

INTEGRATED DEPTH FOR TRAJECTORIES OF AIRBORNE MICROORGANISMS TO ANTARCTICA

BY LUCAS FERNANDEZ-PIANA^{1,a}, ANA JUSTEL^{2,b}  AND MARCELA SVARC^{3,c} 

¹*Departamento de Matemática y Ciencias, Universidad de San Andrés, afpiana@udesa.edu.ar*

²*Department of Mathematics, Universidad Autónoma de Madrid, ana.justel@uam.es*

³*Departamento de Matemática y Ciencias, Universidad de San Andrés and CONICET, msvarc@udesa.edu.ar*

Sets of trajectories that begin or end at the same point, in the form of a bouquet, appear in several real-world problems, such as the dispersion of volcanic ash or forecasts of hurricane paths, among others. Our interest in this type of trajectory focuses on studying the biogeography of airborne microorganisms and their ability to colonise soils recently exposed due to climate change. For these functional data, we introduce a new integrated depth measure ($\mathcal{D}$) that allows finding central and outlier curves in a dataset. First, circular local depths (LD) are calculated in concentric circles around the common point, and in a second step, these values are integrated along the curves, yielding the trajectory depth. Under mild conditions both LD and $\mathcal{D}$ have good properties and are strongly consistent. In addition, we propose an efficient algorithm for working with large datasets. Finally, we apply this new technique to find the main routes followed by air masses carrying microorganisms to Byers Peninsula (Livingston Island, Antarctica).

1. Introduction. Spread direction of a forest fire or the ash from an erupting volcano, identification of areas potentially affected by radiation in a nuclear accident, or hurricane paths that have historically hit a city are some examples of relevant problems in which the data sets are trajectories followed over time (Srinivas et al. (2014), Mirzargar, Whitaker and Kirby (2014), Su et al. (2014)). All these sets have in common that the trajectories describe the position of air parcels or particles, and they all have the same initial or end point. They are also characterised by spreading out in all directions from the common point. Typically, each air parcel or particle travels at a different speed, so the distance travelled in each time interval can vary considerably. We call to this type of set a *bouquet of trajectories*.

In the aerobiology literature, we also find bouquets of trajectories in the simulation of air masses to study the origin or propagation of airborne microorganisms. Galban et al. (2021) analysed five-day back-trajectories transporting microorganisms to the Iberian Peninsula simulated with the NOAA's HYSPLIT modeling system (Stein et al. (2015)). Just by observing the graph in the Figure 1A, different origins can be recognized from Europe, America, or the North Atlantic Ocean. However, visual inspection is inadequate when trying to use the simulated air masses trajectories with HYSPLIT in more global microbial biogeography problems.

Since the 1950s, the Antarctic Peninsula has experienced one of the largest temperature increases on Earth. This warming has contributed to the retreat of glaciers in the region. New ice-free terrains allow the establishment of airborne microorganisms and pave the way for new exotic species that can transform Antarctic terrestrial habitats in a context of climate change. In the study of the circulation patterns of air masses that transport microorganisms to a specific location in Antarctica, the most relevant characteristics of the bouquet cannot be retrieved solely by visual inspection of the map, even with a small fraction of the possible back-trajectories (Figure 1B). Therefore, as in many other statistical problems, summarizing

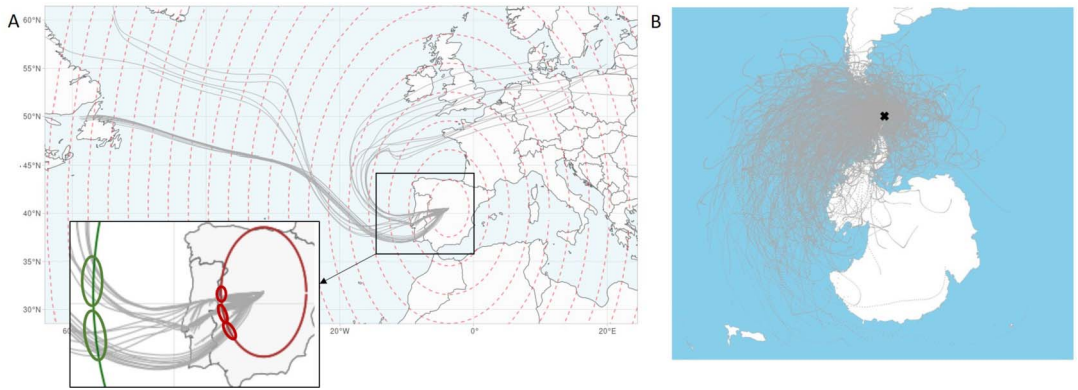


FIG. 1. A Simulated five-day back-trajectories of the air masses arriving at a specific location in Madrid on 23 and 24 May 2018 (in grey). Concentric circles around the end point of the trajectories (in dashed lines). B Bouquet of 600 simulated five-day back-trajectories of air masses with common end point, marked with a black cross, at Byers Peninsula (Livingston Is., South Shetland Is., Antarctica). Arriving at this location between 2005 and 2016.

the trajectory data sets will be the first step to start extracting the information. We consider that the most central trajectories will provide us with valuable information to understand the origin and vulnerability of microbial communities in polar regions. We will attempt to obtain an accurate description of the central trajectories using nonparametric statistics that do not require assumptions about the data distributions. In the case of multivariate or infinite-dimensional data, depth measures play a key role in exploratory data analysis, providing a center-outer ordering of the data that extends the concept of the median.

The aim of this work is to propose a new depth measure for a bouquet of trajectories that takes advantage of their particular geometry. This approach should provide us with new insights into the dispersal of microorganisms in Antarctica as well as being applicable in other studies.

In Section 2 we define new depths for bouquets of trajectories, and we also define a new local depth in $\mathbb{S}^1$. Section 3 is devoted to the numerical implementation of the depths on real and simulated data. Section 4 shows the application of the new tools in identifying the transport patterns of airborne microorganisms to a location in the Antarctic Peninsula region. Concluding remarks are given in Section 5. The study of the properties and asymptotic behaviour of the local depth in $\mathbb{S}^1$ as well as all the proofs are given in Supplementary Material A (Fernandez-Piana, Justel and Svarc (2025a)).

2. Depth for Bouquet of trajectories. A bouquet of trajectories is a set of plane random curves $\Gamma(t)$ with common point of origin (or destination) and finite length. Although trajectories are functional data, existing definitions of depths (Cuesta-Albertos and Nieto-Reyes (2008), Cuevas and Fraiman (2009), López-Pintado and Romo (2011), Dai and Müller (2018), Harris et al. (2021), among others) are not always appropriate since the curves move in any direction from the common point. Furthermore, since all trajectories share the same start/end point, in some regions all trajectories are very close to each other.

As our main interest lies in the position of the 2D curves in the plane independently of the speed at which they move, it is natural to express the trajectories in polar coordinates taking advantage of the common point and the dispersal in any direction. This will allow us to extend to our context the idea behind integrated depths, where the infinite-dimensional problem is reduced to the calculation of depths in low-dimensional projections (Fraiman and Muniz (2001)). The depths will only be measured on the perimeter of concentric circles around the common point and then integrated. An example of such circles can be observed in Figure 1 A. This approach simplifies the calculation of depths and saves computation time.

2.1. Integrated circular depth. Using polar coordinates with reference point the common begin or end, we can express the trajectories of finite length in a bouquet as $\Gamma = (\Phi, \Upsilon) \in C([0, 1], ([0, 2\pi) \times [0, R]))$ with probability measure $\mathcal{P}$. In this notation, $\Phi(t)$ is the direction or angle of the trajectory position at time t with respect to the zero initial direction in counterclockwise orientation, $\Upsilon(t)$ is the distance or radius to the reference point, $\Gamma(0) = (0, 0)$, and $R > 0$ is such that the maximum distance of $\Upsilon(t)$ to $\Gamma(0)$ is smaller than or equal to R .

Assuming that Υ is strictly increasing a.s., for each $r \in [0, R]$ there exists a unique t_r such that $\Upsilon(t_r) = r$ a.s. Then P_r will denote the distribution of the random variable $\Phi(t_r)$, which is the projection of the trajectory in the circle of radius r . The trajectories do not depend on the time parametrization (see Supplementary Material A, [Fernandez-Piana, Justel and Svarc \(2025a\)](#)).

