Bow Echo Alarm System Using Topological Data Analysis

Helene Canot, Philippe Durand, Emmanuel Frénod

Bow Echo Alarm System Using Topological Data Analysis

Hélène Canot[1] Philippe Durand[2] Emmanuel Frénod[1]

Abstract - This study focuses on the extremely severe storm system swept across Corsica early in the morning of August 18, 2022. A derecho refers to thunderstorm episodes that are distinguished by the production of particularly long-lasting, strong and widespread downbursts, generated by an extratropical mesoscale convective system (MCS). The associated radar signatures generally exhibit linear characteristics, with arcuate portions (bow echo). This work studies in particular the life cycle of the bow echo that appeared off the coast of Corsica and caused significant damage on the island as well as victims including five deaths. The bow echo is an exceptional phenomenon that is very difficult to predict because it is very localized. Beyond qualitative observations, methods for detecting the shape of a bow echo from radar images have been studied. We present an alternative method for characterising the bow echo shape using a promising tool based on topological data analysis (TDA) and the measurement of the curvature of the cloud. The use of homological persistence from TDA allows to follow the evolution of the bow echo from its initiation to its maturation just before it hits the Corsican coast. Curvature and persistent homology calculation with the stability property of Wasserstein and Bottleneck distances allows us to detect changes between evolutionary phases. The main idea is to exploit TDA on radar images of this bow echo in order to generate an alert. TDA is one of these new mathematical tools that will help characterize the event.

Keywords - Bow echo, Derecho, Topological Data Analysis, persistence diagram, homology, Bottleneck distance, Wasserstein distance, warning.

[1]Université de Bretagne-Sud, UMR 6205, Laboratoire de Mathématiques de Bretagne Atlantique, F-56000 Vannes, France

[2]Conservatoire National des Arts et Mêtiers, Département de Mathématiques et Statistiques, M2N, 292 rue Saint Matin, 75141 Paris cedex, France

1 Introduction

A storm system, formed during the night of August 17 to 18, 2022 over the north of the Balearic Islands, quickly moved towards the northeast, affecting Corsica early in the morning, producing strong gusts of wind on the surface and causing the death of five people and very significant damage. The system then continued its route towards northern Italy and Austria. Winds of more than 200 km/h hit Corsica. The intensity of this phenomenon can be classified in the category of derechos, it results in intense storms with wind gusts of more than 120 km/h and extended over an area of more than 300 kilometers. It is defined by particularly long-lasting, strong and extensive downbursts generated by an extratropical mesoscale convective systems (MCS) [14] which are organized in thunderstorms with a linear shape lasting hours accompanied by intense precipitation, lightning and hail. Derechos are phenomena that occur mostly in the Great Plains of the United States, in flat areas where available energy is significant and, more rarely, at sea. The associated radar signatures generally have linear characteristics, with arcuate portions (bow echo) or crescent-shaped with a high reflectivity gradient on the convex edge. Signatures can often be observed in the form of undulation of the main precipitating zone, then a bulging associated with meso-depressive. In 1987, R. Johns and W. Hirt [9] presented a series of criteria to define derechos. But it was in 2016 that Corfidi et al [4] proposed revising the definition of a derecho in order to assign it higher size criteria as well as criteria linked to characteristics identifiable thanks to radar images.

The bow echo is associated with strong convective instability and significant near-surface humidity and strong vertical wind shear at low altitude, caused by a rear inflow jet in the middle layers of the atmosphere [11] [25]. During the summer of 2022 a cold drop persisted between Portugal and France. The shift in its position led to the alternation between hot conditions over Portugal and stormy conditions over France. The mesoscale convective system (MCS), formed during the night of August 17 to 18 over the northern Balearic Islands, quickly moved northeast, affecting Corsica early in the morning. The line of storms gradually curved to become an arc echo with the descent of the medium altitude jet on the western flank of the MCS, producing strong gusts of surface wind (record broken with 225 km/h at Marignana). Recently, the severity of the storm that developed over the Mediterranean Sea in August 222 is associated with anthropogenic climate change. and was probably amplified by the extreme marine heatwave [8].

Radar reflectivity signatures have been used in very severe-storm. Fujita [11] discovered in 1978 a bow echo radar signature associated with downbursts. Two distinct scales of bow echo exist the squall line type bow echo (SLBE) and the single-cell type bow echo (CBE). Lee et al. (1992) evaluated in the complete quan-

titative study of the evolution and structure of a CBE. The early evolution of bow echoes from radar observations was studied by Klimowsky et al. [16] (2004). They characterized three initial modes of formation of bow echoes: weakly organized cells, squall line, and super cells. In [17] Klimowski et al. (2003) observe that bow echoes are preceded by thunderstorm mergers roughly $40, 50$ percent of the time.

There are numerous references in the literature regarding the consequences and effects of bow echoes [5], [17], and recently algorithmic tools have been developed for the detection and prediction of bow echoes. In [24] Patil et al. present various algorithmic methods for detecting and monitoring MCS, citing their advantages and limitations. Kamani et al. in [15] proposed an automatic framework to detect Bow echo patterns in radar images with a new skeletonization method and shape matching approaches. Another bow echo shape recognition method is given by Fanlin and Jinyi [10] by measuring the outline of the cloud to find its curvature. Recently Mounier et al. propose a CNN neural network detecting a bow echo using pseudo reflectivity.

In this paper, the life cycle of bow echo is studied with the help Doppler weather radar. The vast majority of studies on the evolution of bow echo concern the United States since it is one of the most affected countries in the world. But other bow echoes are observed in different regions of the world and in particular in Europe [13]. The evolution and the dynamic of the life cycle of the bow echo event on August 18, 2022 in France, are studied using Topological Data Analysis (TDA). TDA is a promising tool in the meteorological field and in particular in severe weather phenomena. The applications of TDA, in particular homological persistence in meteorology, are very recent. Muszynski et al. [23], studies "atmospheric rivers" using topological data analysis and machine learning. Tymochko et al. [28] also used homological persistence to quantify a diurnal cycle of tropical cyclones. A new use of topological data analysis to follow this cycle over time made it possible to detect a cycle repeating approximately every 24 hours. Using persistent homology and through different examples, Strommen et al. [27] assimilate the existence of meteorological regimes in a dynamic system to the existence of a non-trivial topological structure of the system's attractor. Similarly, Maletic et al. [19] explore dynamic systems and demonstrate that that persistent homology locate the holes of the Lorenz'63 system.

