Technometrics

ISSN: 0040-1706 (Print) 1537-2723 (Online) Journal homepage: https://www.tandfonline.com/loi/utch20

Elastic Depths for Detecting Shape Anomalies in Functional Data

Trevor Harris, J. Derek Tucker, Bo Li & Lyndsay Shand

To cite this article: Trevor Harris, J. Derek Tucker, Bo Li & Lyndsay Shand (2020): Elastic Depths for Detecting Shape Anomalies in Functional Data, Technometrics, DOI: 10.1080/00401706.2020.1811156

To link to this article: https://doi.org/10.1080/00401706.2020.1811156

TECHNOMETRICS
2020, VOL. 00, NO. 0, 1–11
https://doi.org/10.1080/00401706.2020.1811156

Taylor & Francis
Taylor & Francis Group

Elastic Depths for Detecting Shape Anomalies in Functional Data

Trevor Harris[a], J. Derek Tucker[b], Bo Li[a], and Lyndsay Shand[b]

[a]Department of Statistics, University of Illinois at Urbana-Champaign, Champaign, IL; [b]Sandia National Laboratories, Albuquerque, NM

ABSTRACT

We propose a new family of depth measures called the elastic depths that can be used to greatly improve shape anomaly detection in functional data. Shape anomalies are functions that have considerably different geometric forms or features from the rest of the data. Identifying them is generally more difficult than identifying magnitude anomalies because shape anomalies are often not distinguishable from the bulk of the data with visualization methods. The proposed elastic depths use the recently developed elastic distances to directly measure the centrality of functions in the amplitude and phase spaces. Measuring shape outlyingness in these spaces provides a rigorous quantification of shape, which gives the elastic depths a strong theoretical and practical advantage over other methods in detecting shape anomalies. A simple boxplot and thresholding method is introduced to identify shape anomalies using the elastic depths. We assess the elastic depth's detection skill on simulated shape outlier scenarios and compare them against popular shape anomaly detectors. Finally, we use hurricane trajectories to demonstrate the elastic depth methodology on manifold valued functional data.

ARTICLE HISTORY
Received July 2019
Accepted August 2020

KEYWORDS
Anomaly detection; Data depth; Functional data; Shape analysis

1. Introduction

As data collection methods rapidly advance, functional data and functional data analysis (FDA) have become more prevalent. Functional data refer to data collected continuously across a compact domain, such as a fixed length of time or region of space, and where an observation is an entire curve or surface over the domain, rather than a single value. Examples of functional data include growth rate curves, electrocardiogram (ECG) data, temperature profiles, imaging data containing geometric shapes, and hurricane trajectories (see Figure 1).

As with traditional data analysis methods, it is critical to perform exploratory data analysis with functional data. Exploratory analysis can reveal significant trends or anomalies, which could bias post-processing analysis such as model fitting. Functional anomalies are of particular interest because of the adverse effects they can have on statistical models. Functional anomalies can also be interesting in their own right and can even be the primary focus of study.

The hurricane trajectories in Figure 1 present a challenge to current functional shape anomaly detectors. Due to the vast distances hurricanes travel, their trajectories resemble paths along the surface of a sphere. Existing methods do not handle spherical valued data, or more generally manifold valued data, so they have to approximate these paths with two-dimensional trajectories. This approximation distorts the distances between curves and consequently has a strong influence on the detection of anomalies. Furthermore, exploratory analysis shows that these trajectories exhibit substantial phase variability that could affect shape outlier detection (Srivastava et al. 2011).

In the functional data setting, identifying functional anomalies, that is, identifying an entire function as an outlier, is not as straightforward as identifying univariate outliers with visualization methods. By definition, a functional anomaly is a function that is significantly more "extreme" in its characteristics than the rest of the functional data. Generally, functional anomalies are categorized into two types: magnitude and shape anomalies (Dai and Genton 2018). Magnitude anomalies are functions that clearly lie outside of the range of all other functions and are usually detected through data visualization methods (e.g., Hyndman and Shang 2010; Sun and Genton 2011; Myllymäki et al. 2017). On the other hand, shape anomalies take on a different shape or pattern than the rest of the data. They are more challenging to identify with visualization methods because they can lie hidden among the rest of the functions (Arribas-Gil and Romo 2014). Examples of shape anomalies include trajectories with more or less curvature, trajectories with more or fewer oscillations, or trajectories sampled from a process having a different mean function than the rest of the data. These are only a few examples; the possible ways a trajectory can be shape outlying are innumerable. Furthermore, the information contained in the shapes of curves "matters a great deal" (Horváth and Kokoszka 2012) and can be quite different from the information contained in their magnitudes. Therefore, it is important to develop methods that can isolate shape information and identify shape outliers from the rest of the data.

Many methods for identifying shape outliers rely on the notion of functional data depth. Functional data depth is a family of methods used to define centrality and induce a

CONTACT Trevor Harris trevorh2@illinois.edu Department of Statistics, University of Illinois at Urbana-Champaign, Champaign, IL 61820.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/TECH.

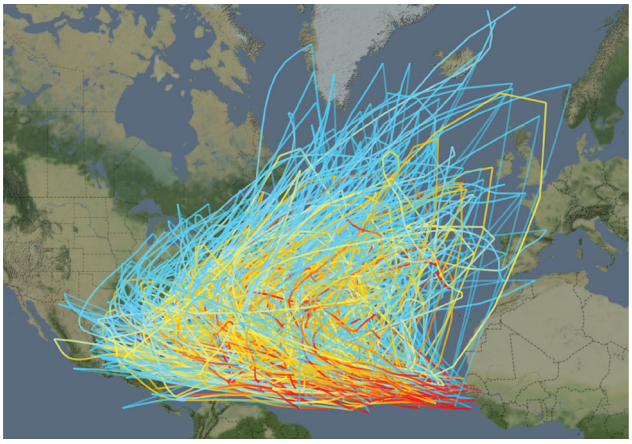

Figure 1. Hurricane trajectories from the HURDAT2 database. Trajectories are colored from red to blue, with red indicating the origin of the hurricane. The familiar "U-shaped" pattern emerges as hurricanes generally start near the coast of Africa, migrate to North America, and then curve back toward Europe. This type of data poses many challenges to existing shape outlier methodology. Trajectories are multivariate (latitude and longitude), exist on the surface of a nonlinear manifold ($\mathbb{S}^2$), and exhibit significant phase and magnitude variability independent of their shape.

center-outward ordering on the sample functions (Liu, Parelius, and Singh 1999; Zuo and Serfling 2000; Mosler and Polyakova 2012). All depth measures rank functions from most central (higher depth values) to least central (lower depth values) and are typically scaled to take on values in [0, 1].

Many outlier detection methods based on depth decompose total depth (or outlyingness) into magnitude and shape depth (or outlyingness). Arribas-Gil and Romo (2014) proposed the outliergram (OG), a visualization tool for the shape and magnitude components of trajectories, using the half-space depth (Tukey 1977) and the band depth (Lopez-Pintado and Romo 2009). Later, Huang and Sun (2019) introduced the total variation depth, which they decomposed into magnitude and shape components. In Dai and Genton (2019), the authors pointed out that these methods, which rely on integrated depth, do not efficiently represent the centrality of functions. To remedy this, Rousseeuw, Raymaekers, and Hubert (2018) and Dai and Genton (2019) simultaneously proposed the concept of directional outlyingess. Directional outlyingness has since been used as the basis of the functional outlier map (Rousseeuw, Raymaekers, and Hubert 2018) and the magnitude-shape plot tool (Dai and Genton 2018). Both of these methods again decompose their depth measures into magnitude and shape components, for the separate identification of magnitude and shape anomalies.

