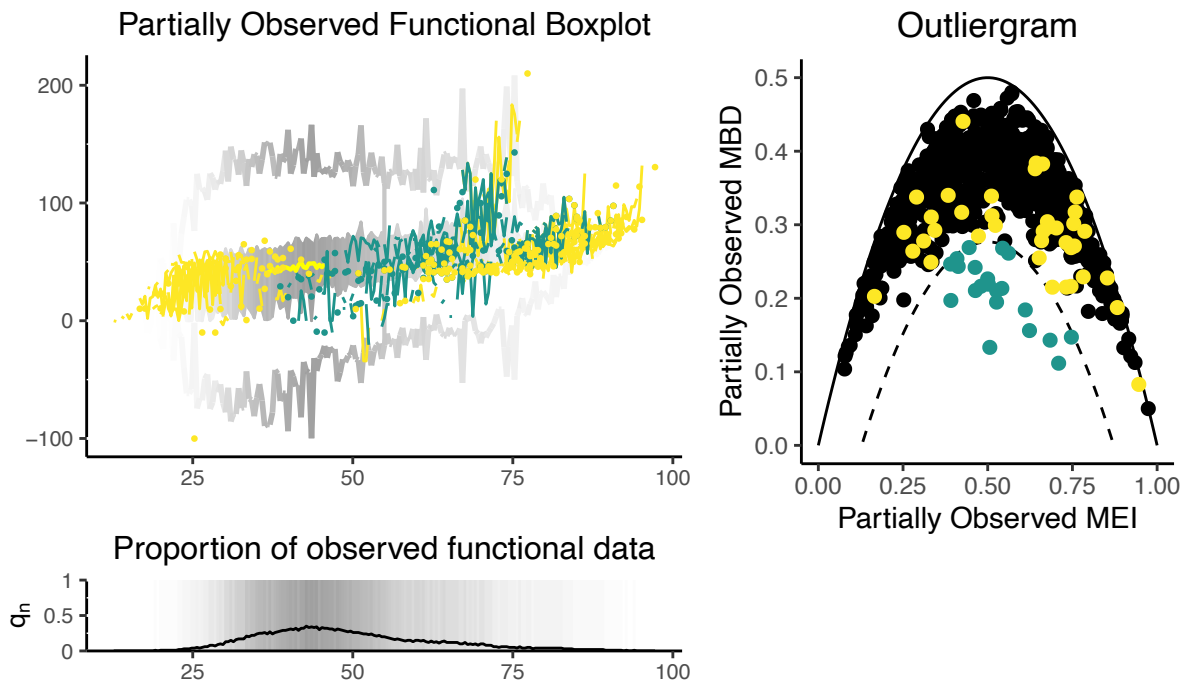


Figure 10: Outlier analysis of the German electricity supply functions. Functional boxplot and outliergram. Green (dark grey in black and white printing) and yellow (light grey in black and white printing) for shape and magnitude outliers respectively.

8.2 Classification of AneuRisk65 data

Table 3 reports the absolute, relative and conditional confusion matrices, estimated according to Leave-one-out Error (L1ER) and Apparent Error Rate (APER), for the classification of the AneuRisk65. All the errors are lower than those obtained in Sangalli et al. (2009).

Table 3: Absolute, relative, and conditional confusion matrices estimated according to Leave-one-out Error (L1ER) and Apparent Error Rate (APER).

L1ER = 13.84%			APER = 10.76 %		
Abs.	Lower	Upper	Abs.	Lower	Upper
Lower	26	3	Lower	28	3
Upper	6	30	Upper	4	30
Rel.	Lower	Upper	Rel.	Lower	Upper
Lower	40.00%	0.04%	Lower	43.07%	4.61%
Upper	9.23%	46.15%	Upper	6.15%	46.15%
Cond.	Lower	Upper	Cond.	Lower	Upper
Lower	89.66%	10.33%	Lower	90.33%	9.67%
Upper	16.66%	83.33%	Upper	11.76%	88.24%

References

- Goldsmith, J., Greven, S., and Crainiceanu, C. (2013). Corrected confidence bands for functional data using principal components. *Biometrics*, 69(1):41–51.
- Kneip, A. and Liebl, D. (2020). On the optimal reconstruction of partially observed functional data. *The Annals of Statistics*, 48(3):1692–1717.
- Kraus, D. (2015). Components and completion of partially observed functional data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(4):777–801.
- Sangalli, L. M., Secchi, P., Vantini, S., and Veneziani, A. (2009). A case study in exploratory functional data analysis: Geometrical features of the internal carotid artery. *Journal of the American Statistical Association*, 104(485):37–48.

MOX Technical Reports, last issues

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

- 49/2022** Botti, M.; Fumagalli, A.; Scotti, A.
Uncertainty quantification for mineral precipitation and dissolution in fractured porous media
- 48/2022** Gregorio, C.; Barbati, G.; Ieva, F.
A wavelet-mixed landmark survival model for the effect of short-term oscillations in longitudinal biomarker's profiles
- 47/2022** Botti, M.; Di Pietro, D.A.; Salah, M.
A serendipity fully discrete div-div complex on polygonal meshes
- 45/2022** Franco, N.; Fresca, S.; Manzoni, A.; Zunino, P.
Approximation bounds for convolutional neural networks in operator learning
- 46/2022** Lucca, A.; Fraccarollo, L.; Fossan, F.E.; Braten, A.T.; Pozzi, S.; Vergara, C.; Muller, L.O.
Impact of pressure guidewire on model-based FFR prediction
- 44/2022** Peli, R.; Dovera, L.; Figuera, G.; Menafoglio, A.; Secchi, P.
Forecasting Oil Production Rates in Primary Depletion using the Physics-based Residual Kriging functional approach
- 43/2022** Zappone E.; Manzoni A.; Gervasio P.; Quarteroni A.
A reduced order model for domain decompositions with non-conforming interfaces
- 40/2022** Fumagalli, A.; Patacchini, F. S.
Well-posedness and variational numerical scheme for an adaptive model in highly heterogeneous porous media
- 42/2022** Gatti, F.; Fois, M.; de Falco, C.; Perotto, S.; Formaggia, L.
Parallel simulations for fast-moving landslides: space-time mesh adaptation and sharp tracking of the wetting front
- 41/2022** Arnone, A.; Ferraccioli, F.; Pigolotti, C.; Sangalli, L.M.
A roughness penalty approach to estimate densities over two-dimensional manifolds