The storm front which formed, in a few minutes, on the western side of the island surprised everyone with its power. In this review, the evolution of the bow echo which struck Corsica will be established through two criteria: persistent homology and the radius of curvature of the arc of the bow echo. Radar images of the life cycle are taken every 10-15 minutes, monitoring of the evolution of the cloud is done in two phases: the first, will make it possible to investigate the formation of

the structure of the bow shape and in the second phase we will monitor the maturity of the cloud. To achieve this, we create the contours of the main cell of the cloud as well as those which correspond to the highest rainfall. The corresponding point clouds are then analyzed according to the persistence diagrams, Bottleneck and Wasserstein distances. These criteria will make it possible to identify shapes, patterns for each image and each time step. In this paper, encouraged by the potential of TDA to identify shapes and patterns, we will focus on the characterization of the evolution of cloud of the bow echo in order to allow an alert to be set at the right time.

This paper is organized as follows : Section 2 summarize the mathematical methods that we used. Section 3 reviews the data set and then presents the different methods used with homology persistence and curvature calculation. Section 4 presents the analysis of the results and we propose an alert in section 5. Finally section 6, we conclude and open up to different perspectives.

2 Mathematical backgroud

The interest in this study is to be able to characterize storm cells, in particular the most severe ones, quantifying them and to follow their evolution. The approach is based on TDA thanks to persistent homology, it shows the duration of persistence of topological properties along a space filtration, and enables pattern recognition. In this paragraph we summarize the definitions and concepts necessary for understanding persistent homology.

2.0.1 Simplicial homology

Here we present a version of homology where the space is described by a simplicial complex. Consider a set $\mathcal{P} \in \{p_0, ..., p_k\}$ of points in the space $\mathbb{R}^d$. An affine combination $\sum_i \lambda_i p_i$ is said to be convex when all scalars λ_i are positive and $\sum_i \lambda_i = 1$. The set of convex combinations forms the convex hull of $\mathcal{P}$. A *k-simplex* σ is the convex hull of k + 1 affinely independent points $\{p_0, ..., p_k\}$ of $\mathbb{R}^d$. We then denote $\sigma = conv(p_0, ..., p_k)$ and we say that σ is of *dimension k*. For example 0-simplex is a point, a 1-simplex is a line segment, a 2-simplex is a triangle, and a 3-simplex is a tetrahedron. A *face* of the *k-simplex* is the convex hull of a non-empty subset of $\{p_0, ..., p_k\}$. When τ is a face of σ, we denote $\tau \leq \sigma$. Let V be a set of vertices, a *simplicial complex* on V is a set K composed of subsets of V (the simplices) such that any $\tau \in \sigma$, nonempty, $\tau \in K$. The points p_i which

generate the *simplices* of K are called the vertices of K. the dimension of a simplex $dim(\sigma)$ is the number of vertices of σ minus 1.

If K is a *simplicial complex*, we denote by K_k the set of simplices of dimension k, we define the k-chains of K, these are linear combinations of k-dimensional simplex. It is convenient to work on the field $(\mathbb{Z}/2\mathbb{Z}, +, x)$ of integers modulo 2. For $0 \leq k \leq n$ the k-chains are written formally by:

$$c = \sum_{i=1}^{n_k} \lambda_i.\sigma_i, with : \lambda_i \in \mathbb{Z}/2\mathbb{Z} \tag{1}$$

and form vector spaces over the field $\mathbb{Z}/2\mathbb{Z}$. $(\sigma_1, ..., \sigma_{n_k})$ is the set of *k-simplices* of K.
We can define the boundary of the k-dimensional simplex $\sigma_k = conv(p_0, p_1, ...p_k)$:

$$\partial_\sigma = \sum_{i=1}^{k} conv(p_0, p_1, ..., \widehat{p_i}, ...p_k) \tag{2}$$

where $conv(p_0, p_1, ..., \widehat{p_i}, ...p_k)$ denotes the $(k-1)$ face of σ not containing the vertex p_i. The *boundary operator* ∂_k expands to a linear map from $\mathcal{C}_k(K)$ to $\mathcal{C}_{k-1}(K)$. $\mathcal{C}_k(K)$ is the linear application which associates at the *k-chain* the sum of the edges of its simplices: $\partial c = \Sigma \lambda_i \partial \sigma_i$. We can verify that the boundary of the boundary of a simplex is zero, for all $k \in N$:

The kernel and image of each linear map, ∂_k are vector subspaces.

For $0 \leq k \leq n$, the k-cycles $Z_k(K)$ represent the kernel of ∂_k :

We have the inclusion: $B_k(K) \subset Z_k(K)$. A topological defect is the obstruction for an k-cycle to be an k-boundary: it results in the non-nullity of the *quotient vector space* (homology vector space $H_k(K)$) given by :

In this study, we favor H_1 group which measures the number of holes in the space. The data from cloud cells are represented by point clouds see fig 1 a loop merges for the radius $\epsilon = 0.35$ and it dies for $\epsilon = 1.4$.

2.0.2 Filtration of a simplicial complex

The persistence of a simplicial complex K is calculated according to a filtration of the complex, which consists of ordering all of its simplices in order to decompose it into sub-levels. A filtration of K is defined by a sequence (K_i) of sub-complices of K, ordered by the inclusion:

Topological persistence on the K complex according to this filtration consists in the study of the evolution of the topology of the K_i sub-complexes. The essential

point in the calculation of persistence therefore lies in the chosen *filtration* method. For example in our case we use *Vietoris-Rips filtration*. First we introduce the *Vietoris-Rips Complex*, which contains simplices of bounded dimension.