Other outlier detection methods proposed in the recent literature account for the geometry of the functions. Kuhnt and Rehage (2016) developed the functional tangential angle (FUNTA) pseudo-depth based on the tangential angles of the intersections of the centered data. Nagy, Gijbels, and Hlubinka (2017) proposed two modifications of previous depth notions to better identify shape outliers by emulating derivatives with multidimensional projections. Xie et al. (2017) separated the variability of functional data into amplitude and phase components, using the registration methods of Srivastava et al. (2011), and displayed this variability using independent boxplots for each component. Xie et al. (2017) showed how treating the phase and amplitude components of trajectories separately could greatly improve the detection rate of shape outliers. Their method, however, falls short of fully characterizing the shape distribution and instead relies on an optimization procedure to approximate the boundary of the inlier distribution.

Finally, there are methods based on functional principal component analysis (FPCA), which extract the features of normal shapes and detect anomalies by finding functions with abnormal features. These methods include a step-wise functional outliers detection test (Yu, Zou, and Wang 2012) and an FPCA score based distance test (Ren, Chen, and Zou 2017). FPCA based feature extraction can be quite powerful for detecting many types of shape anomalies, but FPCA is also known to be deficient when temporal variability is present (Srivastava et al. 2011; Tucker, Wu, and Srivastava 2013). Furthermore, these methods are not designed to explicitly identify shape anomalies separately from magnitude anomalies. Because we assume that both temporal variability and magnitude variability are present and are nuisance properties that mask shape, the problems and data considered here are quite different from the ones considered by FPCA methods. For those reasons, these methods were not included in the comparisons in Section 5.2.

In this article, we introduce a new family of depth measures, called the elastic depths, based on the elastic shape distances used in Xie et al. (2017). We first use the elastic distances to directly measure the distance between the shapes of individual trajectories. This allows us to define a notion of data depth that appropriately captures the distribution of the trajectories' shapes and allows for the explicit identification of shape outliers. Current literature falls short of fully representing the shape distribution and instead uses either a surrogate for shape or only approximates the shape distribution. Explicitly using the shape distribution, as we do in this article, also confers several practical benefits, namely

1. Finding the inner quartile regions and outlier bound is trivial due to our depth based representation. The outlier bound does not require any optimization procedures and has only a single input parameter. We demonstrate in Section A.5 of the online supplement that the performance of our method is relatively insensitive to the value of the input parameter.
2. In Sections A.3 and 5.2, we empirically demonstrate the highly competitive and often superior outlier detection skill of the elastic depth based boxplots. We show that the elastic depths are consistently the top performing detection method across many different shape classes and that they can identify shape outliers in the presence of substantial translation and phase noise.
3. Shape distributions can easily be generalized to manifold valued trajectories, such as the hurricane trajectories (Figure 1), because the elastic depths are based purely on distance. Different manifolds merely require different distance metrics. In Section 6, we show the elastic depths applied to the Hurricane trajectories and in Section A.7 of the online supplement we show additional examples of the elastic depths applied to $\mathbb{R}$ and $\mathbb{R}^2$ valued data.

2. Background

2.1. Data Depth

Data depth is a general notion of measuring the centrality of observations with respect to a distribution. In the FDA literature, data depth is the dominant method used to define centrality and induce ordering on a function space. Given a distribution P on a function space F, a depth function maps each trajectory $f \in F$ to a value in $[0, 1]$, such that the closer a trajectory is to the center of P, the higher its depth value is. If this mapping is monotonic, that is, higher depth values necessarily mean higher centrality, then the depth function induces a center-outward ordering on the function space F with respect to P. This makes depth a natural framework for evaluating the outlyingess of observations. High depth values mean an observation is very close to the center of P so, conversely, low depth values mean an observation is very far from the center of P. Therefore, trajectories with extremely low depth values, as compared to the rest of the distribution, are likely to be outlying or anomalous.

2.2. Elastic Shape Analysis

Elastic shape analysis (ESA) is a collection of techniques for registering functional data through a phase-amplitude separation procedure and for performing statistical analysis on the separated phase and amplitude components (Kurtek, Srivastava, and Wu 2011; Srivastava et al. 2011; Tucker, Wu, and Srivastava 2013). Phase and amplitude represent two orthogonal components of a function's variability. The amplitude component represents variability in shape, where shape refers to the properties of a function that remain unchanged under the shape preserving transformations: rotation, translation, scaling, and phase (Srivastava and Klassen 2016). The phase component represents the "domain" or "timing" variability of the trajectories. Because amplitude is invariant to these phase transformations, amplitude is distinct from the usual concept of magnitude. Magnitude measures the size of the observed realization of a trajectory while amplitude measures the size of the trajectories shape.

The distinguishing feature of ESA is the use of the square root slope function (SRSF) for registration (Kurtek, Srivastava, and Wu 2011). For real valued trajectories, the SRSF bijectively maps, up to an additive constant, a real valued function f to its normalized gradient $f'/\sqrt{|f'|}$. Under ESA, two real valued trajectories are registered by elastically deforming the domain of one function such that the L^2 distance between the SRSFs of the two functions is minimized (Section 2.3). The amount of elastic deformation needed to register two functions is measured by the phase distance (Section A.1), while the residual L^2 distance between the SRSFs, post registration, defines the amplitude distance between them (Section 2.3). Together the amplitude and phase distance are known as the elastic distances. The key insight of ESA is that by registering SRSFs, instead of trajectories directly, the amplitude distances are proper metrics and they are invariant to the shape preserving transformations. Thus, amplitude distance can be used to define the distance between the shapes of functions.

Later the square root slope velocity function (SRVF) (Srivastava et al. 2011) was introduced to register $\mathbb{R}^n$ valued trajectories and the transported square root slope velocity function (TSRVF) (Su et al. 2014) was introduced to register Riemannian manifold valued trajectories, such as trajectories observed on the unit sphere $\mathbb{S}^2$. These notions allow us to calculate amplitude distances between multivariate functions and manifold valued functions, respectively. We present the details for computing amplitude distances for $\mathbb{R}$ valued trajectories in Section 2.3. The details for $\mathbb{R}^n$ valued and $\mathbb{S}^2$ valued trajectories are deferred to the appendix (Section A.1).

The advantages of the ESA approach to shape analysis have previously been shown in the works of Srivastava et al. (2011), Kurtek, Srivastava, and Wu (2011), Tucker, Wu, and Srivastava (2013), and Su et al. (2014). ESA rigorously defines the shape space for a given class of trajectories and then defines a way to construct a proper distance metric on that shape space. The ESA based metrics are preserved under the shape preserving transformations: translation, scale, rotation, and reparameterization (phase). This improves theoretically over alternative shape metrics, such as Huang and Sun (2016), Dai and Genton (2018), and Dai and Genton (2019), that do not guarantee invariance or equivariance to shape transformations. The ESA framework is also general enough to apply to data observed in $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{S}^2$, and any Riemannian manifold $\mathcal{M}$ that has an intrinsic metric. This is important for our motivating example, Atlantic hurricane trajectories (Figure 1), which are observed on the surface of a sphere.

2.3. Amplitude Distance for $\mathbb{R}$ Valued Functions

Let $F_R = \{f : [0,1] \mapsto \mathbb{R}, f \text{ differentiable}\}$ be the class of differentiable trajectories on $[0, 1]$ mapping to $\mathbb{R}$. The SRSF was introduced in Srivastava et al. (2011) as the following transformation on trajectories $f \in F_R$:

Definition 2.1. Let f be a differentiable trajectory in F_R, the SRSF of f is

$$q_f(t) = \frac{f'(t)}{\sqrt{|f'(t)|}}.$$

As was shown in their article, the SRSF is a bijective mapping, up to an additive constant, from the space F_R to the space of square integrable functions L^2. This means that for two functions $f, g \in F_R$, the norm on L^2

$$||q_f - q_g||_2 = \sqrt{\int_0^1 |q_f(t) - q_g(t)|^2 dt}, \tag{1}$$

where q_f, q_g are f and g's associated SRSFs, is a proper distance between f and g. This norm is particularly important for shape analysis because it is phase invariant (Srivastava and Klassen 2016). That is, for any phase function $\gamma \in \Gamma$

$$||q_{f\circ\gamma} - q_{g\circ\gamma}||_2 = ||q_f - q_g||_2,$$

where $f \circ \gamma(t) = f(\gamma(t))$, $\forall t \in [0, 1]$. Technical descriptions of phase functions γ and phase space Γ are can be found in Section A.1 or (Srivastava et al. 2011), but γ functions essentially act to deform the domain $[0, 1]$.

