AIMS Mathematics, 11(6): 18122–18147.
DOI: 10.3934/math.2026737
Received: 07 April 2026
Revised: 04 June 2026
Accepted: 10 June 2026
Published: 18 June 2026

https://www.aimspress.com/journal/Math

Research article

Topological characterization of ecological dynamics via persistent homology

Merve Kahraman Ariman*

Department of Mathematics, Faculty of Arts and Sciences, Izmir University of Economics, Izmir, Turkey

* **Correspondence:** Email: merve.ariman@ieu.edu.tr; Tel: +905314910728.

Abstract: Conventional linear methods provide valuable insights into ecological population dynamics but may not fully capture their underlying geometric complexity. We present a large-scale topological characterization of ecological population dynamics using persistent homology applied to 500 time series from the BioTIME database, spanning marine, terrestrial, and freshwater ecosystems. Time-delay embedding and Vietoris–Rips filtration yield two classes of topological invariants: Betti numbers β_k, which count persistent topological features, and persistence entropy H_k, which quantifies their distributional complexity. These invariants quantify multiscale cyclic organization in a manner that complements spectral and autoregressive approaches. Three principal components capture 92.6% of topological variance, revealing that ecological attractor geometry is fundamentally low-dimensional. Realm membership explains less than 0.01% of this variance, demonstrating that habitat type imposes negligible constraints on dynamical complexity relative to within-realm heterogeneity, a finding that challenges the widely assumed structuring role of environmental context. An exceptionally strong coupling ($\rho = 0.989$) between β_k and H_k reflects an information-theoretic bound $H_k \leq \log_2(\beta_k)$. These results support shared dynamical mechanisms, including density dependence, predator-prey interactions, and life history trade-offs, as primary determinants of attractor topology, and they establish persistent homology as a noise-robust complement to conventional methods for comparative ecological analysis.

Keywords: persistent homology; topological data analysis; ecological time series; population dynamics; attractor geometry; phase space reconstruction; BioTIME; dynamical complexity
Mathematics Subject Classification: 55N31, 62R40, 94A17, 92D40, 37N25

1. Introduction

Ecological systems often exhibit rich, nonlinear temporal dynamics arising from interactions among species, environmental variability, density-dependent feedbacks, and stochastic disturbances. The

recognition that ecological dynamics cannot be fully understood through linear or equilibrium-based models has reshaped theoretical and empirical ecology over the past five decades. A key turning point occurred with May's seminal demonstration that even simple, one-dimensional population models are capable of producing bifurcations, period-doubling cascades, and chaotic trajectories under deterministic conditions [1]. This finding fundamentally altered how ecologists interpret temporal fluctuations, revealing that irregular population patterns could arise from intrinsic nonlinearities rather than solely from environmental noise. Subsequent theoretical and empirical studies strengthened this perspective, showing that nonlinear processes can generate oscillations, multistability, and sensitive dependence on initial conditions across a wide variety of ecological systems [2, 3].

Understanding such complexity was further propelled by developments in dynamical systems theory, including the discovery of deterministic chaos and strange attractors [4] and the introduction of delay-coordinate embedding for attractor reconstruction from scalar time series [5]. Ellner and Turchin [6] demonstrated that nonlinear dynamics can be detected in real population time series, providing evidence that ecological complexity often arises from deterministic density-dependent processes. Over the following decades, ecological forecasting and model-free inference techniques became increasingly influenced by dynamical systems theory. Turchin's synthesis of population dynamics [7] and the emergence of empirical dynamic modeling, particularly convergent cross mapping (CCM), further emphasized the role of nonlinear interactions and feedbacks in shaping ecological trajectories [8, 9]. These approaches rely on the idea that attractor reconstruction provides a powerful lens through which causality, predictability, and system stability can be analyzed.

Despite the impact of nonlinear approaches, many commonly used statistical techniques in ecology still rely on linear assumptions. Autoregressive models, spectral analysis, and linear dimension-reduction techniques provide valuable insights but may not fully characterize the geometric complexity of ecological processes [10]. In practice, estimating nonlinear invariants such as Lyapunov exponents can be challenging for ecological datasets, which typically suffer from noise, short length, irregular sampling, and missing observations. Although Rosenstein et al. [11] and Kantz [12] proposed robust algorithms for estimating Lyapunov exponents from short or noisy time series, these methods remain sensitive to data limitations and parameter choices. Thus, many classical tools for nonlinear time series analysis can be unreliable for ecological applications.

A major breakthrough was achieved with Takens' embedding theorem [13], which formalized the mathematical conditions under which a smooth dynamical system's attractor can be reconstructed from a single observable using delay coordinates. Sauer, Yorke, and Casdagli [14] extended this framework to systems with fractal attractors, establishing conditions involving a box-counting dimension. These results provide a rigorous foundation for attractor reconstruction and empower ecologists to analyze the geometry of dynamical systems even when only a single time series is available.