Given $\epsilon > 0$, the Vietoris-Rips complex $VR(X, \epsilon)$ with vertex set X and the parameter ϵ is defined by:
$\{x_1, x_2, ...x_k\} \in VR(X, \epsilon) \quad \Leftrightarrow \|x_i - x_j\| \leq \epsilon$ for all $i, j \in \{0, 1, ..., k\}$.

As ϵ goes to 0 from ∞ the nested sequence of complexes $VR(X, \epsilon)$ defines the *Vietoris-Rips filtration*.

2.0.3 Persistent homology

Persistent homology was introduce by [7]. Homological persistence on the complex K consists of the study of the evolution of all *subcomplices* K_i of a *filtration*.

We give a *filtration* K_i of a simplicial complex K. We can compute the homology group $H_k(K_i)$. Or for all $j \leq l$ the inclusion $K_j \subset K_l$ induces a linear application. And so the filtration K_i induces a sequence of linear applications

Let $Z_k^j(K)$ and $B_k^j(K)$, the vector subspaces formed from, respectively, the k-dimensional cycles and edges of the subcomplex K_j, so the k-th homology group of K_j is $H_k^j = Z_k^j / B_k^j$. And the *k-th persistent Betti number* of K_l is the dimension of the vector space $H_k^{j,l}(K) = Z_k^j(K) / (B_k^l(K) \cap Z_k^j(K))$.

A homology class c of $H_k(K_j)$ arises in K_j when $c \notin H_k^{j-1,j}(K)$. The homology class disappears in K_l if it vanishes in $H_k(K_l)$ without being zero in $H_k(K_{l-1})$. In particular, in this study, we are interested by 1-dimensional homology which represents the number of loops in the space.

2.0.4 Persistent diagram

Persistence information is represented with *persistence diagrams* (PDs) which summarize the *birth* and the *death* of new holes across Vietoris-Rips filtration. Figures 1 illustrate the mechanism of Vietoris-Rips filtration (VR) with the increasing radius ϵ of the balls.

In the sequence fig 1 we consider the Vietoris-Rips filtration $VR(X, \epsilon)$ which is a process of growing balls of ever increasing radius ϵ around each point, and drawing edges between two points whenever their two respective balls touch for the first time. For radius $\epsilon = 0$ we are reduced to the initial point cloud, each point corresponding to a connected component, which is the 0-dimensional topological feature. As the radius increases $\epsilon = 0.1$ some balls begin to intersect. Then we draw an edge between their center, we build what we call the nerves of the covering.

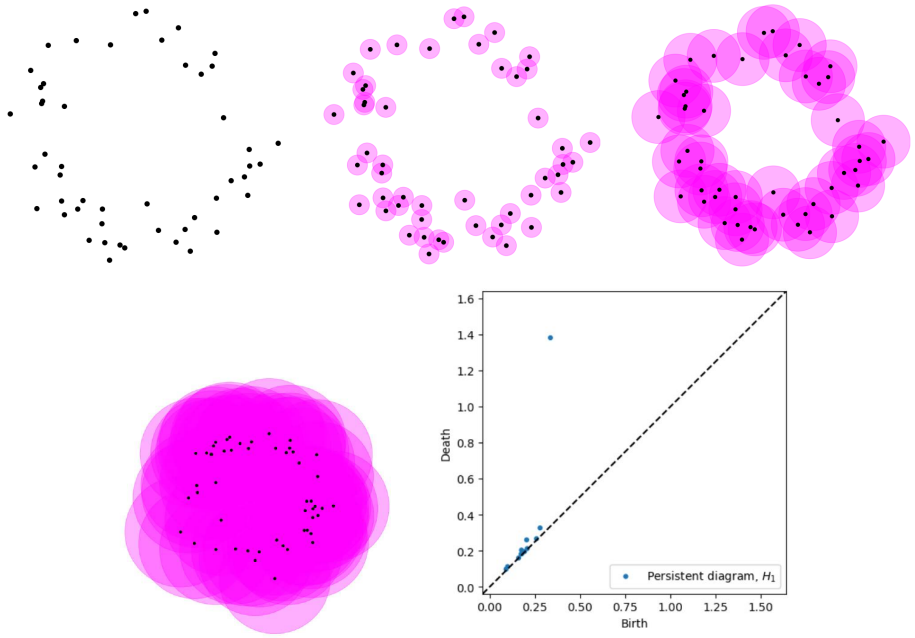

Figure 1: Filtration with increasing values $\epsilon = 0, 0.1, 0.35, 1.4$ (top) of the point cloud X, bottom is the corresponding persistent diagram on the right which each point in 1-dimensional represents the holes

If the intersection of three balls is not empty, we add the simplex constituted by each of the centers of these three balls etc...For each value of increasing ϵ, we associate a simplicial complex and calculate the homology at the resolution epsilon. At $\epsilon = 0.35$ the components have merged into a single connected component; then a new component of dimension 1 i.e a loop has appeared. At $\epsilon = 1.4$ the cycle has been completed which results in its death, all one-dimensional characteristics are dead. These features are visualized in a *persistence diagram*, see right bottom of fig 1 which is a multi-set of points in $\mathbb{R}^2$ with coordinates (a, b) for each feature, with $\epsilon = a$ indicating its *birth time* when the feature appears and $\epsilon = b$ its *death time* when the feature disappears. The difference between death and birth time $b - a$ of a feature is defined as *lifespan* or persistence value. The importance of a topological feature can be also measured in a *lifetime diagram* which represent the same information of *persistence diagram*. To represent a *lifetime diagram* we projected the persistence diagrams onto Euclidean plane. Each point (a_i, b_i) on the persistence diagram is then transformed via a linear transformation $(a_i, b_i - a_i)$. *Persistent diagram* identifies the features and summarizes the changes. This

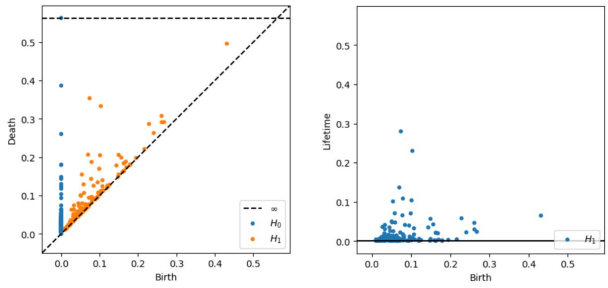

Figure 2: Persistence diagram on the left for H_0 and H_1 features, its corresponding lifetime diagram of H_1 structures on the right

Our interest is in the number of H_1 features which provide information on the number of cells and their surface area. The number of H_1 holes indicates the number of relevant cells. The greater the lifespan of the holes, the greater the surface area of the cell.