Phase invariance means that Equation 1 is measuring some quantity that is independent of the representation, or phase, with

which two functions are observed. This is only true if f and g share a common phase representation, so to find the amplitude distance between two arbitrary $f, g \in F_R$ we need to first place them in phase with each other. That is, we need to find some $\gamma^* \in \Gamma$ such that

$$\gamma^* = \arg\inf_{\gamma\in\Gamma} ||q_f - q_{g\circ\gamma}||_2,$$

so that $||q_f - q_{g\circ\gamma^*}||_2$ measures the difference in their amplitudes. This can be more directly stated by defining the amplitude distance between f and g as in Srivastava et al. (2011).

Definition 2.2 (Amplitude distance). Let f and g be two trajectories in F_R, then the amplitude distance between f and g is

$$d_a(f,g) = \inf_{\gamma\in\Gamma} ||q_f - q_{g\circ\gamma}||_2,$$

where q_f and $q_{g\circ\gamma}$ denote the SRSF's of f and $g\circ\gamma$, respectively.

3. Elastic Depth

3.1. Definition of Elastic Depth

The exact analytic form of the elastic depths will greatly depend on the manifold on which the functional objects live. This is because the elastic depths are a robust summary of the distances between functional objects and the definition of distance between functional objects will inherently depend on the manifold on which they are observed. For instance, the distance between trajectories in $\mathbb{R}^n$ is very different from the distance between trajectories on $\mathbb{S}^2$. For the elastic distances to exist, however, it is only required that the data live on a Riemannian manifold (Section 3.1), such as $\mathbb{R}^n$ or $\mathbb{S}^2$, because the TSRVF of Su et al. (2014) can always be used to construct appropriate phase and amplitude distances. Therefore, we will only assume that our data live on a Riemannian manifold M with an intrinsic metric, so the space of functions we consider is defined as

$$F_M = \{f : [0,1] \mapsto M, \\ f \text{ is differentiable and M is a Riemannian manifold}\}.$$

The amplitude distance between two functions $f_1, f_2 \in F_M$ will generically be denoted as $d_a(f_1, f_2)$ and the phase distance between them as $d_p(f_1, f_2)$. The exact form of these distances is left unspecified because the amplitude distance is highly dependent on the manifold M. See Section 2.3 for the definition of amplitude distance for $\mathbb{R}$ valued trajectories and A.1 for the definitions of phase distance and amplitude distance for $\mathbb{R}^n$ valued and $\mathbb{S}^2$ valued trajectories. We now define the elastic depths for data observed on a manifold M using the associated amplitude and phase distances. Let P denote a distribution supported on the space F_M and suppose we observe a function $f \in F_M$. We first introduce the idea of outlyingness, which describes the degree to which f is an outlier relative to P. We further divide this concept into amplitude and phase outlyingness, using the amplitude and phase distances, respectively. This is done to separately quantify the shape outlyingness and phase outlyingness of f relative to P. Amplitude and phase outlyingness are, respectively, denoted as O_a and O_p and are defined as

$$O_a(f,P) = \inf_{t\in\mathbb{R}^+} \left\{ P(d_a(f,X) \le t) \ge \frac{1}{2} \right\}, \text{ and}$$
$$O_p(f,P) = \inf_{t\in\mathbb{R}^+} \left\{ P(d_p(f,X) \le t) \ge \frac{1}{2} \right\},$$

where X is a random function in F_M and drawn from the distribution P. The outlyingness functions O_a and O_p robustly summarize the pairwise distances between f and all other functions $X \in F_M$. These two functions define a measure of outlyingness such that if $O_a(f,P)$ is large then f is generally dissimilar in amplitude from other functions $X \in F_M$, with respect to the distribution P. Likewise, if $O_a(f,P)$ is small then f is similar in amplitude to other functions $X \in F_M$.

To convert O_a and O_p into depth functions we invert them with the type B depth construction of Zuo and Serfling (2000):

$$D_a(f,P) = (1 + O_a(f,P))^{-1}, \quad (2)$$
$$D_p(f,P) = (1 + O_p(f,P))^{-1}. \quad (3)$$

$D_a(f,P)$ and $D_p(f,P)$ are, respectively, called the amplitude depth and phase depth of f with respect to P. Together we denote them the *elastic depths*. The purpose of inverting the outlyingness functions in this manner is to create bounded measures of centrality, that is, depths, on the amplitude and phase spaces associated with F_M. When depths, such as D_a and D_p, satisfy the properties outlined in Section 3.2, they provide a nonparametric and moment free characterization of the distribution P. Larger depth values indicate higher centrality and low outlyingness, while lower values indicate higher outlyingness. Thus, D_a and D_p provide a simple and rigorous way to identify outliers based on the underlying distribution P.

3.2. Properties

Within the depth literature there have been many desirable properties discussed for both multivariate and functional data depths; see Zuo and Serfling (2000) and Mosler and Polyakova (2012) for comprehensive reviews. These properties ensure that a depth function properly measures the notion of depth or centrality. For instance, a depth function needs to be location and scale invariant (or equivariant) and it should decrease monotonically from a natural point of symmetry. Since our depth is purely for functional data we concentrate on the central properties of Mosler and Polyakova (2012). These properties are established for the amplitude depths because amplitude is the primary concern of shape analysis.

The elastic depths are based on proper distance metrics so they inherit certain properties such as translation invariance and scale equivariance automatically. On some manifolds, such as $\mathbb{R}^2$, scale equivariance can be promoted to scale invariance because the trajectories are constrained to live on an L^2 ball. Invariance to simultaneous reparameterization (simultaneous phase invariance) was shown in Srivastava et al. (2011) for amplitude distances between $\mathbb{R}$ and $\mathbb{R}^n$ valued trajectories and then later extended to $\mathbb{S}^2$ valued trajectories in Su et al. (2014). Consequently, the amplitude depths are also invariant to simultaneous reparameterization.

Other properties, such as phase invariance, maximality of the center, and convex level sets are essential for shape anomaly detection but are not simple corollaries of the amplitude distance. We outline these properties, as they apply to amplitude depth, below. All proofs are deferred to the online supplement Section A.2.

Proposition 3.1 (*Phase invariance*). Let Γ be the space of warping, or phase, functions defined in Section A.1 and let $\gamma \in \Gamma$. Let F_M be the space M-valued differentiable functions as in Section 3.1, let $f \in F_M$ and suppose P is a distribution supported on F_M. Then

$$D_a(f \circ \gamma, P) = D_a(f, P),$$

where $D_a(\cdot, P)$ is the amplitude depth of trajectories on F_M with respect to P.

This property is unique to the elastic depths and ensures that the amplitude depths are invariant to the phase under which each trajectory is observed. This property, in conjunction with translation and scale invariance (equivariance), means that the amplitude depth is invariant to the shape preserving transformations. We can, therefore, say that amplitude depths are appropriately capturing our definition of shape.

Proposition 3.2 (*Maximality of the center*). Let F_M be the space M-valued differentiable functions as in Section 3.1, let $f \in F_M$ and suppose P is a distribution supported on F_M. A trajectory $s \in F_M$ is the amplitude Fréchet median of P if and only if $s = \arg\max_{f \in F_M} D_a(f, P)$, where $D_a(\cdot, P)$ is the amplitude depth of trajectories on F_M with respect to P.