We define a depth measure for a bouquet of trajectories as follows.

DEFINITION 2.1. The integrated circular depth, $\mathcal{D}(\cdot, \mathcal{P}) : C([0, 1], [0, 2\pi) \times [0, R]) \rightarrow [0.5, 1]$, is defined as

$$(2.1) \quad \mathcal{D}(\gamma, \mathcal{P}) = \int_0^R w(r) D(\phi(r), P_r) dr,$$

where $D(\phi(r), P_r)$ is a depth measure in $\mathbb{S}^1$ and $w(\cdot)$ is a nonnegative weight function such that $\int_0^R w(r) dr = 1$.

The univariate depth function D evaluates the centrality of an angle $\phi(r)$ with respect to the distribution P_r . Lower values of D indicate that $\phi(r)$ is a potential outlier with respect to P_r , while higher values of D indicate that $\phi(r)$ is central. The weight function w could be chosen constant or having less weight at the begin of the trajectories, where they usually intertwine, or may depend on the distribution P_r , as indicated by [Nagy et al. \(2016\)](#).

The empirical counterpart of $\mathcal{D}$ is obtained by plug-in, that is, given $\gamma_1 = (\phi_1, \nu_1), \dots, \gamma_n = (\phi_n, \nu_n)$, a random sample of the process Γ ,

$$(2.2) \quad \mathcal{D}_n(\gamma, P_n) = \int_0^R w(r) D_n(\phi(r), P_{n,r}) dr,$$

where P_n and $P_{n,r}$ are the empirical distribution of P and the conditional distribution P_r , respectively.

In the definition of the integrated circular depth, we note that D is the building block of $\mathcal{D}$. This means that, under regular conditions, as far as classical depth properties hold for D , they are inherited by $\mathcal{D}$. This follows from the application of the general theory introduced by [Nagy et al. \(2016\)](#) bridging univariate depths with integrated depths for functional data. This fact applies to measurability, consistency, and the classical properties ([Zuo and Serfling \(2000\)](#)): [P.1] the depths of points are invariant under affine transformations; [P.2] given a proper symmetry definition and a centre which is a symmetry point, then the depth function attains its maximum value at that point; [P.3] depth values decrease as the distance from the centre increases; [P.4] is the property of vanishing at infinity, which makes no sense for D since $\mathbb{S}^1$ is a compact space. In addition, properties [P.5], continuous as a function of ϕ , and [P.6], continuous as a functional of P_r , are desirable.

There are several proposals for depth measures on circular or directional data, mostly extending the classical ideas of univariate depths to the geometry of the circle. The first definitions of depth measures for circular data date from the 1990s, when [Liu and Singh \(1992\)](#) extended the concepts of simplicial and half-space depth, at the time they proposed the arc distance depth. They defined the median in the context of circular data and proved some of the classical depth properties, such as the maximisation at the centre and the decay along

rays. For unimodal symmetric distributions, the three proposals coincide. The main drawback of these depths is the high computational cost. The absence of a well-defined zero direction and a single sense of rotation in circles and spheres led [Agostinelli and Romanazzi \(2013\)](#) to propose a nonparametric approach as a natural framework to describe centrality, dispersion, and outliers in this type of data. Other alternatives fall into the category of distance-based depths, such as the angular Mahalanobis depth ([Ley, Sabbah and Verdebout \(2014\)](#)) and the cosine, chord, and arc length distances defined by [Pandolfo, Paindaveine and Porzio \(2018\)](#).

2.2. Integrated circular local depth. All the previous proposals have in common their good performance for unimodal and symmetric distributions and can be used as D in equation (2.1). However, they are not suitable for multimodal or skewed data, such as that shown for several of the circles in Figure 1A. In this case there are some circles in which the trajectories are concentrated in separate groups. The points of the trajectories that cross the external circle present two modes, while the points of the trajectories that cross the inner circle have three modes. The possibility of dealing with this mixed behavior will also be considered in the calculation of the depth restricted to circles.

We propose to measure locally the depth in the circles to reveal possible multimodality in the underlying distribution following the ideas in the most recent literature on local depths. We define the circular depth by restricting it to a local neighbourhood of the point. The main idea is that a local depth should behave as a global depth in the neighbourhood of each point. There are several proposals in the literature, most of them using the idea of restringing a global depth to a local environment ([Agostinelli and Romanazzi \(2011\)](#), [Agostinelli and Romanazzi \(2013\)](#) and [Paindaveine and Van Bever \(2013\)](#)). Another approach is to define the depth with nonparametric kernel density estimates, which are suitable for retrieving the several modes in the distribution ([Cuevas, Febrero and Fraiman \(2007\)](#), [Sguera, Galeano and Lillo \(2014\)](#), [Sguera, Galeano and Lillo \(2016\)](#)). Finally, [Fernandez-Piana and Svarc \(2022\)](#) propose an integrated local depth for data on Banach spaces and also analyse how to adapt classical depth properties to local contexts. All these proposals are for multivariate or functional data and need to be adapted for circular data.

Inspired by the ideas of [Paindaveine and Van Bever \(2013\)](#) and [Fernandez-Piana and Svarc \(2022\)](#), we provide a new definition of a local depth in the circle based on the quantile function. Previous to define a circular local depth, we adopt the convention of angles $[0, 2\pi)$. Let Θ be a random variable in $\mathbb{S}^1$ with distribution P with zero initial direction, counterclockwise orientation, and cumulative distribution function F_0 . For any $\theta \in [0, 2\pi)$, we denote F_θ to the cumulative distribution function of $\Theta_\theta = \Theta - \theta$. Hence, for $\lambda \in [0, 2\pi)$, $P(0 \leq \Theta_\theta \leq \lambda) = F_\theta(\lambda)$. The quantile function is $Q(u, F_\theta) = \inf\{x : F_\theta(x) \geq u\}$ for $0 < u < 1$.

DEFINITION 2.2. The circular local depth at locality level $\beta \in (0, 1]$ is a function $LD^\beta(\cdot, P) : [0, 2\pi) \rightarrow [0.5, 1]$, given by

$$(2.3) \quad LD^\beta(\theta, P) = \frac{\pi^2}{\pi^2 + \lambda_{\beta, \theta}^+ \lambda_{\beta, \theta}^-},$$

where

$$(2.4) \quad \lambda_{\beta, \theta}^+ = Q(\beta/2, F_\theta) \quad \text{and}$$

$$(2.5) \quad \lambda_{\beta, \theta}^- = 2\pi - Q(1 - \beta/2, F_\theta).$$

The definition of $LD^\beta(\theta, P)$ does not depend on the circle orientation since if it is oriented in the clockwise direction $\lambda_{\beta, \theta}^+$ would be given by equation (2.5) and $\lambda_{\beta, \theta}^-$ by equation (2.4).

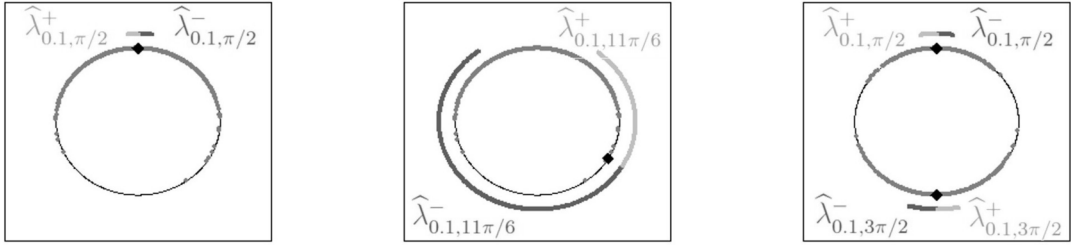


FIG. 2. Left and centre panels: 1000 observations generated with a von Mises distribution with centre at $\pi/2$ and concentration parameter $\kappa = 10$. The clockwise arcs are in dark grey, and the counterclockwise arcs are in light grey. The left (centre) panel shows the arcs for the calculation of $LD^{0.1}$ at $\theta = \pi/2$ ($\theta = 11\pi/6$). Right panel: 1000 observations generated with a mixture of von Mises distributions with centres at $\pi/2$ and $3\pi/2$ and concentration parameters $\kappa = 10$ and $\kappa = 5$, respectively, the mixing fraction is 0.5. Using the same colour key, we show the arcs for $\theta = \pi/2$ and $3\pi/2$ for the calculation of $LD^{0.1}$.