2.0.5 Bottleneck and Wasserstein distances

The next step is to compare the persistence diagrams so that we can identify their similarity for each data set. The similarity between two persistence diagrams is measured by the Bottleneck and Wasserstein distances. Computing these distances will allow us to identify and quantify the changes of structures between them.

Given two persistence diagrams X and Y, let η represents the set of all possible mappings between the structures in the diagram, their bottleneck distance is

Then the Wassertein distance between two diagrams is given by

3 Bow Echo features extraction

3.1 Data

In this paper, we use precipitation radar reflectivity images extracted from the meteociel website [20] whose source comes from the Météo France website [21]. Reflectivity is produced by pulsed Doppler radars, it represents the amount of energy that is returned to the receiver after hitting a target. His images are used to visualize areas of precipitation, without providing information on their nature and their liquid or solid state, in real time. They are the representation of different storm attributes, such as hook echoes, bow echoes, and training echoes. The bow echo is found as an arc segment in the reflectivity pattern and can indicate powerful and destructive winds. Reflexivity is a measure of the power returned to the radar, but it can vary significantly depending on the target hit. The radar reflexivity is disturbed by fixed obstacles such as mountains in the immediate vicinity of the antenna. The intensity of precipitation is determined by a color corresponding to millimeters of water per hour. A front is visualized by a band of precipitation. The blue and green colors correspond to a stratiform front. In an unstable rain front, we observe in the stratiform zones small spots of higher intensity (dark green, yellow, orange or red), which visualize showers. A patch of yellow, orange, or red color is representative of heavy precipitation, red suggests precipitation in the form of hail in most cases.

~

We consider the time-varying radar maps of reflexivity $\mathcal{C} = \{C_0, ..., C_n\}$ which contains the sequence of precipitation maps between 7h and 8h (local time). The goal is to identify important structural changes in the precipitation maps of cloud cells $\mathcal{C}$. In particular, we want to indicate in the Timeline the chronology and the appearance of the bow echo.

3.1.1 Pipeline description

Our pipeline description is defined as follows:

1) Obtain the outline in terms of point cloud of the most important precipitation cells of each radar map of reflexivity. Virulent cells representative of heavy precipitation are materialized in the radar maps of reflexivity by a dark red color corresponding to 60 dZ. Each point in the obtained cloud is characterized by its coordinates.

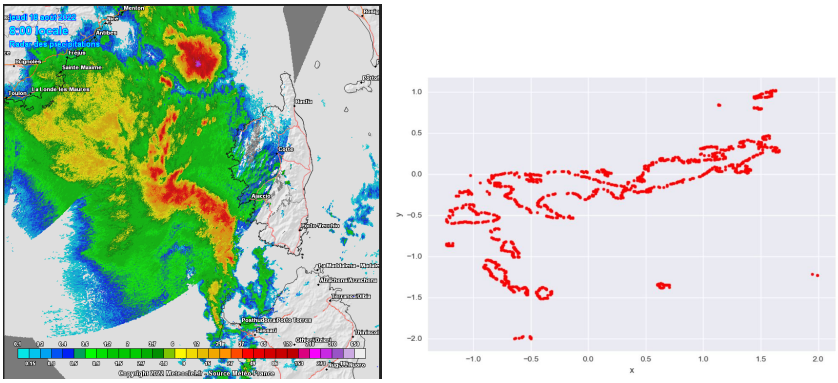

Figure 3: (a) Radar reflexivity of the bow echo at 8 a.m. approaching the Corsican coast, (b) its corresponding point cloud.

2) We evaluate the curvature of the bow echo shape. We calculate the ratio between the arc length and the straight line distance of the two ends of the bow echo.

3) We extract the topological features of the point cloud through the Rips Vietoris filtration by computing the persistence diagrams PD_p in dimension p with $p = \{0, 1\}$. We are particularly interested in the representative H1 structures in our point clouds of storm cells.

4) Compute the Wasserstein and Bottleneck distances in order to compare each pair of persistence diagrams in dimension 1. Measure the similarity between different diagrams using the distance matrix.

5) We plot the upper diagonal elements of the distance matrix.

The radar reflectivity data and the contours of the red part of the bow echo representing the maximum precipitation are shown in Fig 3.

4 Results

4.0.1 Bow echo identification

One of the main characteristics that allows you to identify a bow echo is to measure the contour of the storm cloud for precipitation greater than 40 dBZ. The presence and contours of a bow echo are defined as in L. Zheng [29] et al. and Zhou et al [30]. The arc shape of the Bow Echo is then represented by the ratio between its arc length and the straight line distance of the two ends of the arc describing the cloud. A ratio greater than or equal to 1.2 determines the existence of a bow echo.

In our study, the arc ratio, calculated for each time step, is summarized in the table below:

Time (local)	7h00	7h15	7h35	7h40	7h50	8h00
Arc ratio	0.5	0.62	0.89	1.2	1.32	1.24

We clearly identify a tipping point in the initiation phase from 7h35. With an arc ratio greater than or equal to 1.2 it marks the existence of the bow echo, visible on the reflexivity radar data images. Before the 7h40. threshold, the conditions for obtaining the shape of a bow echo are not acquired but we note that this ratio is increasing to reach 1.2 and more from 7h40.

4.0.2 Persistence diagram features

We analyse the persistence homology applied to the severe precipitation point cloud. We use the Ripser.py library [1] through the Scikit-TDA. Ripser.py is a lean persistent homology package for Python. It is an intuitive interface for Vietoris-Rips filtration with the core of Ripser. Persistence diagrams are generated in figures 4 for H_0 and H_1 features and we summarize the distributions of birth, death, and lifespan values according to the radius from which a structure of edges and hole exists. We visualize the corresponding lifetime diagrams in Figures 5 in 1-dimensional. In Fig 4 the blue and the orange points, respectively, correspond to H_0 and H_1 features. Points on the diagonal are considered noise. The H_0 features, appearing all at the same time at the start of the Vietoris Rips filtration for $\epsilon = 0$, correspond to the outline of the storm cells.