Maximality of the center guarantees that the maximizer of the amplitude depths, denoted the amplitude depth median, is the actual Fréchet median of the distribution. The Fréchet median is the trajectory that minimizes the median distance between itself and all other points in the space. This property ensures that the amplitude depths start their ordering from the true amplitude center of the distribution.

Proposition 3.3 (*Convex level sets*). Let F_M be as in Section 3.1, let $f \in F_M$ and suppose P is a distribution supported on F_M. Let $D_{a,\alpha}(P) = \{f \in F_M : D_a(f, P) \geq \alpha\}$ be the upper level sets for the amplitude elastic depth for all $\alpha \in [0, 1]$. Then $D_{a,\alpha}(P)$ is a convex set. Similarly the upper level sets for the phase elastic depth $D_{p,\alpha}(P) = \{f \in F_M : D_p(f, P) \geq \alpha\}$ are convex for all $\alpha \in [0, 1]$.

Convexity of the level sets implies that depths decrease monotonically from the center of the distribution. In conjunction with maximality of the center, level set convexity guarantees that the elastic depths are measuring centrality in amplitude space and phase space. This property further distinguishes the elastic depths from previous depth notions because they do not directly characterize centrality in the appropriate shape spaces. We use these convex level sets as the theoretical basis for the construction of the depth boxplots (Section 4.1) and for depth thresholding (Section 4.2).

3.3. Estimating Elastic Depths

As in Section 3.2, let F_M be the space of differentiable functions on the Riemannian manifold M and let P represent a distribution supported on F_M. Suppose we observe $f_1, \ldots, f_n \sim P$. The amplitude and phase depths of each f_i, $i \in 1, \ldots, n$, can be estimated empirically using their respective sample outlyingness functions. The sample amplitude and phase outlyingness functions are, respectively, denoted as $O_{a,n}$ and $O_{p,n}$ and are defined as

$$O_{a,n}(f, P_n) = \text{median}\{d_a(f, f_1), \ldots, d_a(f, f_n)\},$$
$$O_{p,n}(f, P_n) = \text{median}\{d_p(f, f_1), \ldots, d_p(f, f_n)\},$$

where P_n denotes the empirical distribution of the functions $f_1, \ldots, f_n$. Using the same construction as before, we invert the sample outlyingness functions into sample depths

$$D_{a,n}(f, P_n) = (1 + O_{a,n}(f, P_n))^{-1}, \quad (4)$$
$$D_{p,n}(f, P_n) = (1 + O_{p,n}(f, P_n))^{-1}, \quad (5)$$

for amplitude and phase, respectively. The following proposition asserts the uniform consistency of this depth estimator.

Proposition 3.4 (*Uniform consistency*). Let F_M be as in Section 3.1, suppose P is a distribution supported on F_M, let $f_1, \ldots, f_n \sim P$, and let P_n represent the empirical distribution of the sample. Then

$$\lim_{n \to \infty} \sup_{f \in F_M} |D_{a,n}(f, P_n) - D_a(f, P)| = 0 \text{ a.s.},$$
$$\lim_{n \to \infty} \sup_{f \in F_M} |D_{p,n}(f, P_n) - D_p(f, P)| = 0 \text{ a.s.},$$

where $D_a(\cdot, P)$ is the amplitude depth of trajectories on F_M with respect to P, $D_{a,n}(f, P_n)$ is the amplitude depth's empirical counterpart, $D_p(\cdot, P)$ is the phase depth of trajectories on F_M with respect to P, and $D_{p,n}(f, P_n)$ is the phase depth's empirical counterpart.

4. Identifying Outliers

Data depth is a natural framework for outlier detection because it provides a center-outward ordering of the data. Functions with very low depth values are strong candidates for outliers because they are statistically far from the center of the distribution. As mentioned in Section 1, there have been many methods, many based on functional depth in some way, for detecting shape anomalies proposed in the literature. These methods typically construct an outlier cutoff boundary on either the depths or the functions and classify any trajectory as an outlier if it exceeds these bounds. In the next two sections, we introduce two simple ways of defining an outlier cutoff point based on elastic depth.

4.1. Depth Boxplots

The first method we introduce is called the depth boxplot, which is a half-boxplot constructed on the elastic depths directly. We showed in Section 3.2 that the elastic depths decrease monotonically from their unique center, as trajectories become more outlying, so using the depths directly does not incur a loss

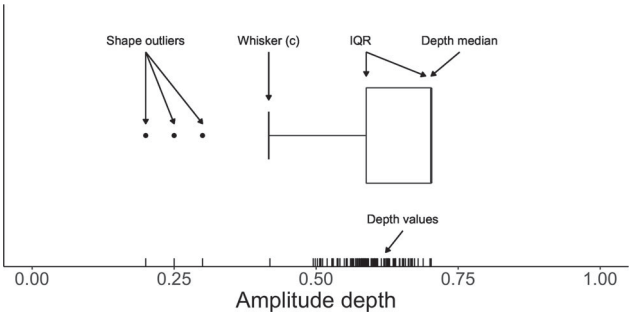

Figure 2. Diagram of the amplitude depth boxplot, created using Algorithm 1 with $k = 1.5$, on example data. The depth median, IQR boundaries, whisker (c), and three shape outliers have all been labeled accordingly. Each of the trajectories' amplitude depths has been plotted along the horizontal axis.

of outlyingness information. Additionally, unlike methods that place bounds on the observed data, using a boxplot on the depth values circumvents the problem of shape outliers being masked due to scale, translation, and phase variability. This is because the boundaries of a depth boxplot correspond to entire central regions on the shape space of functions, and not merely central regions on the projections of functions onto $\mathbb{R}$ (or $\mathbb{R}^2$ or $\mathbb{S}^2$).

Algorithm 1 describes how to construct the amplitude depth boxplot and how amplitude anomalies are identified with the whisker c. Phase anomalies can similarly be defined by substituting amplitude depths for phase depths. The boxplot created in

Algorithm 1: Depth boxplots for finding amplitude outliers

Input : Functions $f_1,\ldots,f_n$ and multiplier k
Output: Outlier status of $f_1,\ldots,f_n$ given k

for $i \leftarrow 1$ **to** n **do**
  Compute amplitude depths $D_{A,n}(f_i, P)$
end
Compute IQR $= \max\{D_{A,n}(f_i, P)\} - \text{median}\{D_{A,n}(f_i, P)\}$
Compute $c = \text{median}\{D_{A,n}(f_i, P)\} - k \times \text{IQR}$.
for $i \leftarrow 1$ **to** n **do**
  if $D_{A,n}(f_i) < c$ **then**
    f_i is an outlier
  else
    f_i is not an outlier
  end
end

Algorithm 1 consists of the following three pieces: the median, the IQR, and the whisker c (Figure 2). The median of the boxplot is the largest depth, because as was shown in Section 3.2, the largest depth corresponds to the median of the distribution. The IQR is the 50% central region because, as per the IQR of univariate data, this range contains the inner 50% of the data. Most importantly is the whisker value c, which determines which trajectories are considered outliers. Any trajectory with an amplitude depth of less than c is considered an anomaly because it is statistically too far from the rest of the data.

The whisker c is determined by a multiplier or inflation factor k. The quantity k is a free parameter that must be set to detect anomalies. In classical univariate boxplots, $k = 1.5$, so as to achieve approximately 99.3% coverage of the boxplot on Gaussian data. This guarantee does not necessarily extend to functional data, but we find empirically that $k \in [1.5, 2.25]$ works well to separate outliers from inliers as long as the sampling frequency is high enough to fully represent the functional data. We investigate the depth boxplot's dependency on k numerically in Section A.5 of the online supplement and find that detection performance is fairly robust to k for a wide range of values and across many types of data. More details on the choice of k can also be found in Section 5.2.

