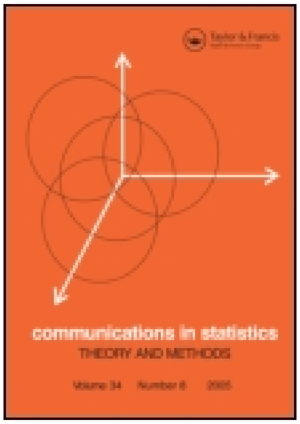

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

Depth Measures for Multivariate Functional Data

Francesca Ieva [a] & Anna M. Paganoni [a]

[a] MOX - Department of Mathematics, Politecnico di Milano, Milano, Italy

Version of record first published: 21 Feb 2013.

To cite this article: Francesca Ieva & Anna M. Paganoni (2013): Depth Measures for Multivariate Functional Data, Communications in Statistics - Theory and Methods, 42:7, 1265-1276

To link to this article: http://dx.doi.org/10.1080/03610926.2012.746368

Communications in Statistics—Theory and Methods, 42: 1265–1276, 2013

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610926.2012.746368

Depth Measures for Multivariate Functional Data

FRANCESCA IEVA AND ANNA M. PAGANONI

MOX – Department of Mathematics, Politecnico di Milano, Milano, Italy

In this article, we address the problem of mining and analyzing multivariate functional data. That is, data where each observation is a set of possibly correlated functions. Complex data of this kind is more and more common in many research fields, particularly in the biomedical context. In this work, we propose and apply a new concept of depth measure for multivariate functional data. With this new depth measure it is possible to generalize robust statistics, such as the median, to the multivariate functional framework, which in turn allows the application of outlier detection, boxplots construction, and nonparametric tests also in this more general framework. We present an application to Electrocardiographic (ECG) signals.

Keywords Depth measures; ECG signals; Multivariate functional data; Rank tests.

Mathematics Subject Classification Primary 62P10; Secondary 92C55.

1. Introduction and Notation

Today the data coming from a biomedical context are frequently functions or images produced by medical devices. This calls for the identification of suitable models and inferential techniques for managing the complexity of such data. For example, a challenging task in functional data analysis is to provide an ordering within a sample of curves to allow the definition of order statistics, such as ranks and L-statistics (see Fraiman and Meloche, 1999). A natural approach to analyze features of functional data is the concept of statistical depth, which provides a measure of centrality or outlyingness of an observation with respect to a given dataset or to a population distribution. Several definitions of depth measures for multivariate data have been proposed and analysed in the literature (for example, see Tukey, 1975; Liu, 1990; Liu and Singh, 1993; Zuo and Serfling, 2000; Zuo, 2003; among others). A generalization to univariate functional data that starts from depth measures for multivariate data is given in Lopez-Pintado and Romo (2009). They also provide an extension of robust statistics to a functional framework. This extension generalizes properties of depth measures which are proved to hold in the multivariate case (see Liu, 1990; Zuo and Serfling, 2000; Serfling, 2006 for further

Received October 28, 2011; Accepted October 13, 2012
Address correspondence to Francesca Ieva, MOX – Department of Mathematics, Politecnico di Milano, p.le Leonardo da Vinci, 32 Milano, 20133 Italy; E-mail: francesca.ieva@mail.polimi.it

details on the multivariate setting). A specific focus on trimmed means for functional data is presented in Fraiman and Muniz (2001). They propose a generalization of some results of Fraiman and Meloche (1999) about multivariate L-estimators.

In this article, we address the problem of exploring and analysing multivariate functional data. That is, each data observation is a set of possibly correlated functions. In particular, our concerns focus on the definition of suitable depth indexes to rank multivariate functional objects and then to make inference on them. In fact, in several applications, the main focuses are to perform an unsupervised clustering of functions arising from different populations and to make inference about the latent differences among clusters, for example analyzing the morphological structure of the curves shape. This is usually implemented without parametric assumptions on the model generating the sample of curves, as in Lopez-Pintado and Romo (2006) and Cuevas et al. (2007). There is also an interest in making inference based on specific summary statistics, as proposed in Li and Liu (2004) for the multivariate setting. Once a depth measure is associated with each univariate/multivariate functional data within a sample, it is possible to rank data as well as to visualize the result of ranking through functional boxplots, as proposed in Sun and Genton (2011), Sun et al. (2012) and generalized in Ieva (2011).

In this article, we address with multivariate functional data. First, we generalize the concept of a depth measure for univariate functional data to the multivariate functional case; this depth measure is derived from averaging univariate centrality measures for functional data in a suitable index. Then, we define suitable generalizations of nonparametric statistics for ranking and classifying multivariate curves as well as making inference on them. We also widen the employment of functional boxplots. This graphical tool is generalized to the more complex case of samples of multivariate functions. Finally, an extension of Wilcoxon rank test based on the order induced by the multivariate functional depth is proposed to test differences between groups of multivariate curves.