It seems that, according to Fig 4, the number of points H_0 and H_1 are more numerous in the first times of initiation of the bow echo at 7h and 7h15. For a radius between 0 and 0.2, the H_0 structures in each persistence diagram mostly disappear, which indicates the formation of most connected components for these values, few structures persist beyond this. From the formation of the bow echo the H_0 structures have a longer lifespan. The same is true for H_1 structures, most H_1 structures were born and died for a radius between zero and 0.1. In order to refine the qualitative comparison of changes in PDs we will count the H_1 homology for each time and for different increasing ϵ which we summarize in the table below. We consider a radius below 0.1 to be noise, they are not counted in the table, we gradually increase the radius from $\epsilon = 0.1, 0.2$ to $\epsilon > 0.3$. We count each H_1 points in the PDs and we measure the maximal lifespan.

The H_1 structures identified by topological data analysis represent storm precipitation cells. The most persistent H_1 structures in the PDs therefore correspond

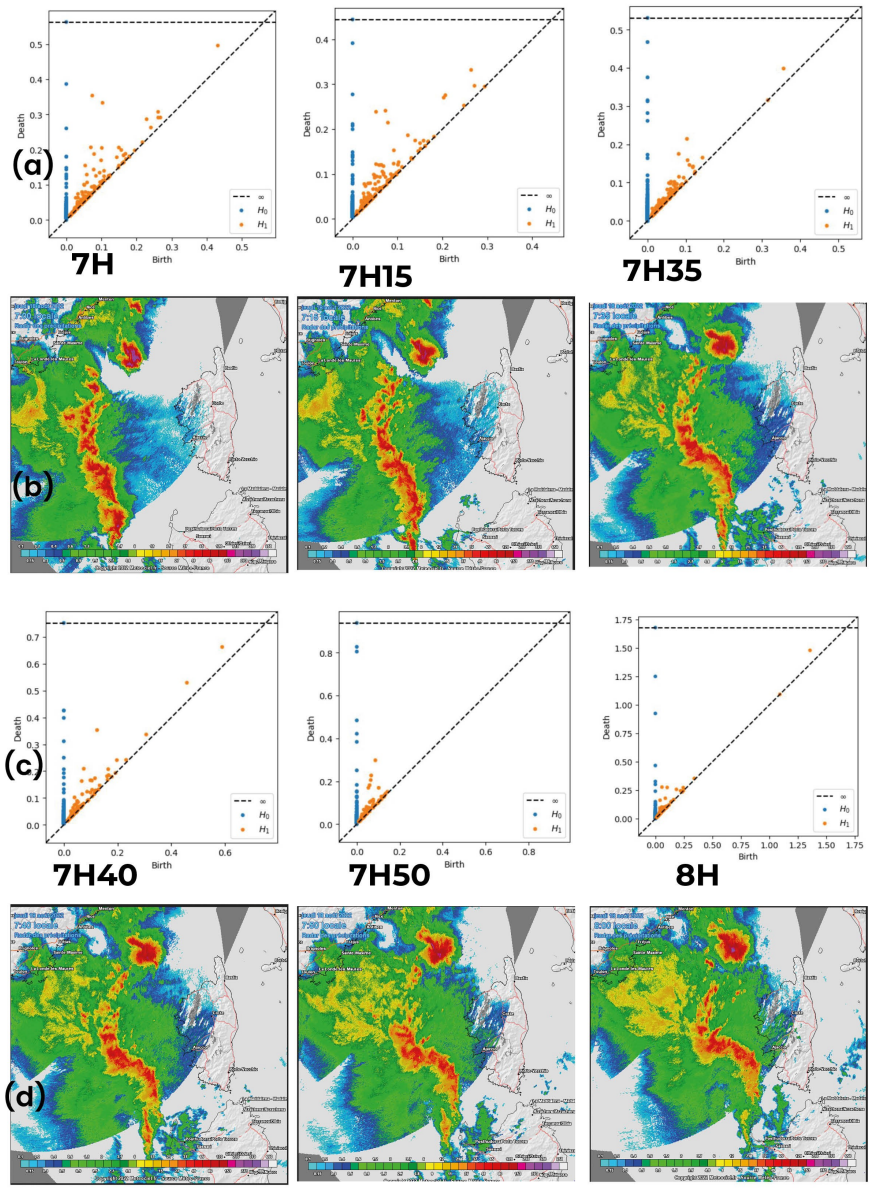

Figure 4: Visual timeline of the bow echo. (b)(d) Radar reflectivity of the forming bow echo with (a)(c) corresponding persistence diagrams.

to storm cells with larger areas. Larger areas are identified by a longer lifespan. The longer the lifespan of a H_1 structures the more likely it is that the H_1 point contains some important information.