In light of Definition 2.2, it can be seen that the depth of θ is inversely proportional to the square geometric mean of $\lambda_{\beta,\theta}^+$ and $\lambda_{\beta,\theta}^-$. These are the shortest arcs that enclose a probability $\beta/2$ to the right and to the left of θ . Since $\lambda_{\beta,\theta}^+$ and $\lambda_{\beta,\theta}^-$ are bounded between 0 and 2π and also $\lambda_{\beta,\theta}^+ + \lambda_{\beta,\theta}^- \leq 2\pi$, then it is easy to see that $0.5 \leq LD^\beta(\theta, P) \leq 1$.

When the distribution P is absolutely continuous, it is immediate to see that for $\beta_1 < \beta_2$, then $\lambda_{\beta_1,\theta}^+ \leq \lambda_{\beta_2,\theta}^+$ and $\lambda_{\beta_1,\theta}^- \leq \lambda_{\beta_2,\theta}^-$.

To illustrate the concept, we generate 1000 random observations according to a von Mises distribution with centre at $\pi/2$ and concentration parameter $\kappa = 10$ (Figure 2 left and centre panels). For computing the circular local depth, we use the empirical version of Definition 2.2

$$LD^{\beta(n)}(\theta, P_n) = \frac{\pi^2}{\pi^2 + \hat{\lambda}_{\beta(n),\theta}^+ \hat{\lambda}_{\beta(n),\theta}^-},$$

where P_n is the empirical distribution of P and

$$\begin{aligned} \beta(n) &= \frac{\lfloor n\beta/2 \rfloor}{n}, \\ \hat{\lambda}_{\beta(n),\theta}^+ &= Q(\beta(n), F_{n,\theta}), \\ \hat{\lambda}_{\beta(n),\theta}^- &= 2\pi - Q(1 - \beta(n), F_{n,\theta}). \end{aligned}$$

When we compute the local depth with $\beta = 0.1$ at the central symmetric point $\pi/2$ and at $11\pi/6$ in the tail of the distribution, $\hat{\lambda}_{0.1,\pi/2}^+$ and $\hat{\lambda}_{0.1,\pi/2}^-$ (Figure 2 left) are smaller than $\hat{\lambda}_{0.1,11\pi/6}^+$ and $\hat{\lambda}_{0.1,11\pi/6}^-$ (Figure 2 centre). Moreover, $\hat{\lambda}_{0.1,\pi/2}^+$ and $\hat{\lambda}_{0.1,\pi/2}^-$ are equal, while $\hat{\lambda}_{0.1,11\pi/6}^+$ is smaller than $\hat{\lambda}_{0.1,11\pi/6}^-$. Then $\pi/2$ is deeper than $11\pi/6$, as expected.

The Figure 2 right panel shows a sample of 1000 observations generated following a bi-modal distribution, a mixture of von Mises distributions with centres at $\pi/2$ and at $3\pi/2$, concentration parameters 10 and 5, and mixing ratio 0.5. We show the arcs around the centres needed to compute the circular local depth with $\beta = 0.10$. Since both centres are symmetry points, the lengths of the clockwise and counterclockwise arcs are the same for each point (Figure 2 right). As the data is more concentrated at $\pi/2$ than at $3\pi/2$, its arcs are shorter.

It is important to note that the new circular local depth in Definition 2.2 enjoys the previously mentioned properties [P.1]–[P.3] and [P.5]–[P.6], properly adapted to local depth (Fernandez-Piana and Svarc (2022)) and circular data. The invariance is rotational, and the maximality at the centre is local and depends on a β -symmetry definition. A formal statement and proofs of these properties appear in the Supplementary Material A (Fernandez-Piana,

Justel and Svarc (2025a)) as well as the measurability and strong consistency theorems and proofs.

In the case that the circular local depth is used as D in the equation (2.1), we define the *integrated circular local depth* at locality level β as

$$(2.6) \quad \mathcal{LD}^\beta(\gamma, \mathcal{P}) = \int_0^R w(r) LD^\beta(\phi(r), P_r) dr.$$

The empirical counterpart $\mathcal{LD}_n^{\beta(n)}(\gamma, P_n)$ is defined straightforward, as in equation (2.2).

A key consequence of proving the measurability and consistency of LD^β is that the properties that hold for LD^β are inherited by $\mathcal{LD}^\beta$. This relevant result, established by Nagy et al. (2016), was given for depth measures, and it can be extended to local depths, provided that they are continuous and bounded. Then it holds that, LD^β is invariant under rotations on Φ , LD^β is continuous as a function of $\gamma \in \Gamma$, and is weakly continuous as a functional of $\mathcal{P}$. For the maximality at the centre, we propose the next symmetry definition for a bouquet of trajectories.

DEFINITION 2.3. The random element $\Gamma = (\Phi, \Upsilon)$ is β – symmetric in γ if $\Phi(r)$ is β – symmetric in $\phi(r)$ for almost all r .

Let $\gamma^* \in \Gamma$ such that $\gamma^*(\theta, r) = \gamma^*(\tilde{\theta}_r, r)$. For every $r \in [0, R]$ and $\Theta_r \in \mathbb{S}^1$ satisfies that $\tilde{\theta}_r$ (resp., $\tilde{\theta}_r + \pi$) is the unique mode (resp., antinode), then

$$\mathcal{LD}^\beta(\gamma^*, \mathcal{P}) = \sup_{\gamma \in \Gamma} \mathcal{LD}^\beta(\gamma, \mathcal{P}).$$

Hence, [P.2] and [P.3] hold. Measurability of $\mathcal{LD}^\beta$ and that $\mathcal{LD}_n^{\beta(n)}$ is a consistent estimate of $\mathcal{LD}^\beta$ also holds (see Supplementary Material A, Fernandez-Piana, Justel and Svarc (2025a)).

3. Implementation. For the implementation of the new depths for bouquet of trajectories, it is necessary to address some practical issues about the weight function and the grid as well as about the locality parameter for the integrated circular local depth. We propose an automatic procedure for β selection, suggest how to fix the grid, and choose appropriate weights to compute the equation (2.2).

3.1. Selection of the locality parameter. Local depth depends on a locality parameter, which must be set to reveal different features of the data. In general, small values of β correspond to local depths close to 1 for large sets. This makes the information provided not very useful. Conversely, when β is close to 1, the local neighbourhoods have high probability, and local features are ignored. For this reason it is crucial to determine a level of locality that allows local structures to be captured.

To illustrate this fact, we generated 2000 random observations from five different distributions (Figure 3, top panel):

M1 $\Theta \sim VM(\mu = \pi/2, \kappa = 10)$.

M2 $\Theta \sim 0.5VM(\mu = \pi/2, \kappa = 10) + 0.5VM(\mu = 3\pi/2, \kappa = 10)$.

M3 $\Theta \sim 0.7VM(\mu = \pi/2, \kappa = 10) + 0.3VM(\mu = 3\pi/2, \kappa = 5)$.

M4 $\Theta \sim 0.5VM(\mu = \pi/2, \kappa = 5) + 0.5VM(\mu = \pi, \kappa = 5)$.

M5 $\Theta \sim 0.3VM(\mu = \pi/2, \kappa = 0.5) + 0.7VM(\mu = 5\pi/4, \kappa = 2)$.