In fact, two are the main goals of the analysis: the first is to point out a suitable method for detection of outliers in a multivariate functional setting, within a sample of curves arising from the same population. The second is to carry out a nonparametric test for comparing samples of multivariate curves and making inference on the differences between the corresponding populations.

A natural application of this theoretical framework is in a biomedical context. For example, applications that deal with cardiovascular diseases diagnosed by Electrocardiographic (ECG) devices. In fact, ECG signals can be considered multivariate functional data with correlated components. Each data describes the same biological event, i.e., the heartbeat. In this context, some interests are classification of groups of curves with similar morphological patterns, multivariate functional outlier detection within a homogeneous group, and classical inference on means and quantiles of subpopulations. From a clinical view point, the first concern is how to carry out a semi automatic diagnosis based only on the morphological deviations from physiological patterns induced by the presence of the disease of interest. The second issue leads to profile "typical" curve expression for each pathology. The third aim is the investigation for the presence of statistically significant differences in the subpopulations of pathological units with respect to physiological ones.

This article is organized as follows. In Sec. 2, a definition of multivariate functional depth is presented and the statistical properties of this depth measure are proved in the more general framework of multivariate functional data. In Sec. 3, an application to ECG signals of healthy patients and patients affected by Bundle Brunch Block Infarction is presented. Finally, Sec. 4 contains conclusions, discussion of results and further developments. The proofs of propositions stated in Sec. 2.1 are included in the Appendix.

2. Band Depth and Inference for Multivariate Functional Data

In this section, a new concept of multivariate functional depth measure is presented and some natural properties are established. Moreover, a modified version of the band depth is given and it is used to construct an extension of functional boxplot to the multivariate setting. Finally, a Wilcoxon non parametric rank test framework is adopted to make inference on samples of multivariate functional data, once they have been ranked according to a suitable multivariate functional index of depth.

2.1. *Band depth for Multivariate Functional Data*

As mentioned in the previous section, a natural tool to analyze and rank data is based on the concept of statistical depth. Statistical depth measures the centrality of a given observation within a group of data providing a center-outward ordering of the set itself. In general, several different definitions of depth can be given (see Zuo and Serfling, 2000). In our case, we refer to the band depth measure for univariate functional data proposed in Lopez-Pintado and Romo (2009) and we introduce one way of extending this notion to the framework of multivariate functional data.

Let X be a stochastic process with a law P taking values on the space $\mathscr{C}(I)$ of real continuous functions on the compact interval I. The graph of a function $f \in \mathscr{C}(I)$ is the subset of the plane $G(f) = \{(t, f(t)) : t \in I\}$. The random band depth of order $J \geq 2$ for a function $f \in \mathscr{C}(I)$ is

$$BD^J_{P_X}(f) = \sum_{j=2}^{J} P_X\{G(f) \subset B(X_1, X_2, \ldots, X_j)\},$$

where $B(X_1, X_2, \ldots, X_j)$ is the random band in $\mathbb{R}^2$ delimited by $X_1, \ldots, X_j$, independent copies of the stochastic process X, defined as

$$B(X_1, \ldots, X_j) = \left\{(t, y(t)) : t \in I, \min_{r=1,\ldots,j} X_r(t) \leq y(t) \leq \max_{r=1,\ldots,j} X_r(t)\right\}, \text{ for } j = 2, \ldots, J.$$

In this article, we propose a new definition of a band depth measure for multivariate functional data, i.e., data generated by a stochastic process $\mathbf{X}$ taking values in the space $\mathscr{C}(I; \mathbb{R}^s)$ of continuous functions $\mathbf{f} = (f_1, \ldots, f_s) : I \to \mathbb{R}^s$.

Definition 2.1. Let $\mathbf{f}$ be a function on I taking values in $\mathbb{R}^s$. The *multivariate band depth measure* is defined as

$$BD^J_{P_\mathbf{X}}(\mathbf{f}) = \sum_{k=1}^{s} p_k BD^J_{P_{X_k}}(f_k), \quad p_k > 0 \text{ for } k = 1, \ldots, s, \quad \sum_{k=1}^{s} p_k = 1. \tag{1}$$

Obviously, this is just one of the possible extensions of the concept of band depth to the multivariate functional setting. This definition allows the analyst to take into account all the components of the multivariate data. In general, the choice of weights in (1) is problem driven. So doing, weights may account for any previous knowledge about the data and/or about the correlation structure between the components of multivariate functional data. In particular, the dependence between data components, which depict the same event from different perspectives, should be summarized in a suitable choice of weights, since they rule how to average depth measures of marginal components. The freedom in weights choice causes a lack of uniqueness in the proposed definition, but guarantees a great adaptivity to the real problem under study.

Let $\mathbf{X}$ be a multivariate random process such that $P(\min_{k=1,\dots,s} \|X_k\|_\infty > M) \to 0$ as $M \to \infty$. Using the properties of the functional depth measure summarized in Lopez-Pintado and Romo (2009), it is easy to prove the following results on the basic properties of the multivariate band depth measure defined in (1).