4.2. Depth Thresholding

Boxplots and other hard cutoffs are not the only way to investigate shape anomalies. The quantiles of the depth distribution itself can also be used to investigate the most extreme data. The elastic depths induce a proper center-outward ordering of the trajectories, so the most extreme, that is, smallest, depth values correspond to the most extreme trajectories. Therefore, if we wanted to view the 5% most extreme functions, we could simply select the functions with the 5% smallest elastic depth values. More generally, the $100 \times (1-p)$% most outlying functions have depth values below the $(1-p)$th quantile of the depth's distribution. This type of thresholding is quite useful in exploratory analysis for comparing, say, the 1%, 5%, and 10% most outlying shapes with the 1%, 5%, and 10% most inlying shapes.

The limitation of thresholding is that it will always select $100 \times (1-p)$% of the data to be outlying, so as an anomaly detector it is insufficient on its own. However, it can be paired with the depth boxplots to produce a more robust depth boxplot. Algorithm 2 extends Algorithm 1 to include a thresholding parameter p so that to be considered an outlier, a function must have a depth value below the whisker c and below the $(1-p)$th quantile of the depth's distribution.

Algorithm 2: Depth boxplots for finding amplitude outliers with thresholding

Input : Functions $f_1,\ldots,f_n$, multiplier k, and threshold p
Output: Outlier status of $f_1,\ldots,f_n$ given k and p

for $i \leftarrow 1$ **to** n **do**
  Compute amplitude depths $D_A(f_i, P)$
end
Compute IQR $= \max\{D_A(f_i, P)\} - \text{median}\{D_A(f_i, P)\}$
Compute $c = \text{median}\{D_A(f_i, P)\} - k \times \text{IQR}$.
Compute $q = (1-p)$th quantile of $\{D_A(f_i, P)\}$
for $i \leftarrow 1$ **to** n **do**
  if $D_A(f_i) < \min\{c, q\}$ **then**
    f_i is an outlier
  else
    f_i is not an outlier
  end
end

The purpose of p is mainly to control the number of false positives when detecting outliers. If $p = 0.95$, then at most 5% of the trajectories will be considered outlying, no matter what the whisker value is. While the whisker is generally sufficient for achieving good coverage of the depth distribution and detecting anomalies, there are situations, such as low sampling frequency (Section A.5 of the online supplement), where the whisker can

fall short unless the multiplier k is set higher. In these situations, p will act to effectively increase the k so that better coverage is achieved.

5. Simulation Study

A simulation study was conducted to comprehensively assess the performance of the elastic depths and the depth boxplots. We compared our method against nine other shape anomaly detectors: the OG (Arribas-Gil and Romo 2014), sequential transformations (ST-T1, ST-T2, ST-D1) (Dai et al. 2018), the functional outlier map (FOM) (Rousseeuw, Raymaekers, and Hubert 2018), total variation depth (TVD) (Huang and Sun 2016), the magnitude-shape (MS) plot (Dai and Genton 2018), the robust functional tangential angle pseudo-depth (rFUNTA) (Kuhnt and Rehage 2016), order extended integrated depth (FDJ and IDJ) (Nagy, Gijbels, and Hlubinka 2017), geometric boxplots (GEOM) (Xie et al. 2017), and directional outlyingness (DIR) (Dai and Genton 2019). In the following sections comparisons to TVD, MS, DIR, and GEOM are included while the rest are deferred to the appendix. These four methods were consistently the strongest competitors across each of the outlier models. We describe their implementation here briefly.

The TVD outliers were found using the `detectOutlier` function in the `TVD` R package with an empFactor = 1.5. MS outliers were found by computing the **MO** and **VO** quantities then using the `cerioli2010.irmcd.test` function in the `CerioliOutlierDetection` R package to compute the boundary with a coverage probability of 99.3%. DIR outliers were found using the authors `dir.out` function with Mahalanobis distance and the default parameters fac = 0.154, and cutoff = 6.91. GEOM outliers were found using the `AmplitudeBoxplot` function in the `fdasrvf` R package using $k = 1$. Implementation details for the detectors that are deferred to the appendix have likewise been deferred to the appendix (Section A.4 of the online supplement). Elastic depths (ED) outliers were identified using the depth boxplots with $k = 1.8$. Boxplots were computed on both the amplitude and phase depths separately. The results using amplitude depth are denoted as ED-A and the results using phase depths are ED-P.

5.1. Simulation Design

We define seven different shape outlyingness scenarios to test the effectiveness of the above shape outlier detectors. Each of these scenarios is represented by one of the seven models detailed below. The first six correspond to amplitude (shape) outliers while the seventh is for phase outliers.

1. Model 1 (Amplitude increase). Main model: $X(t) = \sin(5\pi t) + 4t + e(t) + \delta$ and contamination model: $X(t) = 4\sin(5\pi t) + 4t + e(t) + \delta$, where $t \in [0,1]$, $e(t)$ is a centered Gaussian process with covariance function $\gamma(x,x') = \exp\{-(x-x')^2/0.5\}$, and $\delta \sim N(0,1)$ is a random additive translation term. The purpose of δ is to shift each curve by a random amount so as to mask shape outliers that could accidentally be identified as magnitude outliers.
2. Model 2 (Amplitude decrease). Main model: $X(t) = \sin(5\pi t) + 4t + e(t) + \delta$ and contamination model: $X(t) = \frac{1}{6}\sin(5\pi t) + 4t + e(t) + \delta$, where $t \in [0,1]$, and $e(t)$ is the Gaussian process from Model 1.
3. Model 3 (Mixed polynomials). Main model: $X(t) = t^3 - 2t^2 + 0.5t + e(t)$ and contamination model: $X(t) = 2t^3 + t^2 - 0.5t + e(t)$.
4. Model 4 (Covariance change). Main model: $X(t) = \sin(5\pi t) + 4t + e_1(t) + \delta$ and contamination model: $X(t) = \sin(5\pi t) + 4t + e_2(t) + \delta$, where $t \in [0,1]$ and $e_1(t)$ and $e_2(t)$ are centered Gaussian processes with covariance functions $\gamma(x,x') = \exp\{-(x-x')^2/50\}$ and $\gamma(x,x') = \exp\{-(x-x')^2/2\}$, respectively.
5. Model 5 (Frequency increase). Main model: $X(t) = \sin(2\pi t) + 4t + e_1(t) + \delta$ and contamination model: $X(t) = \sin(12\pi t) + 4t + e(t) + \delta$ where $t \in [0,1]$ and $e(t)$ is the Gaussian process from Model 1.
6. Model 6 (Jump contamination). Main model: $X(t) = \sin(5\pi t) + 4t + e(t) + \delta$ and contamination model: $X(t) = \sin(5\pi t) - 2\mathbb{1}_{(t<T)} + 3\mathbb{1}_{(T\le t)} + 4t + e(t) + \delta$, where $t \in [0,1]$ and T is distributed uniformly on $[0.4, 0.6]$.
7. Model 7 (Phase contamination). Main model: $X(t) = \sin(5\pi t) + 4t + e(t) + \delta$ and contamination model: $X(t) = \sin(5\pi\gamma(t)) + 4\gamma(t) + e(\gamma(t)) + \delta$, where $t \in [0,1]$ and γ is a random phase function from Γ. The functions γ are generated from the first two Fourier basis functions with random amplitudes distributed as $N(0,\sigma)$ on the tangent space to the unit Hilbert sphere. We use $\sigma = 6$ to impose a large amount of phase variability on the contamination model (Figure 5).