M1 is a unimodal von Mises distribution, while M2, M3, and M4 are mixtures of von Mises distributions, with mixing and concentration parameters equal in the cases of M2 and M4 and different in the case of M3. M5 is an asymmetric distribution. For each model, we

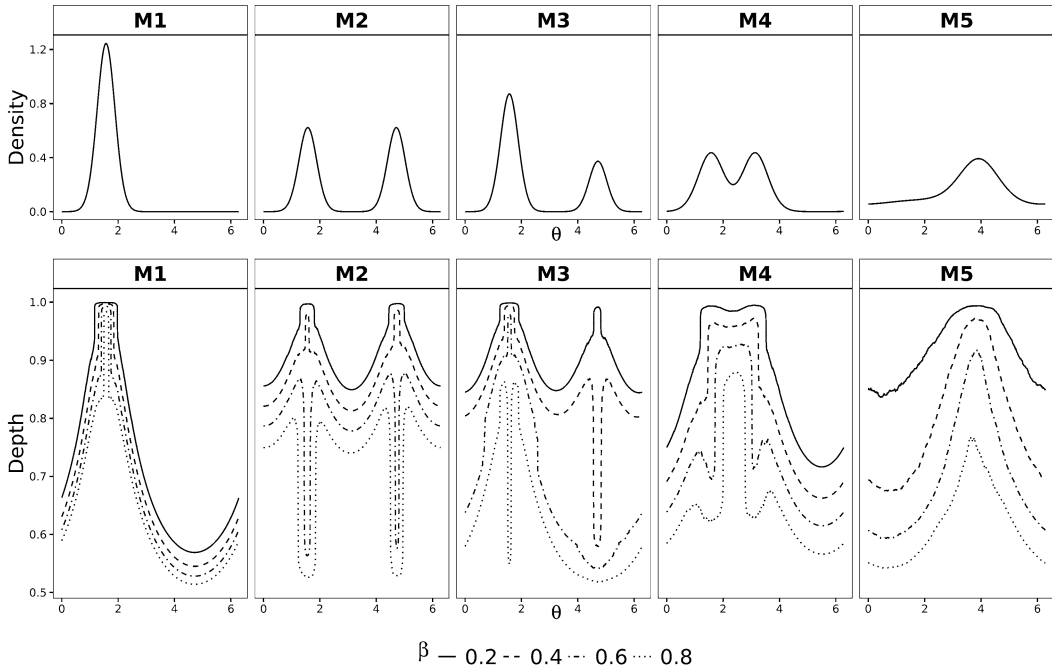


FIG. 3. Top panel: Density functions for M1 to M5. Bottom panel: Circular local depths computed at locality levels $\beta = 0.2, 0.4, 0.6$, and 0.8 for M1 to M5.

computed LD^β with different values of β (0.2, 0.4, 0.6, and 0.8). The results are shown in Figure 3 (bottom panel). As expected, for model M1 the LD^β behaves as a global depth for all values of β , reaching a maximum at the symmetry point. A similar pattern is shown for M5. In the case of mixture models M2, M3, and M4, when the value of β is less than the mixing proportion, the local depth also reaches local maxima in the modes. However, if β is greater than the mixing parameter, the local depth breaks down and has local minima in the modes. In M3 the maximum and/or minimum for each mode appears at different locality levels, depending on the mixing proportions. For M4 the pattern is harder to see than for M2 because the modes are closer together.

Inspired by these examples, we propose a data-driven algorithm to select an appropriate value of β so that LD^β is both informative and able to capture local features. The main idea is that an abrupt change in the function $LD^\beta(\theta)$ indicates that a local mode is no longer being considered as a local central point. The Algorithm 1 is designed to detect the maximum value of β before this happens. The first step is to compute $LD^\beta(\theta)$ in a grid of locality levels $\beta_0(n) < \dots < \beta_q(n)$, for all $\theta_1, \dots, \theta_n$ being a random sample in $\mathbb{S}^1$. In the second step, the differences between the depths for two consecutive values of β are calculated at each θ_i point, and the ordered sequence $sd = (sd_1, \dots, sd_q)$ is obtained, where sd_j is the standard deviation of $\{LD^{(\beta_{j+1}(n))}(\theta_i, P_n) - LD^{(\beta_j(n))}(\theta_i, P_n); i = 1, \dots, n\}$. In the final step, Pettit's test (Verstraeten et al. (2006)), with significance level α , is used to estimate the change point in the series sd . If the test does not reject the null hypothesis of no change point in sd , then any locality parameter can be chosen. For simplicity, we suggest $\beta = 0.2$. For local depth measures for other type of data, 0.2 is the usual recommendation (Paindaveine and Van Bever (2013), Fernandez-Piana and Svarc (2022)).

When we apply Algorithm 1 to automatically select the locality parameter to compute LD^β on each of the five datasets in Figure 3, we obtain $\beta = 0.45$ for M1, M4, and M5, $\beta = 0.49$ for M2, and $\beta = 0.32$ for M3. In all cases we use the input grid $0.01, 0.02, \dots, 0.99$ and the p-values of the Pettit test smaller than $\alpha = 0.1$.

Algorithm 1: β selection

Input: $0 < \beta_0(n) < \dots < \beta_q(n) < 1, \alpha$
Output: β_{opt}
Data: $\theta = \{\theta_1, \dots, \theta_n\}$

```

1 for  $j \leftarrow 0$  to  $q$  do
2   for  $i \leftarrow 1$  to  $n$  do
3      $ld(j, i) = LD^{\beta_j(n)}(\theta_i, P_n)$ 
4 for  $j \leftarrow 1$  to  $q$  do
5    $differ_j \leftarrow ld(j, \cdot) - ld(j-1, \cdot)$ 
6    $sd_j \leftarrow \text{var}(differ_j)^{1/2}$ 
7  $pvalue \leftarrow \text{pettitt.test}(sd)$ 
8 if  $pvalue \leq \alpha$  then
9    $\beta_{\text{opt}} \leftarrow \text{pettitt.test}(sd)(\text{change.point})$ 
10 else
11    $\beta_{\text{opt}} \leftarrow 0.2$ 
    
```

3.2. Grid and weights selection. One of the outstanding characteristics of the bouquet is that the trajectories can have different lengths and, therefore, different maximum distances to the origin. The definition of the grid of concentric circles in which the depths will be calculated and the choice of weights in equation (2.1) will play a predominant role in handling this characteristic in the implementation of the integrated circular local depth in datasets. We propose to give more weight to the depths calculated in the radii crossed by more curves. In this way the weights will usually decrease for larger radii where fewer curves reach.

Given $\gamma_1 = (\phi_1, v_1), \dots, \gamma_n = (\phi_n, v_n)$ a random sample of the process Γ , where trajectory γ_i is observed at distances $v_i^1, \dots, v_i^{l_i}$ from the origin/end point, $\Upsilon = \{v_1^1, \dots, v_1^{l_1}, \dots, v_n^1, \dots, v_n^{l_n}\}$ is the set of all radii in which at least one trajectory was observed in the corresponding concentric circle. For the empirical integration of the circular local depth in $\mathcal{LD}^\beta$, we suggest the grid $r_{[1]} < \dots < r_{[q]}$, where $r_{[i]}$ are quantiles of Υ . The number of concentric circles q should be smaller than $\#(\Upsilon) = \sum_{i=1}^n l_i$; for large $\#(\Upsilon)$, we suggest taking $q = 100$. For this grid the empirical probability that a radius belongs to $(r_{[k]}, r_{[k+1]})$ is equal for all $k = 1, \dots, q-1$. Thus, if the trajectories are of different lengths, the distance between adjacent radii will increase as the radius increases.

Prior to empirically calculating γ_i , with equation (2.2), the trajectory is interpolated at all radii of the quantile grid that are smaller than the maximum distance to the origin/end point. Then

$$\mathcal{D}(\gamma_i, \{\gamma_1, \dots, \gamma_n\}) = \sum_{k=1}^q w(r_{[k]}) D(\phi_i(r_{[k]}), \mathcal{S}_k),$$

where $\mathcal{S}_k = \{\phi_j(r_{[k]}) : \max_{1 \leq i \leq l_j} (v_i^{l_j}) \geq r_{[k]}\}$, and the weights $w_k = w(r_{[k]})$ are proportional to the number of trajectories that cross each concentric circle of radius $r_{[k]}$,

$$w_k = \frac{\sum_{i=1}^n I_{\{v_i^{l_i} > r_{[k]}\}}}{\sum_{k=1}^q \sum_{i=1}^n I_{\{v_i^{l_i} > r_{[k]}\}}} = \frac{\sum_{i=1}^n I_{\{v_i^{l_i} > r_{[k]}\}}}{\sum_{i=1}^n l_i}.$$

If $\max_{1 \leq i \leq l_j} (v_i^{l_j}) < r_{[k]}$, then $D(\phi_i(r_{[k]}), \mathcal{S}_k) = 0$.