Proposition 2.1.

(a) *Let* $T(\mathbf{f}) = \mathbf{A}(t)\mathbf{f}(t) + \mathbf{b}(t)$, *where* $\forall t \in I$ $\mathbf{A}(t)$ *is a* $s \times s$ *diagonal matrix such that* $\mathbf{A_{kk}}(t)$ *are continuous functions in* I, *with* $\mathbf{A_{kk}}(t) \neq 0$, *for each* $t \in I$, *and* $\mathbf{b}(t) \in \mathscr{C}(I; \mathbb{R}^s)$. *Then* $BD^J_{P_{T(\mathbf{X})}}(T(\mathbf{f})) = BD^J_{P_\mathbf{X}}(\mathbf{f})$.
(b) $BD^J_{P_{\mathbf{X}(g(t))}}(\mathbf{f}(g(t))) = BD^J_{P_{\mathbf{X}(t)}}(\mathbf{f}(t))$ *when* g *is a one-to-one transformation of the interval* I.
(c) $\sup_{\min_{k=1,\dots,s}\|f_k\|_\infty > M} BD^J_{P_\mathbf{X}}(\mathbf{f}) \to 0$ *as* $M \to \infty$.
(d) *If* $\forall k = 1, \dots, s$ *the probability distribution* P_{X_k} *on* $\mathscr{C}(I)$ *has absolutely continuous marginal distributions, then* $BD^J_{P_\mathbf{X}}$ *is a continuous functional on* $\mathscr{C}(I; \mathbb{R}^s)$.

If $\mathbf{X}_1, \dots, \mathbf{X}_n$ are independent copies of the stochastic process $\mathbf{X}$, the sample version of (1) can be introduced in order to conduct descriptive and inferential statistical analyses on a set of multivariate functional data $\mathbf{f}_1, \dots, \mathbf{f}_n$ generated by the process $\mathbf{X}$. For any $\mathbf{f}$ in the sample $\mathbf{f}_1, \dots, \mathbf{f}_n$ we can compute the depth as

$$BD^J_n(\mathbf{f}) = \sum_{k=1}^{s} p_k BD^J_{n,k}(f_k),$$

where for the function $f_k \in \mathscr{C}(I)$

$$BD^J_{n,k}(f_k) = \sum_{j=2}^{J} \binom{n}{j}^{-1} \sum_{1 \le i_1 < i_2 < \cdots < i_j \le n} \mathbb{I}\{G(f_k) \subset B(f_{i_1;k}, \dots f_{i_j;k})\}$$

and $\mathbb{I}\{G(f_k) \subset B(f_{i_1;k}, \dots, f_{i_j;k})\}$ indicates if the band determined by $(f_{i_1;k}, \dots, f_{i_j;k})$ contains the whole graph of f. The k component of the vector $\mathbf{f_i}$ is denoted by $f_{i;k}$.

Proposition 2.2. *The sample version of multivariate functional depth is consistent, i.e.,*

$$|BD^J_n(\mathbf{f}) - BD^J_{P_\mathbf{X}}(\mathbf{f})| \to 0, \quad a.s. \text{ if } n \to \infty. \tag{2}$$

Moreover, the convergence is uniform on any equicontinuous set A *of functions on* I *taking values in* $\mathbb{R}^s$.

As proposed in Lopez-Pintado and Romo (2009), and also in this multivariate functional setting, we can move to the analogous of the modified band depth:

$$MBD_n^J(\mathbf{f}) = \sum_{k=1}^{s} p_k MBD_{n,k}^J(f_k), \tag{3}$$

where for the function $f_k \in \mathscr{C}(I)$ the modified band depth measures the proportion of time that the curve f_k is in the band, i.e.,

$$MBD_{n,k}^J(f_k) = \sum_{j=2}^{J} \binom{n}{j}^{-1} \sum_{1 \le i_1 < i_2 < \cdots < i_j \le n} \tilde{\lambda}\{E(f_k; f_{i_1;k}, \ldots, f_{i_j;k})\},$$

where $E(f_k) =: E(f_k; f_{i_1;k}, \ldots, f_{i_j;k}) = \{t \in I, \min_{r=i_1,\ldots,i_j} f_{r;k}(t) \le f_k(t) \le \max_{r=i_1,\ldots,i_j} f_{r;k}(t)\}$ and $\tilde{\lambda}(f_k) = \lambda(E(f_k))/\lambda(I)$ and λ is the Lesbegue measure on I. As stated in Lopez-Pintado and Romo (2009), the values of the modified band depth measure are stable with respect to the choice of J, and in order to be computationally faster we set $J = 2$ and we denote $MBD_n^J(\mathbf{f})$ as $MBD(\mathbf{f})$. Also, in the multivariate functional context the use of the modified band depth measure avoids the problem of ties.