Each of these models, except model 7, was then further contaminated with two additional sources of noise: compositional (phase) noise and magnitude outliers. Compositional noise was added by composing each trajectory with a random phase function generated by the `rgam` function in the `fdasrvf` package with sigma = 0.1. Magnitude outliers were added by randomly shifting 10% of the generated functions by ± 10. Model 7 was only contaminated with magnitude outliers because adding phase noise would destroy the difference in phases that we are trying to detect. These two noise sources introduce a level of realism to our simulations because nuisance phase and magnitude outlyingness are often present when analyzing shapes. The base amplitude outlier models without additional phase and magnitude noise are pictured in Figure 3.

5.2. Contamination by Multiple Anomalies

We considered the case when 10% of the data is outlying in shape. We compared the performance of the detection methods on the seven outlier models using the F_1 score (Chinchor 1992) for outlier classification. The F_1 score is a comprehensive measure of classification accuracy that considers both the precision (positive predictive value) and the recall (true positive rate) of a detection method. A method that perfectly classifies all outliers as outliers and all inliers as inliers will have an F_1 score of 1. Methods that do not perfectly classify will have F_1 scores less than 1. The F_1 score was traditionally defined as the harmonic mean of precision and recall, but it can also be expressed in terms of true positive (TP), false negative (FN), and

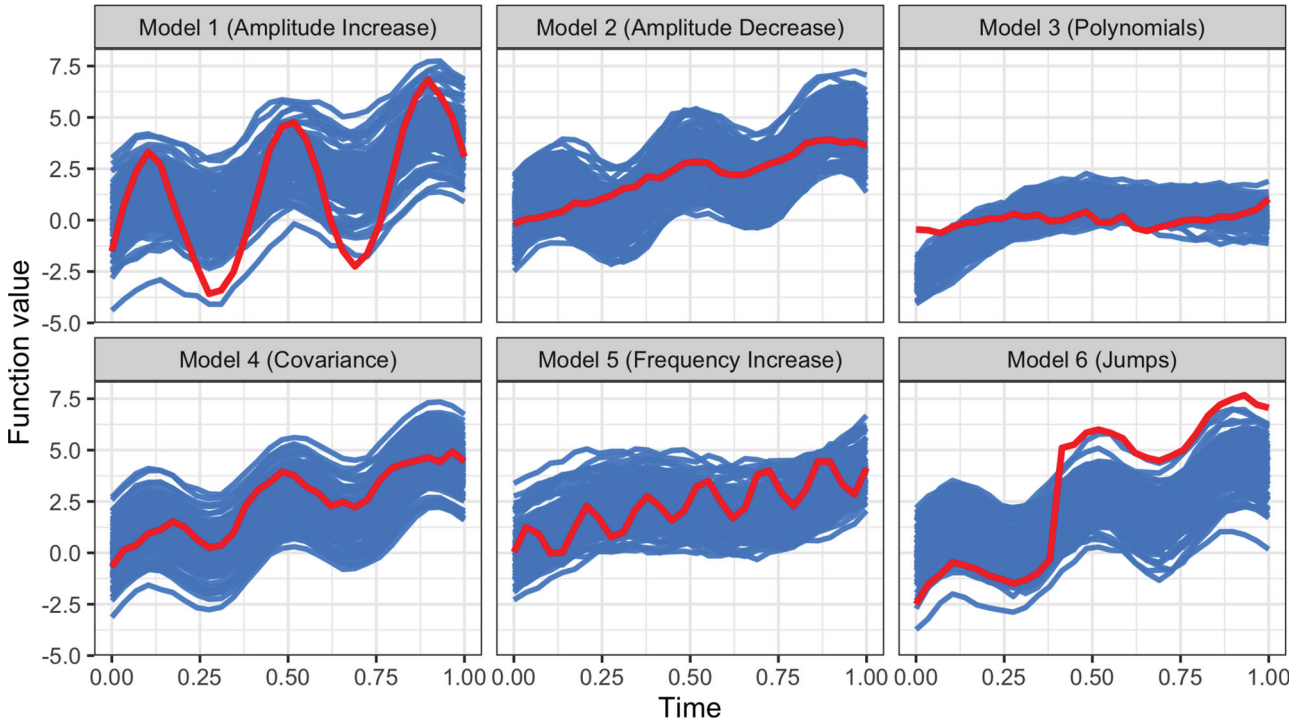

Figure 3. Main model (blue solid lines) versus Contamination model (red lines) in each of the amplitude outlier models.

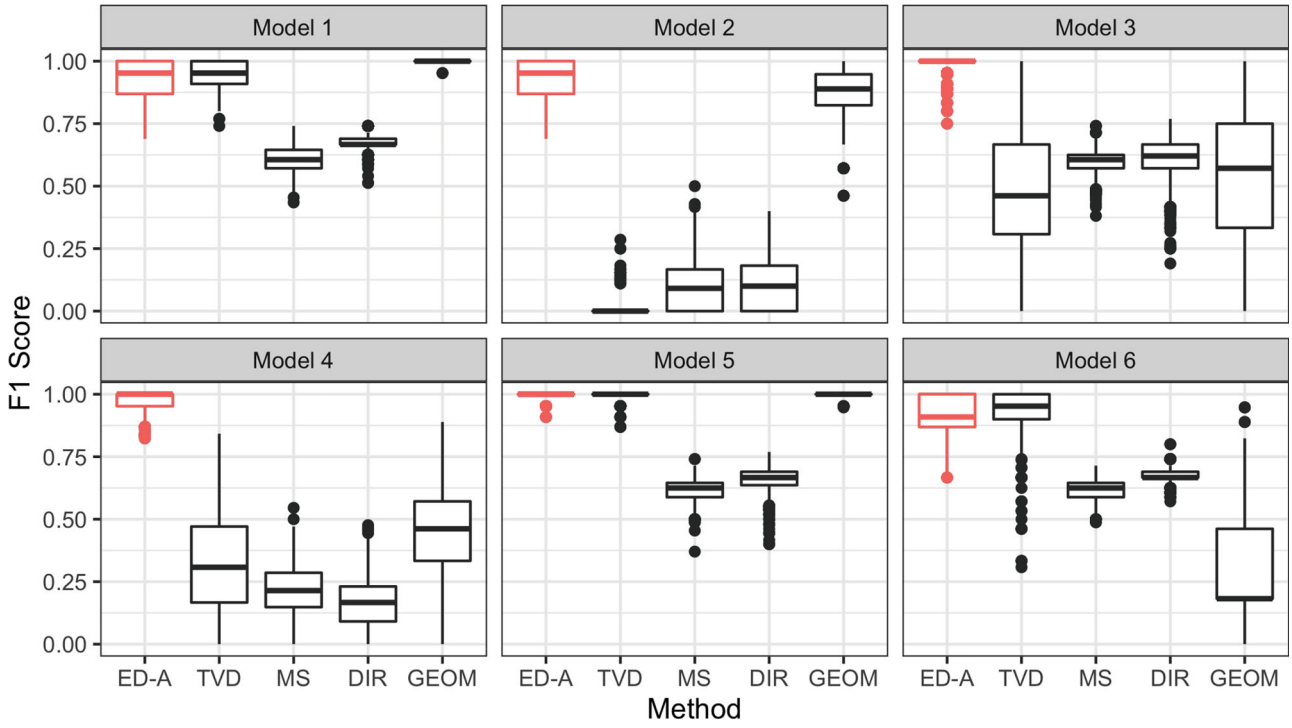

Figure 4. F1 score comparison of the top five models (ED-A, MS, DIR, TVD, and GEOM) on each of the six amplitude outlier models.

false positives (FP) quantities:

$$F_1 = \frac{2\text{TP}}{2\text{TP} + \text{FN} + \text{FP}}.$$

90 inlying trajectories and 10 outlying trajectories were sampled from the main model and contamination model, respectively. Compositional noise and magnitude outliers were again added to each of the models, except for model 7 (phase contamination) where only magnitude outliers were added. Trajectories were sampled on an equidistant 30 point grid over $[0, 1]$ and 1000 simulations were performed for each of the six models. The results for the top models are summarized in Figure 4. Full results for all considered models are available in the appendix.