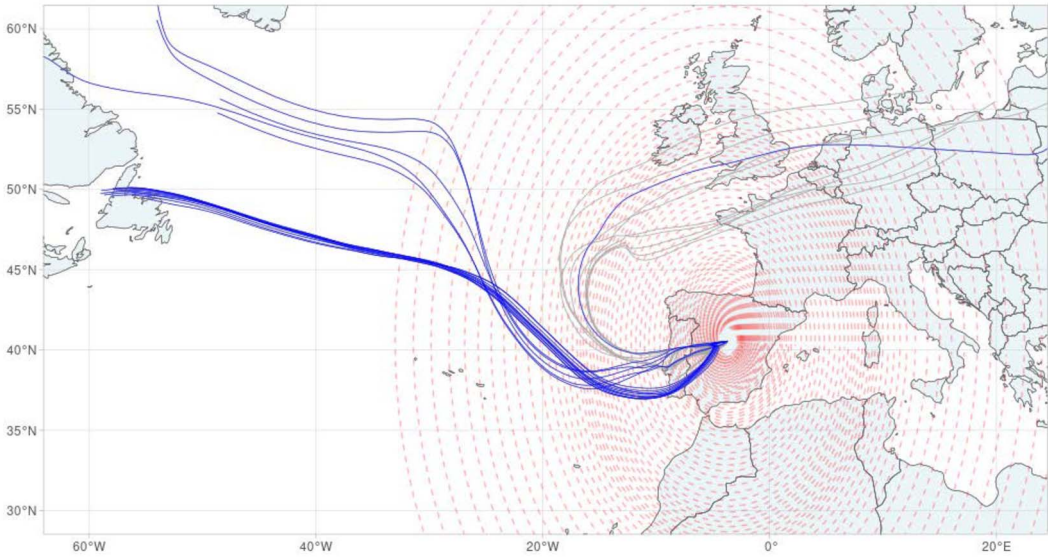


FIG. 4. *Madrid dataset with the 10% deepest back-trajectories (dark lines).*

This proposal can be modified according to a subject matter expert might be most interested in with respect to a particular bouquet of trajectories.

Following the guidelines given in Sections 3.1 and 3.2, we calculate $\mathcal{LD}^\beta$ for the dataset presented in the [Introduction](#), where the trajectories in the bouquet arrived to Madrid from three different regions. In Figure 4 the 10% deepest curves (dark), reveal this behaviour. It is important to note that these three groups are not separated in all concentric circles. For example, the western and northwestern air masses merge at about $25^\circ\text{W } 44^\circ\text{N}$.

3.3. Depth computation in simulated examples. In the following we calculate the integrated circular local depth for artificial datasets that are simulated from numerical models containing different cluster structures or outliers. We analyse the performance of $\mathcal{LD}^\beta$ under four synthetic scenarios: models A and B with two and three balanced clusters of trajectories, respectively, model C with two unbalanced clusters, and model D with a single group contaminated with 10% of outliers, which are curves of different lengths and directions. The curves are given by a series of 30 pairs of points, $\gamma = \{p_1, \dots, p_{30}\}$, generated by the parametric model,

$$\begin{aligned}\theta_{i+1} &= \theta_i + \eta_{i+1}\delta, \\ p_{i+1} &= p_i + \tau(\cos(\theta_{i+1}), \sin(\theta_{i+1})) \quad \text{for } i = 1, \dots, 29,\end{aligned}$$

where $p_1 = (0, 0)$ is the common origin/end for all trajectories and θ_1 sets the first direction of the trajectory. At each step, it is updated by adding $\delta\eta_i$, where η_i are independent random variables $\text{Bernoulli}(q)$. Higher values of δ correspond to more pronounced changes in the direction of the curve. The parameter τ modifies the amplitude at each update; higher values of τ result in longer trajectories.

Figure 5 shows the four datasets that we generated with the parameters θ_1 , δ , q , τ , and n (the number of trajectories) in Table 1 for the four different models. For each of these datasets, we compute the integrated circular local depth with the input β -grid $(0.10, 0.20, \dots, 1)$ and $\alpha = 0.1$ in Algorithm 1. The radii of the concentric circles in which we integrate correspond to equiprobable regions of probability 0.01, from 0.1 to 0.9 percentiles.

The deepest trajectories represented with dark lines in the upper panels appear in the centre of the two groups in models A and C, the three groups in B, and the nonoutlier group in D.

TABLE 2
Mean proportion of outliers detected, with the corresponding standard deviation (sd), for different contamination rates

Contaminantion rate, ζ (%)	1	5	10	15
Average outlier detection (sd)	0.98 (0.13)	0.97 (0.07)	0.95 (0.07)	0.85 (0.08)

datasets. A common strategy to overcome this problem in high-dimensional spaces is to consider integrated depths based on one-dimensional projections. In our proposal we follow this idea, but the computation of some one-dimensional depths in the circle can be very slow on a personal computer, even for mild-size datasets.

To evaluate the computational time of $\mathcal{LD}^\beta$, we generated different datasets using model A and varying the sample size between 100 and 1200. Figure 7 shows that as the sample size increases, the time required grows linearly up to 1.8 seconds for 1200 trajectories.

We repeated the same procedure replacing LD^β with the classical Tukey circular depth (TCD), which is implemented in the R package `circular` (Agostinelli and Lund (2022)). Figure 7 shows the computation time for LD^β and TCD as the one-dimensional depth. TCD shows an exponential growth in computation time, measured in seconds, as the sample size increases, requiring more than 24 hours for 1200-trajectory bouquets. The computational cost with LD^β is practically negligible compared to the computational cost with TCD.

We completed the simulation by analysing the behaviour of $\mathcal{LD}^\beta$ with denser sampling trajectories. For Models A to D, we considered the same trajectories but sampled at 1000 points. We ran 100 replicates and calculated in each one the correlation between the depths $\mathcal{LD}^\beta$ for the trajectories sampled at 30 and 1000 points. In each case the correlation was greater than 0.98. Table 3 shows that increasing the number of points in each trajectory by more than 30 times at most doubles the computational cost.

All the calculations were performed on a personal computer with a 16-core Intel Core i7-10700K 3,80 GHz processor with 16-cores processor and 64 GB of RAM. The R package `tcd`, which contains the routines we use to do the calculations, is freely available from the Github package repository. Also, the code is presented in the Supplementary Material C (Fernandez-Piana, Justel and Svarc (2025c)).

4. Where do the microorganisms that reach Antarctic Peninsula come from?. The recent deglaciation in some areas of the Antarctic Peninsula has left large areas of bare soil for the first time in thousands of years. This phenomenon allows us to try to understand the origin of microbial life in pristine regions. Considering that an important part of the movement of

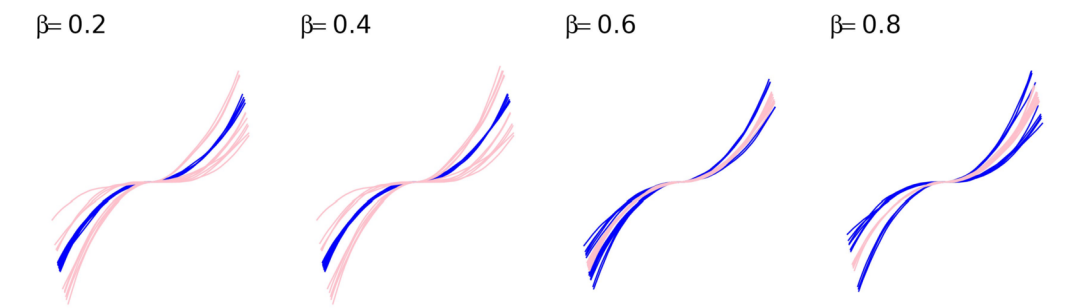


FIG. 6. Dataset generated following Model A. The integrated circular depth is computed for $\beta = 0.2, 0.4, 0.6, 0.8$. In each case the top 10% deepest curves appear with dark lines (in blue), while the 10% most outlying curves appear with light lines (in pink).