Given the multivariate band depth measure defined in (3), a sample of multivariate functional data $\mathbf{f}_1, \ldots, \mathbf{f}_n$ can be ranked. We will denote $\mathbf{f}_{[i]}$ the sample curve associated with the i-th largest depth value, so $\mathbf{f}_{[1]} = \operatorname{argmax}_{\mathbf{f}\in\{\mathbf{f}_1,\ldots,\mathbf{f}_n\}} MBD(\mathbf{f})$ is the *median* (deepest and more central) curve, and $\mathbf{f}_{[n]} = \operatorname{argmin}_{\mathbf{f}\in\{\mathbf{f}_1,\ldots,\mathbf{f}_n\}} MBD(\mathbf{f})$ the most outlying one.

2.2. *Multivariate Functional Boxplot and Outlier Detection*

The idea of generalizing the concept of functional boxplot to multivariate functional data is based on the new definition of multivariate functional depth measure given in (3) which takes into account simultaneously the behaviour of all the s components of $\mathbf{f}$, and weighing the depth measures of the marginal components in a suitable way. The same idea is used when the goal is to carry out multivariate functional outlier detection. For example, to robustify a training set adopted in unsupervised classification algorithms (see Ieva et al., 2012). In fact, the definition of *multivariate functional outlier* is strictly connected to the definition of depth measure and, in particular, to the choice of the weights used to define the multivariate functional depth measure. In general, the meaning of the term *outlier* changes as long as the different types of data are considered. In our case, an outlier among ECG curves could be a signal very different from others in terms of phase/amplitude and morphology. In this case, the concept of outlier has to take into account outlyingness on all the components (i.e., the leads), eventually weighting them in different way according to the importance of each lead within the diagnosis process, since all its components are describing the same biological event (i.e., the heartbeat). That is why we define an average index such the one proposed in (1).

The following steps should be implemented on multivariate curves sample $\mathbf{f}_1, \ldots, \mathbf{f}_n$.

1. For each statistical unit j, $j = 1, \ldots, n$, compute the value of depth measure $MBD(\mathbf{f}_j)$;

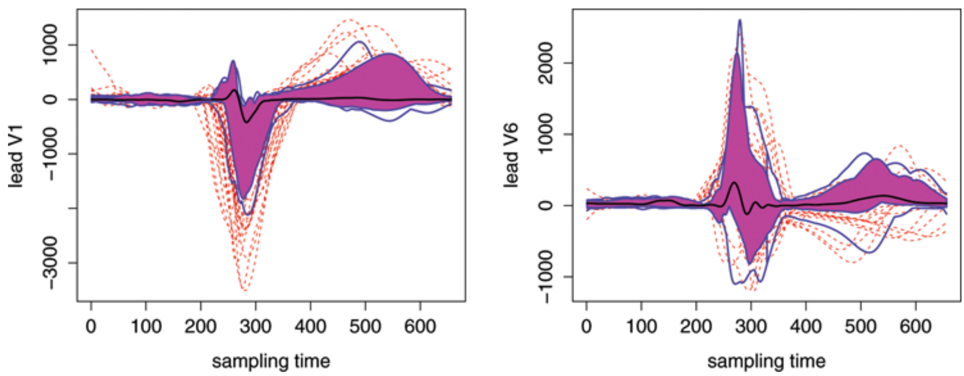

Figure 1. Functional boxplots (only leads V1 and V6 are represented) of 150 ECG traces. The central bands (grey area), fences (solid lines), and outliers (dotted lines) of each lead are defined as described in Sec. 2.2, according to the ranking induced by $MBD_n^J(\mathbf{f})$ defined in (3), and all the leads are weighted equally. (color figure available online)

2. Rank the multivariate functions $\mathbf{f}_j(t)$ according to the value of multivariate depth measure and define outliers as those curves that, for at least one t, are outside the fences obtained by inflating by h times the envelope defined by the $\alpha\%$ of the central region. In particular, the $\alpha\%$ central region for the component f_k determined by a sample of curves is defined as

$$\mathscr{C}_\alpha = \left\{ (t, y(t)) : \min_{r=1,\ldots,\lceil \alpha n \rceil} f_{[r];k}(t) \leq y(t) \leq \max_{r=1,\ldots,\lceil \alpha n \rceil} f_{[r];k}(t) \right\},$$

where $\lceil \alpha n \rceil$ is the smallest integer greater than or equal to αn. In the following, we set $\alpha\% = 50\%$ and $h = 1.5$.
3. Visualize the functional boxplot of each component, building the envelope of the 50% deepest functions and then the functional boxplot according to the ranking arising from the multivariate index previously pointed out.

Notice that this algorithm defines outliers according to the multivariate index of depth, which takes into account simultaneously the depth of all components of the multivariate function. This implies that the envelope of the central region is composed of the same $\alpha\%$ most central curves, with respect the multivariate index of depth, in each component (see also Lopez-Pintado and Romo, 2007).