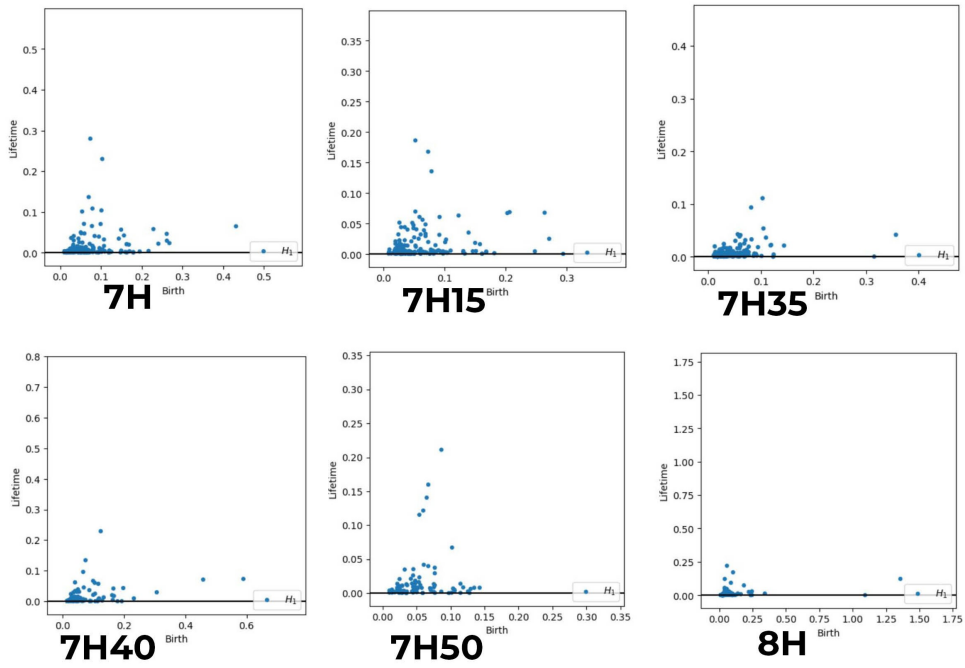

Figure 5: Distribution of H_1 structures in the lifetime diagrams.

The H_1 structures in the lifetime diagrams gives us the indication of the size of the loops. The greater lifespan of the loop is the size of the loop which will give an indication of the size and surface of the storm cell.

Table 1: Numbers of H_1 points as a function of the radius ϵ.

height t	Total number H_1	$0.1 < \epsilon < 0.2$	$0.2 < \epsilon < 0.3$	$\epsilon > 0.3$	Lifespan
$7h$	174	45	7	3	0.28
$7h15$	137	27	8	1	0.17
$7h35$	264	13	1	2	0.11
$7h40$	103	20	4	1	0.23
$7h50$	103	19	2	0	0.21
$8h00$	90	18	4	1	0.22

The PDs in fig 4 at 7h and 7h15 look identical with the same dispersion for H_1 structures, however according to the lifetime diagram their lifetimes are different. At 7h, two significant points with a lifespan of 0.28 are detected in the lifetime diagram. They correspond respectively to the two principal isolated storm cells in the squall line visible in the radar image. We note an inflection in the following PD for 7h15, three loops emerge but with a lifetime of 0.17, significantly shorter than at 7h00. We see on the radar image that the two isolated convective cells have indeed merged and we notice a rear-inflow notch at the back of the cloud synonymous with strong winds. This difference in the topological results marks the evolution of the squall line at the very beginning of the formation of the bow echo.

The PD at 7h35 therefore reveals in H_1 a change in structure, we see significant topological differences between the PD at 7h35 and those at 7h00 and 7h15. It detects a difference in aggregation compared to previous times with much less dispersion in the H_1 points which marks the fusion of the cells. We thus obtain significant signatures in the H_1 distributions for each point cloud of the bow echo evolution. The differences in these distributions allow us to identify the tipping point when the bow echo shape appears. The lifetime of the two significant H1 points is even shorter than those of 7:00 and 7:15. This topological change signals the formation of the bow echo confirmed by the curvature of the arc which continues to increase and the radar image where we see the beginning of an arc with an organized structure with completely fused cells.

We note in table 1, for a radius between 0.2 and 0.4, the number of loops is much less important for the times of 7h40, 7h50 and 8h. The results for these three times according to table 1 are very close. It is also noted, in addition to the number of H_1 structures, that the lifetime of the maximum structures is on average equal to 0.22. Another change visible in the last two maps for 7h50 and 8h: the distribution H_1 is more aggregated less dispersed than the other PDs. Now the curvature of the arc gradually increases to reach the threshold which characterizes a bow echo. The similarity in structural and therefore topological characteristic marks the maturation stage from 7h40 of the bow echo.

4.0.3 Comparing topologies with Bottleneck Wasserstein distances

We compute the Wasserstein and Bottleneck distances in 1-dimensional between the persistence diagrams representative of cloud cells. They allow for pairwise comparison of persistence diagrams.

The interest of these two distances lies in their stability [3]. A small perturbation in the point cloud dataset induces a small change in the persistence diagram. The Wasserstein distance pins down average behaviors while the Bottleneck dis-

tance pins the extreme behaviors. the persistence diagram. The bottleneck distance is the cruder of the two distances and is the reason Wasserstein distances between diagrams are preferred.

Table 2: The Bottleneck distance

Time	7h00	7h15	7h35	7h40	7h50	8h00
7h00	0	0.1126	0.1394	0.0942	0.1062	0.1754
7h15		0	0.0678	0.0828	0.0566	0.0611
7h35			0	0.1145	0.0834	0.0857
7h40				0	0.0553	0.0747
7h50					0	0.0611
8h00						0

Table 3: The Wasserstein distances

Time	7h00	7h15	7h35	7h40	7h50	8h00
7h00	0	1.2017	1.2137	1.1052	1.1327	1.3493
7h15		0	0.7921	1.3034	1.1084	1.2805
7h35			0	0.8174	0.6412	0.7841
7h40				0	0.8134	0.9331
7h50					0	0.8241
8h00						0

Table 4: Ratio of Wasserstein to Bottleneck distances

Time	7h00	7h15	7h35	7h40	7h50	8h00
7h00	0	10.7	8.7	11.7	10.7	7.7
7h15		0	13.6	15.7	19.6	20.95
7h35			0	7.16	7.7	9.2
7h40				0	14.7	12.5
7h50					0	13.6
8h00						0

We compare the variations of the different similarity measures according to the distance matrices. We expect that the distances between the PD1 persistence diagrams corresponding to increasingly distant times are less and less close to zero. However, we see here in the Bottleneck distance matrix that some results do not follow a coherence for the diagrams 7h - 7h40 or 7h - 7h50 for example or the lines 7h15 and 7h35. The results of the Wasserstein distance matrix seem more coherent with the expected results. Note that the ratio of Wasserstein distance to Bottleneck distance is about 10 except for the line 7h15 where the ratio is almost 2 times higher, this being due to the results of the Bottleneck distance on this line.