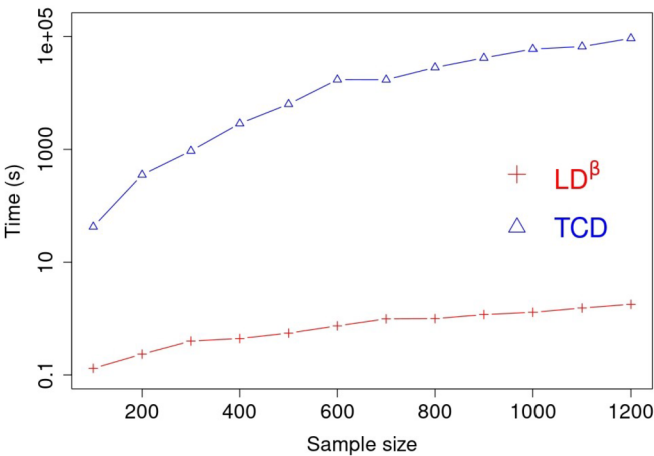


FIG. 7. Computation time (log scale) for the integrated circular local depth calculated with LD^β (+) and Tukey depth (Δ).

microorganisms occurs through the atmosphere, we analyse the circulation of air masses that have reached the Byers Peninsula (Livingston Island, South Shetland Islands, Antarctica) over the last two decades. Byers Peninsula is an Antarctic Specially Protected Area (ASPA No. 126) because of its terrestrial, aquatic, and coastal ecosystems. It is the largest ice-free area in maritime Antarctica with very low human presence. Acquiring this knowledge will help us to understand the dynamics of emerging microorganism communities in the soils of this region in future studies.

The dataset contains 104, 448 five-day back-trajectories simulated with the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPPLIT, Stein et al. (2015)), all with a common endpoint at Byers Peninsula ($62^\circ39'S$ $61^\circ04'W$). Between 2005 and 2016, four curves were generated daily (00z, 06z, 12z, 18z), each time at six vertical initialization (0.25, 0.5, 0.75, 1.25, 1.5, and 1.75 fraction of mixed boundary layer) for 121 consecutive points (hourly). Back-trajectories were cut when the height was zero metres above the surface. The global Data Assimilation System (GDAS, NOAA) at 0.5 degrees was used for the input meteorological data that HYSPPLIT uses in the back-trajectories computation. A subset of 600 randomly selected back-trajectories is shown in Figure 1B. Even for this small subset of back-trajectories, we observe that it is not possible to retrieve the most salient features of the bouquet solely by visual inspection of the map.

We calculated the LD^β for the entire dataset and for each year separately. The locality parameter β was chosen in a data-driven way by applying Algorithm 1 for each of the 71 selected radii. The 10% of the smallest radii were removed to control the noise generated at the arrival point at Byers Peninsula, and the 10% of the largest radii were also discarded to avoid favouring the longest trajectories. Figure 8 shows that the deepest and most outlying

TABLE 3
Mean time in seconds needed to compute LD^β for each model for the trajectories sampled at 30 points (OT) and the trajectories sampled at 1000 points (DT)

Model	A	B	C	D
OT	0.18	0.26	0.14	0.10
DT	0.32	0.50	0.23	0.15

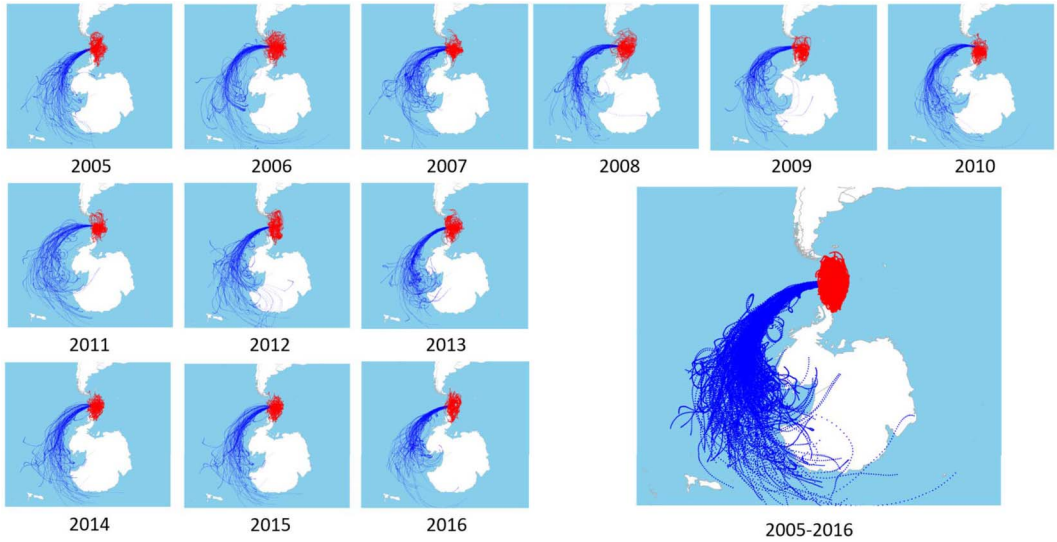


FIG. 8. Each small map shows the 1% deepest wind back-trajectories in blue (dark) and the 1% most outlying back-trajectories in red (light), yearly from 2005 to 2016. The large map on the bottom right shows the 0.3% deepest and most outlying wind back-trajectories for the entire period, maintaining the colour key.

wind back-trajectories for each year and for the complete dataset describe a clear and robust pattern. The central curves came from the Ross Sea in West Antarctica, which is consistent with the fact that the Antarctic Peninsula is subject to Southern Westerly Winds associated with the Antarctic Circumpolar Current (Gonzalez and Vasallo (2020)). These deepest back-trajectories circumvent Antarctica, making it difficult for colonising microorganisms to enter from the continent. Furthermore, most of the shallower back-trajectories are outliers, showing erratic behaviour with movements in all directions, that are associated with the frequent cyclones that cross the area. These back-trajectories are not of interest for the problem of microbial transport because they are short and remain close to the peninsula. They are, therefore, not capable of introducing alien microorganisms to colonise the area under study.

The same analysis using a global depth instead of a local, that is, setting $\beta = 0.8$, shows that the main patterns are similar, as expected, since no distinct groups emerge (Figure 1, Supplementary Material A, Fernandez-Piana, Justel and Svarc (2025a)). However, in general, they show less dispersion, especially at the beginning of the back-trajectories. In contrast, the choice of the data-driven locality parameter allows us to capture more variability in the behaviour of the back-trajectories, showing more cases in which the curves came within the Antarctic continent, allowing for internal colonisation. When considering the fixed locality parameter $\beta = 0.2$, the results are very similar to those obtained with the data-driven choice; see Figure 2 in Supplementary Material B (Fernandez-Piana, Justel and Svarc (2025b)).

As the seasonal behaviour of the Southern Westerly Winds can influence the back-trajectories, we reproduced the analysis for the summer (December-January-February) and winter (June-July-August) back-trajectories separately. This fact is related to the semiannual oscillation, which is stronger at the equinoxes and weaker at the solstices. Figure 9 shows the deepest back-trajectories for each year and season. All the central back-trajectories of the air masses came from the west, except in the winter of 2010, when they mostly moved in the opposite direction. A unique pattern of aerial transport of microorganisms is identified, except in the summers of 2006, 2008, 2010, and the winter of 2016, where the deeper back-trajectories separate into two groups highlighting different origins. In years 2007, 2011, and 2015, the central air masses came very close to the American terrestrial ecosystems, in some cases passing over land, which may imply that in these years there could have been a greater

flow of alien species toward Antarctica. These years are part of the two periods (2007–08, 2010–11) with strong La Niña events (1.5 to 1.9 SST anomaly in the Oceanic Niño Index), and the only period (2015–16) with a very strong El Niño event (≥ 2.0) between 2005 and 2016. The Oceanic Niño Index (ONI) is the standard index used by the NOAA Climate Prediction Center (U.S. National Oceanic and Atmospheric Administration) to classify El Niño and La Niña events. This relationship alerts us to a possible increase in the vulnerability of Antarctic microbial communities to future changes in ENSO (El Niño–Southern Oscillation) due to climate change.