An example of multivariate functional boxplot is shown in Fig. 1. The presence of a high number of outliers can be appreciated in the picture; in fact the 150 traces correspond to 100 physiological and 50 pathological units. So this explorative graphical tool calls for a statistical quantitative procedure to test differences between subgroups of data.

2.3. *Robust Statistics and Rank Test*

By taking the order of the sample curves to be the one induced by multivariate functional depth measure, the definition of *trimmed mean* in Fraiman and Muniz (2001) can be extended to multivariate functional data. We can also generalize to this framework a non parametric rank test to compare two samples of multivariate functions. In particular, consider a sample $\mathbf{f}_1, \ldots, \mathbf{f}_n$ generated according to a distribution $P_{\mathbf{X}}$ and another sample $\mathbf{g}_1, \ldots, \mathbf{g}_m$ generated according to a distribution $P_{\mathbf{Y}}$. We want to test differences between the two populations. Assume that there is a third reference sample, say $\mathbf{h}_1, \ldots, \mathbf{h}_N$, from one of the two populations, for example $P_{\mathbf{X}}$. Calculate the rank of each element of sample $\mathbf{f}_1, \ldots, \mathbf{f}_n$ and of the sample $\mathbf{g}_1, \ldots, \mathbf{g}_m$ with respect to the reference sample $\mathbf{h}_1, \ldots, \mathbf{h}_N$. In particular, let $R(P_N, \mathbf{f}_i)$ be the proportion of $\mathbf{h}_j$'s with MBD less than or equal to the MBD of $\mathbf{f}_i$, where the MBD is computed with respect to the reference sample $\mathbf{h}_1, \ldots, \mathbf{h}_N$. An analogous definition is assumed for $R(P_N, \mathbf{g}_i)$. Then order these values, $R(P_N, \mathbf{f}_i)$ and $R(P_N, \mathbf{g}_i)$, from the smallest to the highest giving them a rank from 1 to $n + m$. This induces a rank on the functions $\mathbf{f}_1, \ldots, \mathbf{f}_n, \mathbf{g}_1, \ldots, \mathbf{g}_m$. According to Liu and Singh (1993), we can apply the Wilcoxon test to the induced ranks. In particular, the lower the depth the lower the rank. The proposed test statistic R is the sum of the ranks of the second sample $R(P_N, \mathbf{g}_1), \ldots, R(P_N, \mathbf{g}_m)$. According to the null hypothesis (H_0) there are no differences between the distributions generating the data. Hence, $R(P_N, \mathbf{g}_1), \ldots, R(P_N, \mathbf{g}_m)$ can be viewed as a random sample size m drawn without replacement from the set $(1, \ldots, n + m)$, and we reject H_0 for values of R too small. For large values of n and m it is possible to use a Normal approximation (see Li and Liu, 2004). The presence of ties is treated as explained in Liu and Singh (1993) and Lopez-Pintado and Romo (2009). Such test represents a quantitative method for carrying out inference in a supervised multivariate functional clustering framework. On the other hand, in the unsupervised clustering case, it can also be seen as a way to test if the process generating the outliers, pointed out by the functional boxplot, can be considered as different from the process generating the curves of the $\alpha\%$ most central region.

3. An Application to ECG Signals

In Ieva et al. (2012), a statistical framework for analysis and classification of ECG curves starting from their sole morphology is proposed. The main goal of this article is to identify, from a statistical perspective, specific ECG patterns which could benefit from an early invasive approach. In fact, the identification of statistical tools capable of classifying curves using only their shape could support an early detection of heart failures, not based on usual clinical criteria. In order to do this, a real time procedure consisting of preliminary steps like reconstructing signals, wavelets denoising, and removing biological variability in the signals through data registration is tuned and tested. Then, a multivariate functional k-means clustering of reconstructed and registered data is performed. When testing the new procedures the performance of classification method is validated through cross validation. Hence, it is mandatory to find a suitable training of the algorithm on data. This would robustify classification algorithm and would improve reliability of prediction. The procedure proposed in the previous section is an effective way to reach this goal. In fact, it leads to select as training set the proportion of multivariate curves whose depth is high. Considering the ECG of the j-th patient as a 8-variate function $\mathbf{f}_j = (f_{j;1}, \ldots, f_{j;8})$, the $f_{j;k}$, $(k = 1, \ldots, 8)$ correspond to the eight leads I, II, V1, V2, V3, V4, V5, and V6. Then the procedure discussed in Sec. 2 is applied in order to carry out functional boxplots and to perform outlier detection for two different groups: physiological and pathological patients, i.e., people affected by a particular kind of heart disease, called Bundle Branch Block (BBB). It is easy to