Figure 4 shows that, across the six amplitude outlier models, the elastic depths maintained the highest average F_1 score. The amplitude depth based boxplots have an average F_1 score of around 0.95–1.0, indicating that in each scenario, regardless of

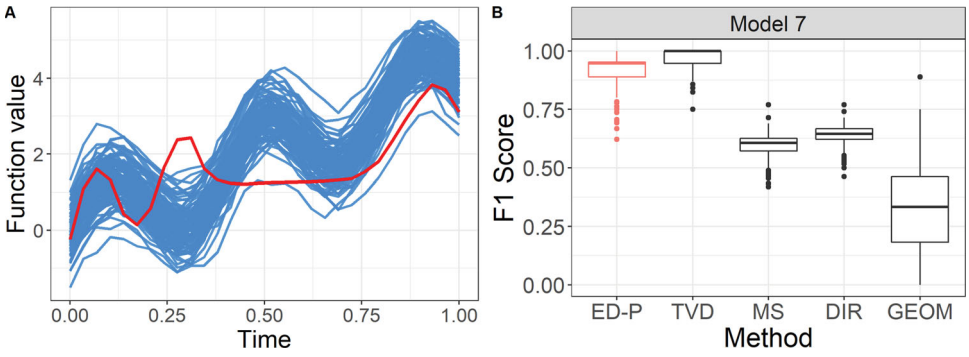

Figure 5. Panel A: Example inlier trajectories (blue) versus a phase outlying trajectory (red). Panel B: F_1 score comparison of the top five models (ED-P, MS, DIR, TVD, and GEOM) on the phase outlier model. ED-P and TVD both have nearly perfect detection rates, with TVD slightly winning out in this simulation.

outlier or inlier type, the amplitude depths can achieve near-perfect detection. Consistently high performance of the elastic depths is notable because existing methods, while strong in some cases, suffer major losses of power in others. For instance, on some models, such as model 3 (mixed polynomials) and model 4 (covariance change), the amplitude depth based boxplot was the only method able to detect the shape outliers consistently. Even GEOM, which uses the elastic distances, was unable to consistently separate these outliers from the inliers. Together, the results show that because the elastic depth based boxplots use the shape distribution, albeit indirectly via data depth, they can consistently and skillfully detect a wide variety of shape outliers. They do not generally suffer a loss of power due to compositional (phase) noise, translation noise, presence of magnitude outliers, or even inlier and outlier type.

Figure 5 shows that, across the top phase outlier detection models, ED-P, the phase counterpart of ED-A, has near-perfect detection skill. The TVD, however, maintains a slightly higher F_1 score. Our conjecture is that the data generation process may have incidentally induced non-phase based differences into the phase outliers. TVD, MS, and DIR could potentially take advantage of this non-phase information to improve their score, whereas ED-P and GEOM could not because they only use phase information. See a small simulation example in Section A.6 in the online supplement.

We also investigated the sensitivity of the boxplots to the parameter k via a simulation study in Section A.5 of the online supplement. Overall, we found that the coverage of the boxplot on the inlying uncontaminated data is insensitive to the value of k. Across each model, the coverage of the boxplots steadily increases from about 95% at $k = 1$ to around 100% at $k = 3$. Some models required higher values of k to achieve the desired 99% coverage, which we found to be due to under sampling, that is, sampling below the Nyquist rate, of the observed trajectories. We recommend avoiding this potential issue by ensuring adequate sampling of the trajectories when possible or using $k \approx 2$ when this is not possible.

6. Hurricane Trajectories

Our motivating example (Figure 1) comes from the National Hurricane Center's (NHC) Atlantic Hurricane Database (HURDAT2) (Landsea and Franklin 2013). The NHC assimilates all observations, real time and post-storm, for each tropical cyclone to estimate and record its characteristics and path across the Atlantic Ocean. The HURDAT2 database contains records for 979 tropical cyclone paths of various lengths, shapes, sizes, orientations, and placements. We only consider storms with at least 25 observations because only those paths had sufficient time to develop. The storms were then further subset to include only those originating in the ocean and north of South America.

The typical path of a hurricane is "U" shaped, starting in Africa then cutting across the eastern United States and finally heading back east toward Europe. Due to the vast distances hurricanes travel it would be inappropriate to treat them as lying on a Euclidean plane. Instead, we consider them as trajectories on the surface of a unit sphere $\mathbb{S}^2$. We used the elastic depth boxplots with a depth threshold of 0.05 to limit the number of amplitude outliers to fourteen; the top four of which are pictured in Figure 6.

Each of the top four outliers is markedly different from the standard "U" shape. They exhibit an atypical spiraling behavior as they meander across the Atlantic. The identification of shape outliers helps climate scientists further investigate what causes the trajectories to be anomalous. It can be very important for improving the accuracy of hurricane prediction algorithms if the dynamics which produce anomalies are well understood. Further data examples on $\mathbb{R}$ and $\mathbb{R}^2$ valued trajectories can be found in Section A.7 of the online supplement.

7. Discussion

In this article, we proposed a new class of functional depths based on the elastic distance metrics and showed how they may be used to detect shape outliers. The theoretical properties of our new elastic depth were investigated, and it was shown that they satisfy most key properties required of a depth metric. These include translation, scale, rearrangement, phase invariance (equivariance), maximality of the center, convex level sets and monotonicity from the center. Rearrangement invariance and phase invariance were particularly crucial for detecting shape anomalies because they allowed the amplitude depth to measure centrality independent of phase.

We demonstrated the empirical performance of our method together with nine competing methods using extensive simulation studies. It was shown that our method attains the overall highest average F_1 score across all models. On each of the six amplitude models the amplitude depth based boxplots had an

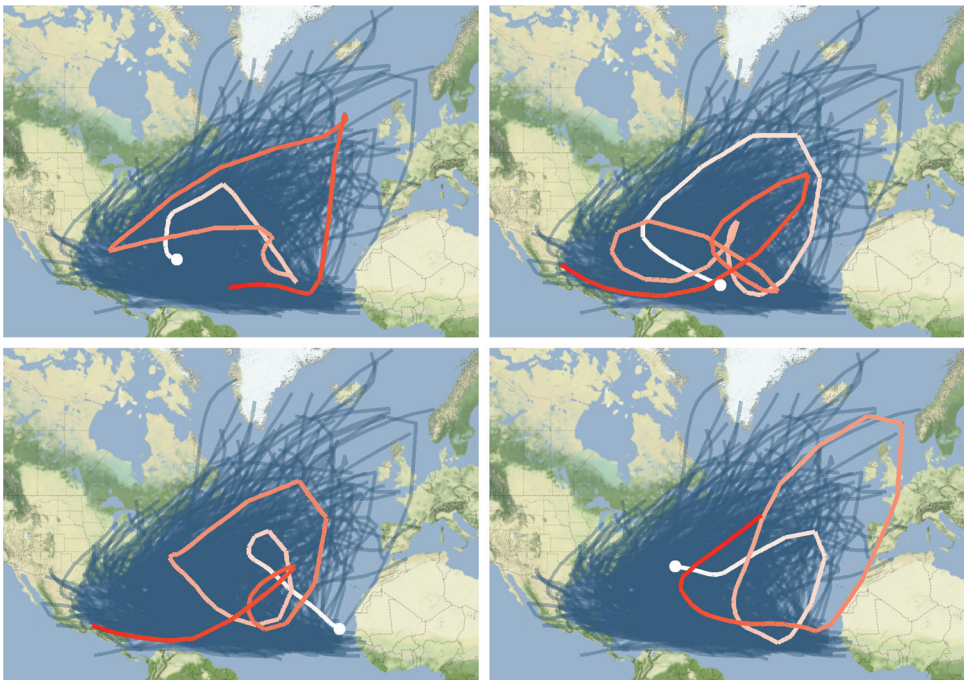

Figure 6. Four most shape outlying hurricane trajectories from the HURDAT2 data overlaid on the entire dataset. The starting point for each track is marked by a point and trajectories become progressively darker as they develop.

average F_1 score of around 0.95–1.0, indicating that in each scenario, regardless of outlier or inlier type, the amplitude depths achieved near-perfect detection on average. On some models, such as model 3 (mixed polynomials) and model 4 (covariance change), the amplitude depth based boxplot was the only method able to detect the shape outliers consistently. Together, these results demonstrate the power of using the shape distribution, albeit indirectly via data depth, to detect shape outliers.