A particular phenomenon can be observed in the winter of 2010, where the southeasterlies became the dominant winds. This finding is consistent with the results obtained by [Turner et al. \(2016\)](#), which indicate a period of absence of warming in Antarctica between 1998 and 2014 (the end of the study). During this period there were stronger easterly and southeasterly near-surface currents toward the Antarctic Peninsula, associated with katabatic winds. These winds are particularly strong in winter, when they tend to push sea ice toward the east coast of the peninsula. Over the past 20 years, the sea ice-cover has increased, producing a cooling effect due to changes in atmospheric circulation. Visual inspection of all the winter wind back-trajectories (see Figure 3 in the Supplementary Material B, [Fernandez-Piana, Justel and Svarc \(2025b\)](#)) reinforces the findings of the method, which are in line with Turner's results. It can be seen that easterly and southeasterly winds were frequent in winter and that this phenomenon is more pronounced in 2010. The result persists, even when we repeated the analysis with a global depth (considering $\beta = 0.8$). Moreover, we have also analysed what happens when we examine the deepest 25% of the curves. In some years, such as 2008 or 2014, different origins emerge and new patterns are identified for the deepest curves. In other cases, such as 2013 and 2016, where different origins were already highlighted for 5% of the deepest curves, these patterns are reinforced (see Figure 4 in the Supplementary Material B, [Fernandez-Piana, Justel and Svarc \(2025b\)](#)).

In conclusion, the analysis of the back-trajectories reaching the Antarctic Peninsula throughout the year during this period tells us that the main origin of the microorganisms is not particularly sensitive to the occurrence of a short period of cooling followed by a very strong El Niño. However, when we consider only the winter back-trajectories, the air masses from the interior of the continent are more relevant during the years of absence of warming, becoming the most central in 2010. The study we have carried out also provides some evidence of the possible influence of El Niño and La Niña phenomena on the arrival of new invasive species capable of thriving in the current climate change context.

5. Concluding remarks. In this paper we study the patterns of air mass back-trajectories in Antarctica concerning microbial dispersal. The deeper curves allow us to identify the main air transport routes of airborne microorganisms to Byers Peninsula, a specially protected area on the Antarctic Peninsula dedicated to scientific research. This study will also help future works about the origin and evolution of the sampled soils and microbial mats in this very sensitive and unique location in the Maritime Antarctica.

For this aim we introduce new notions of depth measures for trajectories that share the same origin or endpoint and that spread in all possible directions. The main idea is to define an integrated depth where the integral variable is the radius of concentric circles from the common point. Thus, a new local depth measure for circular data is introduced. Desirable theoretical properties are obtained for both the circular local depth and the integrated depth. An efficient data-driven algorithm is presented, which is available online and allows the handling of large datasets.

The problem studied raises extensions and future challenges beyond the scope of this paper. The integrated depth definition can be extended to problems where the trajectories belong

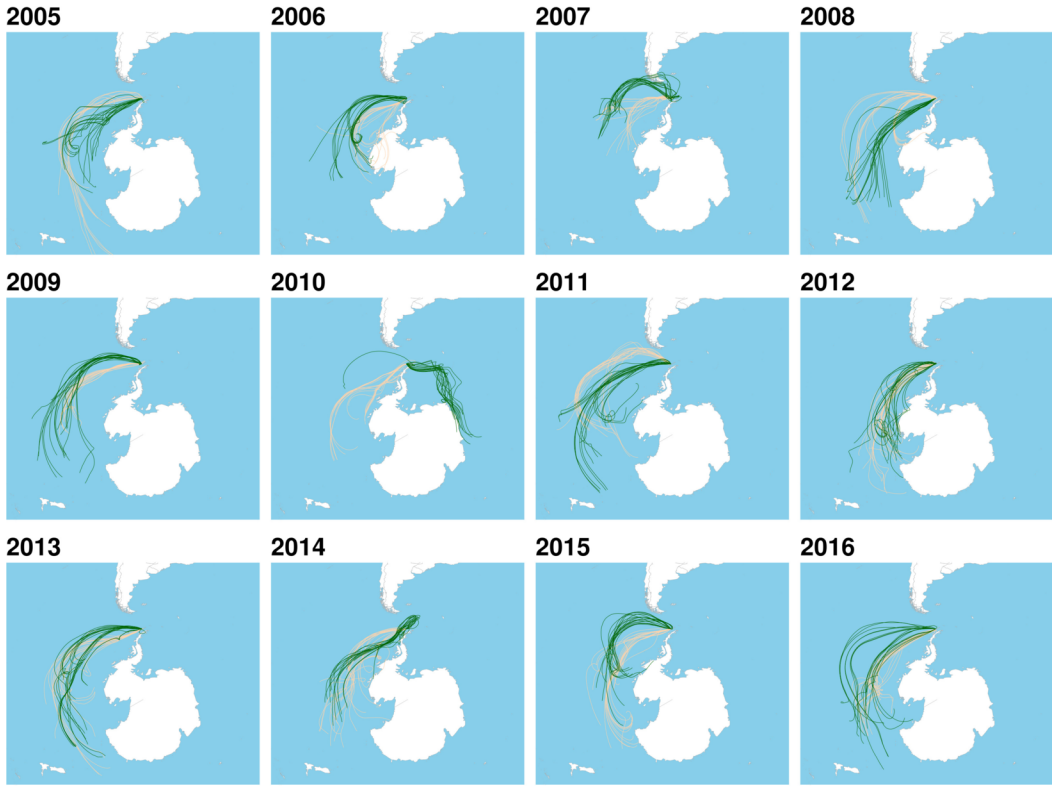


FIG. 9. In light lines (yellow) are the 5% of deepest curves for summer, and in dark lines (green) are the 5% of deepest curves for winter, for each year between 2005 and 2016.

to a two-dimensional smooth connected Riemannian manifold. Then the slicing for the integral depth is in an expanding sequence of equidistant one-dimensional subvarieties, where the shape is given by the metric associated to the manifold. If the one-dimensional subvariety is not a circle, the polar coordinate parameterisation makes no sense, but the quantiles can still be defined, so the λ definition should be reformulated accordingly. Computational aspects should be adapted ad hoc. This methodology can be extended beyond the problems of air and particle transport. An example could be the study of the density of cart networks in irregular land. This method can also be applied to the problem of in-store product placement of complementary items, considering as centres the most attractive promotions.

Furthermore, it would be appealing to explore some other ideas for circular local depth by considering a kernel density depth. An important point in this case will be the choice of the bandwidth, which is particularly challenging in the presence of multimodality.

Finally, there are many situations where the start or end point of the bouquet varies in time, for example, when measuring emissions from a moving vehicle. In these cases it would be interesting to be able to describe temporal patterns of the main features of these bouquets of trajectories with time dependencies.

Acknowledgments. The authors acknowledge the computer resources, technical expertise, and assistance provided by Centro de Computación Científica at the Universidad Autónoma de Madrid (CCC-UAM). Our thanks to Sergi González from WSL Institute for Snow and Avalanche Research SLF and Pablo Sanz from CCC-UAM. We also gratefully acknowledge the NOAA Air Resources Laboratory (ARL) for the provision of the HYSPLIT transport and dispersion model and/or READY website (<https://www.ready.noaa.gov>) used in this publication.

Funding. This research has been funded by grant PID2020-116520RB-I00 by the Spanish Ministry of Science and Innovation/State Agency of Research MCIN/AEI.

SUPPLEMENTARY MATERIAL

Supplementary material A: Integrated depth for trajectories of airborne microorganisms to Antarctica (DOI: [10.1214/25-AOAS2041SUPPA](https://doi.org/10.1214/25-AOAS2041SUPPA); .pdf). This supplementary material contains the proofs for the results given in Section 2.

Supplementary material B: Integrated depth for trajectories of airborne microorganisms to Antarctica (DOI: [10.1214/25-AOAS2041SUPPB](https://doi.org/10.1214/25-AOAS2041SUPPB); .pdf). This supplementary material provides further details on the case presented in Section 4.