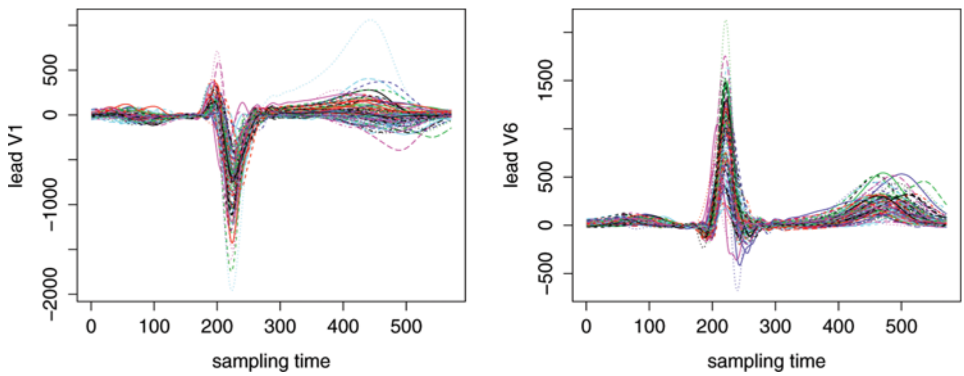

Figure 2. Row signals (leads V1 and V6) of the 100 physiological patients. (color figure available online)

detect this pathology through observing shape modifications of the ECG pattern; this disease divides in Right Bundle Branch Block (RBBB) and Left Bundle Branch Block (LBBB) according to the side of the heart it affects. In the following, we will consider a sample of 100 physiological signals and 50 pathological ones, where the latter come from patients affected by LBBB. In Figs. 2 and 3, the row data showing leads V1 and V6 of Normal and LBBB signals, respectively, are depicted. Figures 4 and 5 show the corresponding functional boxplots, (see Ieva et al., 2012 for details on statistical analysis and procedures). Functional boxplots are produced according to the ranking induced by the multivariate functional index where the weights p_k, $(k = 1, \ldots, 8)$ are all equal to 1/8, weighting all the leads equally.

Since there is a common ranking of all components of $\mathbf{f}_j$s, induced by the multivariate index of depth, as noticed in Sec. 2.2, the central band of the functional boxplot is defined by the same statistical units in each component. The multivariate functional index of depth defined in (3) takes jointly into account the order of each component (lead) of the multivariate function (ECG). This is the main and most important difference between functional boxplots reported in Figs. 4 and 5 and those we would have obtained by simply asking for functional boxplots of

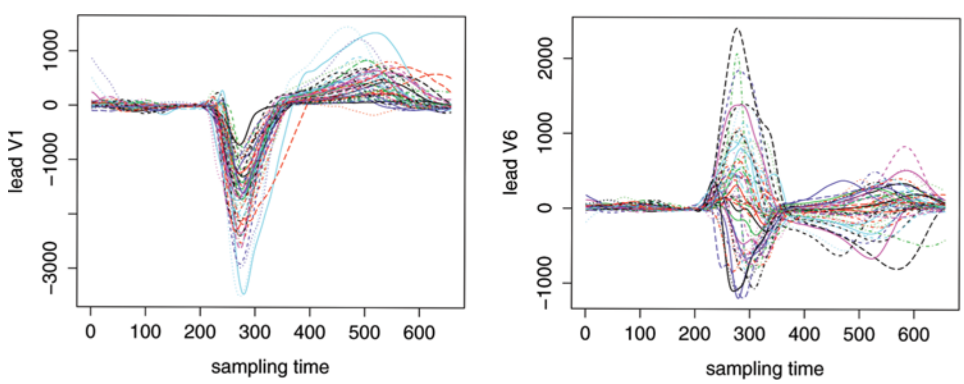

Figure 3. Row signals (leads V1 and V6) of the 50 pathological patients. (color figure available online)

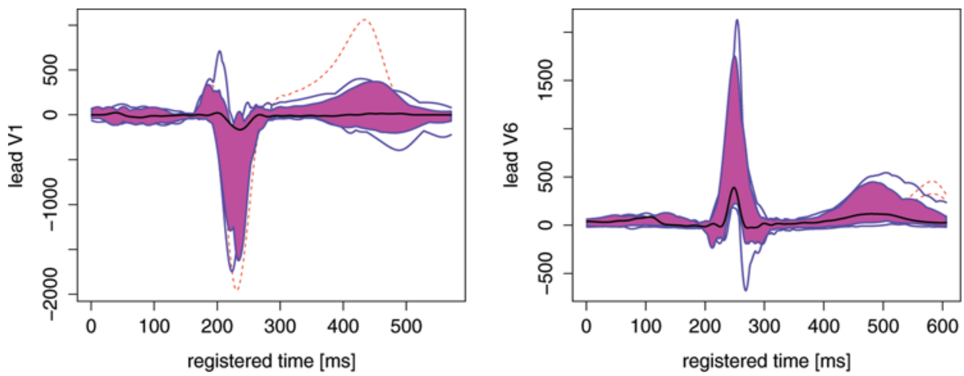

Figure 4. Functional boxplots of ECG leads V1 and V6 of the 100 physiological patients. The central bands (grey area), fences (solid lines), and outliers (dotted lines) of each lead are defined as described in Sec. 2.2, according to the ranking induced by $MBD_n^J(\mathbf{f})$ defined in (3) and using the same weight for each lead. (color figure available online)