We now focus on the results of the upper diagonal of the tables for both distances. They appear on both graphs

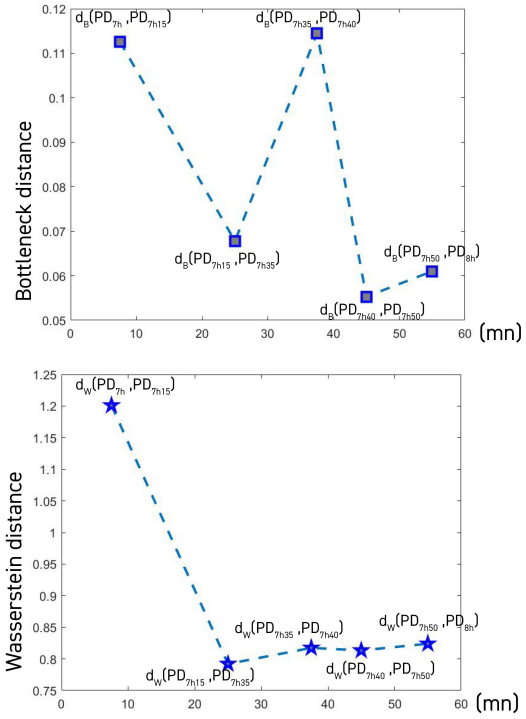

Figure 6: Graph of the Bottleneck distance (top), (bottom) graph of the Wasserstein distance

It is noted in Figure 6 that the Wasserstein distances are of the order of 0.8 for the entire bow echo formation process, they admit linear variations. While the Bottleneck distances showed a significant variation between the persistent diagrams at 7h35 and 7h40 compared to the Wasserstein distance. Both distances captured

some distinct differences during the initiation phase but there is no perturbation for both distances during the maturation phase of the bow echo.

5 Discussion

5.1 Severe thunderstorm warning

Among the difficulties to be understood in terms of bow echo forecasting are the ability to predict when it will form and when it begins the formation process at which time to trigger the alert while minimizing false alarms. Despite the difficulty of the decision it is essential to warn the population of each bow echo. The alert system will therefore represent a compromise defined according to the risks incurred in the event of non-detection of a dangerous event, the consequences in the event of a false alert, and the anticipation time necessary to ensure the safety of people. In our case, the alert system will be based on the results of data topology analysis and calculation of the arc ratio in order to position ourselves on the triggering of the bow echo alert. Two types of alert will be issued: a *potential risk alert* intended to alert people when impressive storm cells form in a line and already present a dangerous character, and a *confirmed risk alert* corresponding to an imminent risk.

First of all, it should be noted that the bow echo was born and developed near the Corsican coast. The difficulty now in the short term is to assess the ideal moment to trigger the confirmed risk alert. First we set the potential risk alert at 7h: a cloud structure with several potentially dangerous cells is in place but the arc has not yet formed. Given the visible difference in the persistence diagrams at 7h35 with more aggregated H_1 structures with a shorter lifespan, an arc begins to emerge. The bow echo gradually evolves at 7h40 the arc is more marked thanks to a marked rear entry while the arc ratio reaches the threshold of 1.2. From 7h40 to 8h00, the H_1 structures are less numerous but with a maximum lifespan for the most persistent ones of the same order. Moreover the Wasserstein distance between these three persistence diagrams $d_W(PD_{7h40}, PD_{7h50})$ and $d_W(PD_{7h50}, PD_{8h})$ gives us equivalent results around 0.8. At 7h50 the curvature has increased and the arc ratio is now greater than 1.2, the last two persistent diagrams at 7h50 and 8h are close. All this suggests that the bow echo acquires the maturation stage from 7h40. So the *confirmed risk alert* will therefore be triggered at 7h40 with an imminent risk producing its effects, likely to harm their physical integrity.

5.2 Limitations

One of the limitations of this example is obviously the few reflexivity radar images in order to follow the gradual evolution of the formation of the arc. This would allow to refine the decision to trigger the confirmed risk alert and reduce false alerts. The other difficulty in the example of the bow echo of August 18, 2022 is its formation very close to the Corsican coast, we see that the decision of imminent danger could only be taken very shortly before its arrival on the island. Case studies on other bow echoes are needed to ensure the suitability and accuracy of the persistent homology method to detect the bow echo.

6 Conclusion and perspectives

In this paper we present a method for analysing the evolution of a bow echo using TDA. In addition to the measurement of the arc curvature, the one-dimensional persistent homology allowed to characterize the formation and maturation of the bow echo shape. The results showed a robust implementation that allowed to identify changes in the evolution of this severe storm. By working on reflexivity radar images, this method of measurement and homological persistence made it possible to generate an alert.

Now regarding the prospects and future work, there are multiple options. Topological data analysis is a promising and little-explored new tool in meteorology. The first thing is to improve our model by including other parameters. The azimuthal shear values ââfrom 0 km to 3 km and from 0 km to 6 km must be taken into account as well as the humidity rate and the water temperature. All this could allow to improve the model and launch the alert in a better period. An extrapolation between two times of radar images would be necessary thanks to AI in order to follow the evolution of the arc as closely as possible. Topological data analysis in this field is still in its infancy, there is still a huge field to explore. Due to its ability to exploit large amounts of data, affected by missing values ââeven with the presence of noise, TDA has a strong potential that is inexpensive in terms of computation. Coupled with machine learning and in particular with neural networks, it will be capable of identifying relevant characteristics, helping with classification and prediction and thus opening up the field of possibilities.

References

[1] Bauer. U, Ripser: efficient computation of Vietoris-Rips persistence barcodes, J. Appl. Comput. Topol., Journal of Applied and Computational

Topology,2021, Vol 5, p 391–423, DOI: https://doi.org/10.1007/s41468-021-00071-5.

[2] https://geometrica.saclay.inria.fr/team/Fred.Chazal/

[3] Cohen-Steiner, D. and Edelsbrunner, H. and Harer, J.L. Stability of persistence diagrams. Discrete Computational Geometry 37(2007), 103-120.