Supplementary material C: Integrated depth for trajectories of airborne microorganisms to Antarctica (DOI: [10.1214/25-AOAS2041SUPPC](https://doi.org/10.1214/25-AOAS2041SUPPC); .zip). The R library is provided in this supplementary material.

REFERENCES

- AGOSTINELLI, C. and LUND, U. (2022). *R Package ‘Circular’: Circular Statistics*.
- AGOSTINELLI, C. and ROMANAZZI, M. (2011). Local depth. *J. Statist. Plann. Inference* **141** 817–830. [MR2732952 https://doi.org/10.1016/j.jspi.2010.08.001](https://doi.org/10.1016/j.jspi.2010.08.001)
- AGOSTINELLI, C. and ROMANAZZI, M. (2013). Nonparametric analysis of directional data based on data depth. *Environ. Ecol. Stat.* **20** 253–270. [MR3068658 https://doi.org/10.1007/s10651-012-0218-z](https://doi.org/10.1007/s10651-012-0218-z)
- CUESTA-ALBERTOS, J. A. and NIETO-REYES, A. (2008). The random Tukey depth. *Comput. Statist. Data Anal.* **52** 4979–4988. [MR2526207 https://doi.org/10.1016/j.csda.2008.04.021](https://doi.org/10.1016/j.csda.2008.04.021)
- CUEVAS, A., FEBRERO, M. and FRAIMAN, R. (2007). Robust estimation and classification for functional data via projection-based depth notions. *Comput. Statist.* **22** 481–496. [MR2336349 https://doi.org/10.1007/s00180-007-0053-0](https://doi.org/10.1007/s00180-007-0053-0)
- CUEVAS, A. and FRAIMAN, R. (2009). On depth measures and dual statistics. A methodology for dealing with general data. *J. Multivariate Anal.* **100** 753–766. [MR2478196 https://doi.org/10.1016/j.jmva.2008.08.002](https://doi.org/10.1016/j.jmva.2008.08.002)
- DAI, X. and MÜLLER, H.-G. (2018). Principal component analysis for functional data on Riemannian manifolds and spheres. *Ann. Statist.* **46** 3334–3361. [MR3852654 https://doi.org/10.1214/17-AOS1660](https://doi.org/10.1214/17-AOS1660)
- FERNANDEZ-PIANA, L., JUSTEL, A. and SVARC, M. (2025a). Supplement A to “Integrated depth for trajectories of airborne microorganisms to Antarctica.” <https://doi.org/10.1214/25-AOAS2041SUPPA>
- FERNANDEZ-PIANA, L., JUSTEL, A. and SVARC, M. (2025b). Supplement B to “Integrated depth for trajectories of airborne microorganisms to Antarctica.” <https://doi.org/10.1214/25-AOAS2041SUPPB>
- FERNANDEZ-PIANA, L., JUSTEL, A. and SVARC, M. (2025c). Supplement C to “Integrated depth for trajectories of airborne microorganisms to Antarctica.” <https://doi.org/10.1214/25-AOAS2041SUPPC>
- FERNANDEZ-PIANA, L. and SVARC, M. (2022). An integrated local depth measure. *ASIA Adv. Stat. Anal.* **106** 175–197. [MR4426853 https://doi.org/10.1007/s10182-021-00424-6](https://doi.org/10.1007/s10182-021-00424-6)
- FRAIMAN, R. and MUNIZ, G. (2001). Trimmed means for functional data. *TEST* **10** 419–440. [MR1881149 https://doi.org/10.1007/BF02595706](https://doi.org/10.1007/BF02595706)
- GALBAN, S., JUSTEL, A., GONZALEZ, S. and QUESADA, A. (2021). Local meteorological conditions, shape and desiccation influence dispersal capabilities for airborne microorganisms. *Sci. Total Environ.* **10** 146653.
- GONZALEZ, S. and VASALLO, F. (2020). Antarctic climates. In *Encyclopedia of the World’s Biomes* (M. I. Goldstein and D. A. DellaSala, eds.) **2** 595–605.
- HARRIS, T., TUCKER, J. D., LI, B. and SHAND, L. (2021). Elastic depths for detecting shape anomalies in functional data. *Technometrics* **63** 466–476. [MR4331447 https://doi.org/10.1080/00401706.2020.1811156](https://doi.org/10.1080/00401706.2020.1811156)
- LEY, C., SABBAB, C. and VERDEBOUT, T. (2014). A new concept of quantiles for directional data and the angular Mahalanobis depth. *Electron. J. Stat.* **8** 795–816. [MR3217789 https://doi.org/10.1214/14-EJS904](https://doi.org/10.1214/14-EJS904)
- LIU, R. Y. and SINGH, K. (1992). Ordering directional data: Concepts of data depth on circles and spheres. *Ann. Statist.* **20** 1468–1484. [MR1186260 https://doi.org/10.1214/aos/1176348779](https://doi.org/10.1214/aos/1176348779)
- LÓPEZ-PINTADO, S. and ROMO, J. (2011). A half-region depth for functional data. *Comput. Statist. Data Anal.* **55** 1679–1695. [MR2748671 https://doi.org/10.1016/j.csda.2010.10.024](https://doi.org/10.1016/j.csda.2010.10.024)
- MIRZARGAR, M., WHITAKER, R. and KIRBY, R. (2014). Curve boxplot: Generalization of boxplot for ensemble curves. *IEEE Trans. Vis. Comput. Graph.* **20** 2654–2663.
- NAGY, S., GIJBELS, I., OMELKA, M. and HLUBINKA, D. (2016). Integrated depth for functional data: Statistical properties and consistency. *ESAIM Probab. Stat.* **20** 95–130. [MR3528619 https://doi.org/10.1051/ps/2016005](https://doi.org/10.1051/ps/2016005)

- PAINDAVEINE, D. and VAN BEVER, G. (2013). From depth to local depth: A focus on centrality. *J. Amer. Statist. Assoc.* **108** 1105–1119. [MR3174687](#) <https://doi.org/10.1080/01621459.2013.813390>
- PANDOLFO, G., PAINDAVEINE, D. and PORZIO, G. C. (2018). Distance-based depths for directional data. *Canad. J. Statist.* **46** 593–609. [MR3902616](#) <https://doi.org/10.1002/cjs.11479>
- SGUERA, C., GALEANO, P. and LILLO, R. (2014). Spatial depth-based classification for functional data. *TEST* **23** 725–750. [MR3274472](#) <https://doi.org/10.1007/s11749-014-0379-1>
- SGUERA, C., GALEANO, P. and LILLO, R. (2016). Functional outlier detection by a local depth with application to NO_x levels. *Environ. Res. Risk Assess.* **30** 1115–1130.
- SRINIVAS, C. V., VENKATESAN, R., BASKARAN, R., RAJAGOPAL, V. and VENKATRAMAN, B. (2014). Regional scale atmospheric dispersion simulation of accidental releases of radionuclides from Fukushima Dai-ichi reactor. *Atmos. Environ.* **61** 66–84.
- STEIN, A., DRAXLER, R., ROLPH, G., STUNDER, B., COHEN, M. and NGAN, F. (2015). NOAA HYSPLIT atmospheric transport and dispersion modeling system. *Bull. Am. Meteorol. Soc.* **96** 2059–2077.
- SU, J., KURTEK, S., KLASSEN, E. and SRIVASTAVA, A. (2014). Statistical analysis of trajectories on Riemannian manifolds: Bird migration, hurricane tracking and video surveillance. *Ann. Appl. Stat.* **8** 530–552. [MR3192001](#) <https://doi.org/10.1214/13-AOAS701>
- TURNER, J., LU, H., WHITE, I. et al. (2016). Absence of 21st century warming on Antarctic Peninsula consistent with natural variability. *Nature* **535** 411–415.
- VERSTRAETEN, G., POESEN, J., DEMAREE, G. and SALLES, C. (2006). Long-term (105 years) variability in rain erosivity as derived from 10-min rainfall depth data for Ukkel (Brussels, Belgium): Implications for assessing soil erosion rates. *J. Geophys. Res.* **111**. D22109.
- ZUO, Y. and SERFLING, R. (2000). General notions of statistical depth function. *Ann. Statist.* **28** 461–482. [MR1790005](#) <https://doi.org/10.1214/aos/1016218226>