each lead. This difference is confirmed and can be appreciated also considering the number of curves labeled as outlier in the first and in the second case, i.e., adopting procedures where functional boxplots are based on multivariate index of depth or taking lead-by-lead outliers, respectively. In our case, curves labeled as multivariate functional outliers are always a subset of those identified as outlier by the lead-by-lead outlier detection procedure. In general, taking jointly into account the depth of each component of the multivariate curve, the number of outlying curves decreases.

As described in Sec. 2.3, using the order of the sample curves induced by the multivariate functional depth measure, it is possible to generalize to this framework a non parametric rank test to compare two samples of multivariate functions. Actually, we adopt the rank test to check for differences in the underlying process generating the LBBB curves with respect to the physiological ones. To apply the

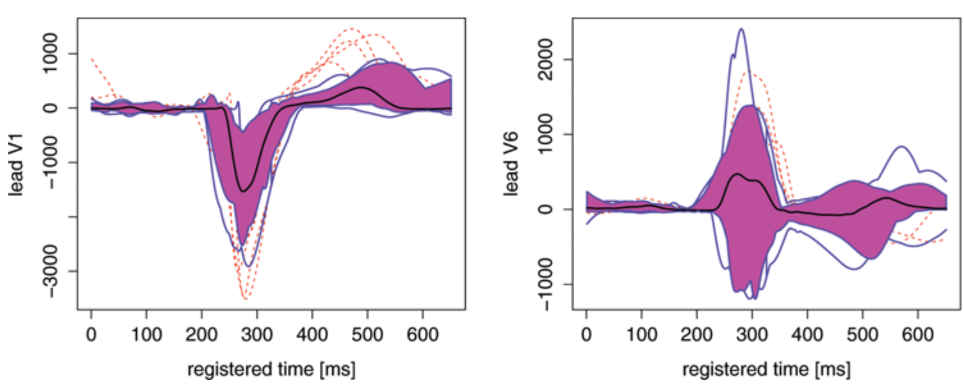

Figure 5. Functional boxplots of leads V1 and V6 of the 50 pathological (Left Bundle Brunch Block) ECGs. The central bands (grey area), fences (solid lines), and outliers (dotted lines) of each lead are defined as described in Sec. 2.2, according to the ranking induced by $MBD_n^J(\mathbf{f})$ defined in (3) and using the same weight for each lead. (color figure available online)

rank test, we considered 50 ECGs randomly chosen from the physiological traces as the reference group to compute the ranks of the remaining 50 physiological and 50 LBBB traces. The procedure has been repeated 20 times to avoid bias selection in the choice of the reference group. The average of the $p-$values of the tests is less than 10^{-16}. Therefore, we can conclude that there exist significant differences between the two groups. The statistical evidence remains very strong (average $p-$value still less than 10^{-16}) if we compute the depth measure (3) setting $(p_1, \ldots, p_8)$ equal to $(1/10, 1/10, 2/10, 1/10, 1/10, 1/10, 1/10, 2/10)$, stressing the weight of leads V1 and V6, since they are the most important for carrying out the LBBB diagnosis, as confirmed by cardiologists. That is, there is a strong evidence that the LBBB signals arise from a different latent process.

4. Conclusions

In this work, we generalize the notion of depth for functional data presented in Lopez-Pintado and Romo (2009) to the multivariate functional case and we also define a new multivariate functional index of depth which is able to jointly take into account the depth of the multivariate functional data on each component. This provides a center-outward ordering criterion for a sample of multivariate functions. Extensions and proofs of the statistical properties of the new index are also provided, as well as for the modified version. A generalization of the nonparametric test to this framework has been adopted to carry out inference in a supervised clustering context. Finally, the application of the new index to a real case of ECG signals has been presented and discussed, highlighting how the methodology works effectively both in detecting outliers and in distinguishing between samples arising from different underlying latent processes. Further developments of this work will be focused on the study of ECG traces arising from numerical simulations. We will assess if a simulated multivariate signal can be considered as coming from a reference population of signals, computing its depth in the way described above. This will be a new method for validating numerically simulated ECGs from a statistical perspective.