[4] Corfidi. Stephen. F, Coniglio. Michael. C. Cohen. Ariel. E, and Mead. Corey. M. A Proposed Revision to the Definition of "Derecho". Bulletin of the American Meteorological Society, 01 Jun 2016 DOI: https://doi.org/10.1175/BAMS-D-14-00254.1 Page(s): 935â949

[5] Davis. C, Atkins. N, Bartels. D, L. Bosart et al., "The bow echo and mcv experiment", Bulletin of the American Meteorological Society, vol. 85, no. 8, p. 1075, 2004.

[6] Edelsbrunner, H., Harer, J., Computational Topology: an Introduction, American Mathematical Society, 2010 .

[7] H. Edelsbrunner, D. Letscher and A. Zomorodian, Topological persistence and simplification, Discrete Comput. Geom. 28 (2002) 511-S533.

[8] González-Alemán, J. J., D. Insua-Costa, E. Bazile, S. GonzÃ¡lez-Herrero, M. Marcello Miglietta, P. Groenemeijer, and M. G. Donat, 2023: Anthropogenic Warming Had a Crucial Role in Triggering the Historic and Destructive Mediterranean Derecho in Summer 2022. ull. Amer. Meteor. Soc., 104, E1526âE1532, https://doi.org/10.1175/BAMS-D-23-0119.1.

[9] Johns, R. H., and W. D. Hirt, 1987: Derechos: Widespread convectively induced windstorms. Wea. Forecasting, 2, 32â49, doi:10.1175/1520-0434(1987)002/0032:DWCIW.2.0.CO.2.

[10] Fanlin. Z and Jinyi. H. Automatic Detection of Bow Echoes from Weather Radar Images 2018 13th World Congress on Intelligent Control and Automation (WCICA) IEEE

[11] Fujita. T,1978: Manual of downburst identification for project NIMROD. Satellite and Mesometeorology Research Paper 156, 104 pp.

[12] giotto-tda: A Topological Data Analysis Toolkit for Machine Learning and Data Exploration, Tauzin et al, J. Mach. 22.39 (2021): 1-6.

[13] Goulet, L., 2015: Bow echoes: Conceptual schemes and European relevance. The European Forecaster, No. 20, 22â31, http://www.euroforecaster.org/newsletter20/meteofr2.

[14] Houze. R. Jr,2004: Mesoscale convective systems Révérend Géophys. ,42, RG4003,https://doi.org/10.1029/2004/000150.

[15] Kamani, M., Farhat. F, Wistar. S, and Wang. J, 2016: Shape matching using skeleton context for automated bow echo detection. 2016 IEEE Int. Conf. on Big Data (Big Data), Washington, DC, IEEE, 901â908, https://doi.org/10.1109/BigData. 2016.7840685.

[16] Klimowsky B. A, Hjelmfelt. M. R, and Bunkers. M. J, "Radar observations of the early evolution of bow echoes", Weather and Forecasting, vol. 19, pp. 727â734, 2004

[17] Klimowski. B. A, Bunkers. M. J, Hjelmfelt. M. R, and Covert. J. N. , "Severe convective windstorms over the northern high plains of the United States", Weather and Forecasting, vol. 18, no. 3, pp. 502â519, 2003.

[18] Carbone. R, Lee. Wen Chau and Wakimoto. R. The evolution and structure of a "Bow echo microburst" event. Part II: the bow echo.(1992). Monthly weather review. Volume 120.

[19] Maletic S, Zhao Y, Rajkovic M (2016) Persistent topological features of dynamical systems. Chaos. https:// doi.org/10.1063/1.49494 72.

[20] https://www.meteociel.fr. La météo en temps réel et prévisions météo pour la France, Observations météo, modèles numériques et logiciels météo (GFS, ECMWF, UKMO, GEM, AROME, ARPEGE, ...

[21] https://meteofrance.com. PREVISIONS METEO FRANCE - Site Officiel de Météo-France

[22] Mounier, A., L. Raynaud, L. Rottner, M. Plu, P. Arbogast, M. Kreitz, L. Mignan et B. TouzÃ©, 2022 : DÃ©tection des échos d'arc dans les prévisions à l'échelle kilométrique à l'aide d'un réseau de neurones convolutifs . Artif. Intell. Système de terre. 1, e210010, https://doi.org/10.1175/AIES-D-21-0010.1.

[23] Muszynski G, Kashinath K, Kurlin V, Wehner M (2019) Topological data analysis and machine learning for recognizing atmospheric river patterns in large climate datasets. Geosci Model Dev. https://doi. org/ 10. 5194/ gmd- 12-613- 2019

[24] Patil, V., S. Das, and A. Phadke, 2019: Methods for mesoscale convective systems detection and tracking: A survey. 2019 10th Int. Conf. on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, IEEE, 1â7, https://doi.org/10.1109/ICCCNT45670.2019.8944656.

[25] Przybylinski, R., 1995: The bow echo: Observations, numerical simulations, and severe weather detection methods. Wea. Forecasting, 10, 203â218, https://doi.org/10.1175/1520-0434(1995)010,0203:TBEONS.2.0.CO; 2.

[26] Pucik, T., 2022: The derecho and hailstorms of 18 August 2022. European Severe Storms Laboratory, accessed 11 May 2022, https://www.essl.org/cms/thederecho-and-hailstorms-of-18-august-2022.

[27] Strommen, K., Chantry, M., Dorrington, J. et al. A topological perspective on weather regimes. Clim Dyn 60, 1415â1445 (2023). https://doi.org/10.1007/s00382-022-06395.

[28] Tymochko S, Munch E, Dunion J, Corbosiero K, Torn R (2020) Using persistent homology to quantify a diurnal cycle in hurricanes. Pattern Recognit Lett 133:137â143. https:// doi. org/ 10. 1016.

[29] Zheng, Y. G., Chen, J. (2013). A climatology of deep convection over South China and the adjacent waters during summer. Journal of Tropical Meteorology, 19, 1â15.

[30] Zhou, A., Zhao, K., Xu, X., Liu, Q., Ding, Z., Huang, H., et al. (2023). A climatological study on the two types of bow echoes over South China. Journal of Geophysical Research: Atmospheres, 128, e2023JD038720. https://doi.org/10.1029/2023JD038720