The simulation results, of course, depend on the boxplot multiplier k. We recommend setting $k = 2$ as the default value. We found through empirical studies (see online supplement Section A.5) that any k value between 1.5 and 2.25 provides nearly the same level of detection skill. The value $k = 2$ had the most favorable trade-off between false positives and false negatives across all of the outlier models in the simulation. Larger values (up to around 2.5–2.8) are helpful when the sampling frequency of the trajectories is too low (see online supplement Section A.5). If the sample size is large then the depth quantiles, p, can also be used to help set a minimum value for k. That is, choose k large enough to make the whisker value less than the $(1 - p)$th quantile of the amplitude depth. This guarantees that the shapes of the inner $p \times 100\%$ of trajectories will be considered as inliers.

Finally, we showed how the elastic depths may be used to identify shape outliers in functional data observed on the unit sphere $\mathbb{S}^2$. We used the HURDAT2 hurricane track database and identified the four most shape outlying trajectories. We found that these trajectories' paths were remarkably different from the standard U-shaped paths that hurricanes normally follow. Further applications in the online supplement A.7 demonstrate the elastic depths on $\mathbb{R}$, $\mathbb{R}^2$ valued functional data. These data examples illustrated the simplicity and consistency with which the elastic depths may be applied, regardless of the underlying geometry of the space.

Supplementary Materials

Supplementary results: PDF file "Supplementary results" containing detailed descriptions of the elastic distances, proofs for all results in Section 3.2, additional simulations, and two additional real data examples. (pdf file).

Source code: Zip file "Source code" containing code to produce all figures and results in the article and supplementary file. (zip file)

R-package for elastic depth: R-package "elasticdepth" containing code to compute the elastic depths and depth boxplots for $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{S}^2$ valued trajectories. Available through GitHub: `trevor-harris/elasticdepth`. (GNU zipped tar file)

Funding

This article describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government. This work was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. Li's research is also partially supported by the National Science Foundation through awards AGS-1602845 and DMS-1830312.

References

Arribas-Gil, A., and Romo, J. (2014), "Shape Outlier Detection and Visualization for Functional Data: The Outliergram," *Biostatistics*, 15, 603–619. [1,2,7]

Chinchor, N. (1992), "MUC-4 Evaluation Metrics," in *Proceedings of the 4th Conference on Message Understanding*, Association for Computational Linguistics, pp. 22–29. [7]

Dai, W., and Genton, M. G. (2018), "Multivariate Functional Data Visualization and Outlier Detection," *Journal of Computational and Graphical Statistics*, 27, 923–934. [1,2,3,7]

——— (2019), "Directional Outlyingness for Multivariate Functional Data," *Computational Statistics & Data Analysis*, 131, 50–65. [2,3,7]

Dai, W., Mrkvicka, T., Sun, Y., and Genton, M. G. (2018), "Functional Outlier Detection and Taxonomy by Sequential Transformations," arXiv no. 1808.05414. [7]

Horváth, L., and Kokoszka, P. (2012), *Inference for Functional Data With Applications* (Vol. 200), New York: Springer. [1]

Huang, H., and Sun, Y. (2016), "Total Variation Depth for Functional Data," arXiv no. 1611.04913. [3,7]

——— (2019), "A Decomposition of Total Variation Depth for Understanding Functional Outliers," *Technometrics*, 61, 445–458. [2]

Hyndman, R. J., and Shang, H. L. (2010), "Rainbow Plots, Bagplots, and Boxplots for Functional Data," *Journal of Computational and Graphical Statistics*, 19, 29–45. [1]

Kuhnt, S., and Rehage, A. (2016), "An Angle-Based Multivariate Functional Pseudo-Depth for Shape Outlier Detection," *Journal of Multivariate Analysis*, 146, 325–340. [2,7]

Kurtek, S., Srivastava, A., and Wu, W. (2011), "Signal Estimation Under Random Time-Warpings and Nonlinear Signal Alignment," in *Proceedings of Neural Information Processing Systems (NIPS)*. [3]

Landsea, C. W., and Franklin, J. L. (2013), "Atlantic Hurricane Database Uncertainty and Presentation of a New Database Format," *Monthly Weather Review*, 141, 3576–3592. [9]

Liu, R. Y., Parelius, J. M., and Singh, K. (1999), "Multivariate Analysis by Data Depth: Descriptive Statistics, Graphics and Inference" (with discussion and a rejoinder by Liu and Singh), *The Annals of Statistics*, 27, 783–858. [2]

Lopez-Pintado, S., and Romo, J. (2009), "On the Concept of Depth for Functional Data," *Journal of the American Statistical Association*, 104, 718–734. [2]

Mosler, K., and Polyakova, Y. (2012), "General Notions of Depth for Functional Data," arXiv no. 1208.1981. [2,4]

Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H., and Hahn, U. (2017), "Global Envelope Tests for Spatial Processes," *Journal of the Royal Statistical Society*, Series B, 79, 381–404. [1]

Nagy, S., Gijbels, L., and Hlubinka, D. (2017), "Depth-Based Recognition of Shape Outlying Functions," *Journal of Computational and Graphical Statistics*, 26, 883–893. [2,7]

Ren, H., Chen, N., and Zou, C. (2017), "Projection-Based Outlier Detection in Functional Data," *Biometrika*, 104, 411–423. [2]

Rousseeuw, P. J., Raymaekers, J., and Hubert, M. (2018), "A Measure of Directional Outlyingness With Applications to Image Data and Video," *Journal of Computational and Graphical Statistics*, 27, 345–359. [2,7]

Srivastava, A., and Klassen, E. P. (2016), *Functional and Shape Data Analysis*, New York: Springer. [3]

Srivastava, A., Klassen, E. P., Joshi, S. H., and Jermyn, I. H. (2011), "Shape Analysis of Elastic Curves in Euclidean Spaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33, 1415–1428. [1,2,3,4]

Su, J., Kurtek, S., Klassen, E., and Srivastava, A. (2014), "Statistical Analysis of Trajectories on Riemannian Manifolds: Bird Migration, Hurricane Tracking and Video Surveillance," *The Annals of Applied Statistics*, 8, 530–552. [3,4]

Sun, Y., and Genton, M. (2011), "Functional Boxplots," *Journal of Computational and Graphical Statistics*, 20, 316–334. [1]

Tucker, J. D., Wu, W., and Srivastava, A. (2013), "Generative Models for Functional Data Using Phase and Amplitude Separation," *Computational Statistics and Data Analysis*, 61, 50–66. [2,3]

Tukey, J. (1977), *Exploratory Data Analysis*, Reading, MA: Addison-Wesley. [2]

Xie, W., Kurtek, S., Bharath, K., and Sun, Y. (2017), "A Geometric Approach to Visualization of Variability in Functional Data," *Journal of the American Statistical Association*, 112, 979–993. [2,7]

Yu, G., Zou, C., and Wang, Z. (2012), "Outlier Detection in Functional Observations With Applications to Profile Monitoring," *Technometrics*, 54, 308–318. [2]

Zuo, Y., and Serfling, R. (2000), "General Notions of Statistical Depth Function," *The Annals of Statistics*, 28, 461–482. [2,4]