Appendix

Proof of Proposition 2.1. Part (a). Using Definition 2.1 and the property (1) of Theorem 3 in Lopez-Pintado and Romo (2009) we have

$$BD^J_{P_{T(\mathbf{X})}}(T(\mathbf{f})) = \sum_{k=1}^{s} p_k BD^J_{P_{A_{kk}X_k+b_k}}(A_{kk}f_k + b_k) = \sum_{k=1}^{s} p_k BD_{X_k}(f_k) = BD^J_{P_{\mathbf{X}}}(\mathbf{f}).$$

The diagonality requirement on matrix A means that the multivariate functional depth measure $BD^J_{P_{\mathbf{X}}}(\mathbf{f})$ is invariant as regards affine transformations of each component taken one by one, without combining different elements of the multivariate function.

Part (b) follows directly from property (2) of Theorem 3 in Lopez-Pintado and Romo (2009).

Part (c)

$$\sup_{\min_{k=1,\ldots,s}\|f_k\|_\infty > M} BD^J_{P_{\mathbf{X}}}(\mathbf{f}) = \sup_{\min_{k=1,\ldots,s}\|f_k\|_\infty > M} \sum_{k=1}^{s} p_k BD_{X_k}(f_k)$$

where each term in the sum is summed over components that go to zero as M goes to infinity.

Part (d) this point also follows directly from property (4) of Theorem 3 in Lopez-Pintado and Romo (2009).

Proof of Proposition 2.2.

$$\begin{aligned} \left|BD^J_n(\mathbf{f}) - BD^J_{P_{\mathbf{X}}}(\mathbf{f})\right| &= \left|\sum_{k=1}^{s} p_k BD^J_{n,k}(f_k) - \sum_{k=1}^{s} p_k BD_{X_k}(f_k)\right| \\ &\leq\leq \sum_{k=1}^{s} p_k \left|BD^J_{n,k}(f_k) - BD_{X_k}(f_k)\right| \end{aligned} \tag{4}$$

and each term of the sum in the last term of (4) goes to zero as stated in Theorem 4 of Lopez-Pintado and Romo (2009). Also, the uniformity of convergence on the equicontinuous sets is a straightforward extension of the univariate functional case.

Acknowledgments

This work is within PROMETEO (PROgetto sullarea Milanese Elettrocardiogrammi Teletrasferiti dallExtra Ospedaliero). The authors wish to thank Mortara Rangoni Europe s.r.l. for having provided data.

References

Cuevas, A., Febrero, M., Fraiman, R. (2007). Robust estimation and classification for functional data via projection-based depth notions. *Computat. Statist.* 22:481–496.

Fraiman, R., Meloche, J. (1999). Multivariate L-estimation. *Test* 8:255–317.

Fraiman, R., Muniz G. (2001). Trimmed means for functional data. *Test* 10:419–440.

Ieva, F. (2011). Outlier detection for training sets in an unsupervised functional classification framework: An application to ECG signals. *Proc. 17th EYSM*. Portugal: Lisbon.

Ieva, F., Paganoni, A. M., Pigoli, D., Vitelli, V. (2012). Multivariate Functional Clustering for the Morphological Analysis of ECG Curves. *J. Roy. Statist. Soc. Ser. C (Appl. Statist.)*. Prepublication version online DOI: 10.1111/j.1467-9876.2012.01062.x.

Li, J., Liu, R. (2004). New nonparametric tests of multivariate locations and scales using data depth. *Statist. Sci.* 19:686–696.

Liu, R. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* 18: 405–414.

Liu, R., Singh, K. (1993). A quality index based on data depth and multivariate rank tests. *J. Amer. Statist. Assoc.* 88(421):252–260.

Lopez-Pintado, S., Romo, J. (2006). Depth-based classification for functional data. In: Liu, R. Y., Serfling, R., Souvaine, D. L. Eds., *Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications*. Dimacs Series in Discrete Mathematics and Theoretical Computer Science. Vol. 72. American Mathematical Society, pp. 103–120.

Lopez-Pintado, S., Romo, J. (2007). Depth-based inference for functional data. *Computat. Statist. Data Anal.* 51(10):4957–4968.

Lopez-Pintado, S., Romo, J. (2009). On the concept of depth for functional data. *J. Amer. Statist. Assoc.* 104(486):718–734.

Serfling, R. (2006). Depth functions in nonparametric multivariate inference. In: Liu, R. Y., Serfling, R., Souvaine, D. L. Eds., *Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications*. Dimacs Series in Discrete Mathematics and Theoretical Computer Science. Vol. 72. American Mathematical Society.

Sun, Y., Genton, M. G. (2011). Functional boxplots. *J. Computat. Graph. Statist.* 20:316–334.

Sun, Y., Genton, M. G., Nychka, D. W. (2012). Exact fast computation of band depth for large functional datasets: How quickly can one million curves be ranked? *Statistics* 1:68–74.

Tukey, J. (1975). Mathematics and picturing data. *Proc. 1975 Int. Congr. Math.* 2:523–531.

Zuo, Y., Serfling, R. (2000). General notions of statistical depth function. *Ann. Statist.* 28(2):461–482.

Zuo, Y. (2003). Projection based depth functions and associated medians. *Ann. Statist.* 31:1460–1490.