Delay embedding transforms a univariate time series into a higher-dimensional point cloud that approximates the system's underlying attractor. Whereas many traditional nonlinear analyses summarize this cloud using scalar metrics, a growing body of work emphasizes the importance of characterizing the global geometric and topological structure of reconstructed attractors. Topological data analysis (TDA) offers a powerful toolkit for achieving these goals. TDA captures mesoscale and global geometric features, such as connected components, loops, and voids, that reflect organization in the underlying dynamics. The introduction of persistent homology [15, 16] made it possible to quantify how these features appear and disappear across scales, yielding robust summaries called persistence

diagrams. Importantly, persistent homology is robust under data perturbations, as shown by Cohen-Steiner, Edelsbrunner, and Harer [17], making it particularly suitable for ecological datasets that often include measurement noise.

In recent years, TDA has emerged as a promising framework for analyzing time series. Bubenik's persistence landscapes [18] transformed persistence diagrams into functional representations, enabling direct statistical inference. Perea and Harer [19] demonstrated that sliding-window embeddings combined with persistent homology can detect periodicity and quasiperiodicity in noisy signals with remarkable sensitivity. Khasawneh and Munch [20] showed that persistent homology can differentiate between periodic, quasiperiodic, and chaotic regimes in dynamical systems, highlighting the potential of TDA for characterizing nonlinear behavior beyond the reach of traditional indicators. More recently, persistent homology has been applied to detect bifurcations in dynamical systems through Betti numbers and CROCKER plots [42] to analyze temporal networks via zigzag persistence [43] and to characterize the topology of nonlinear time series through phase space reconstruction [44], showing that TDA remains an active and growing framework for dynamical systems analysis. Closely related to the present work, Strommen et al. [37] applied persistent homology to characterize atmospheric weather regimes, equating the existence of dynamical regimes with nontrivial topological structure (connected components and loops) in the reconstructed attractor, including oscillatory modes such as the North Atlantic Oscillation.

Beyond dynamical systems, TDA has been successfully applied to a variety of scientific fields including time-series classification [21], materials science [22], and neuroscience [23]. These applications demonstrate that topological signatures can capture structural organization in complex, high-dimensional data that is not readily accessible through classical statistical approaches. Despite these methodological advancements, a fundamental question remains largely unexplored: Do different environmental realms impose fundamental constraints on the dynamical complexity of population fluctuations? Marine, terrestrial, and freshwater ecosystems differ profoundly in their physical properties, thermal stability, nutrient dynamics, and trophic structure. Marine environments exhibit greater thermal buffering and more stable abiotic conditions compared to freshwater and terrestrial systems, potentially reducing the number of concurrent environmental periodicities that drive population oscillations. Conversely, freshwater systems experience pronounced seasonal temperature variation, hydrological cycles, and nutrient pulses that may generate more complex multiscale dynamics. If habitat type constrains attractor geometry, we would expect systematic differences in topological complexity across realms. Alternatively, if common mechanisms such as density dependence, predator prey interactions, and life history trade-offs dominate population dynamics regardless of environmental context, topological structure should vary primarily with species-specific characteristics rather than broad habitat categories. Although spatial topology approaches have been applied to characterize ecological network connectivity and quality across landscapes [35], the present study employs a fundamentally different framework: persistent homology applied to temporal attractor reconstruction, targeting the dynamical geometry of population trajectories rather than the spatial network structure.

We address three interrelated questions that bridge dynamical systems theory and ecological understanding. First, can persistent homology capture ecologically meaningful variation in population dynamics across diverse systems? Second, do marine, terrestrial, and freshwater ecosystems exhibit systematically different topological signatures, or does within-realm heterogeneity dwarf between-

realm differences? Third, what geometric features of attractor structure can persistent homology reveal that complement conventional spectral and autoregressive methods? By applying topological data analysis to hundreds of time series spanning major environmental realms, we can assess whether habitat type imposes detectable constraints on dynamical complexity or whether population dynamics are governed primarily by taxon-specific processes operating across environmental boundaries.

Given the complexity, noise, and limited length of many ecological time series, TDA offers a compelling alternative perspective for understanding ecological dynamics.

Alongside these developments, ecological research has increasingly sought reliable early-warning indicators of regime shifts and tipping points [24, 25], motivating the search for complementary analytical frameworks that can characterize the geometric structure of population dynamics.

Despite these methodological advancements, there remains a lack of large-scale comparative studies that examine the topological structure of ecological time series across diverse ecosystems. Most previous applications of TDA to time series have focused on controlled experiments or individual ecological systems. Yet, biodiversity monitoring programs have generated extensive ecological time series covering multiple environmental realms. The BioTIME database [26] contains hundreds of population and abundance time series from marine, freshwater, and terrestrial ecosystems worldwide. Although this dataset has been widely used to study biodiversity change, turnover, and compositional dynamics, its potential for understanding dynamical geometry remains largely unexplored.

In this study, we integrate delay-coordinate embedding with persistent homology to quantify the topological complexity of ecological time series from three major environmental realms: marine, freshwater, and terrestrial ecosystems. Using 500 time series from BioTIME, we reconstruct attractors for each system, compute persistent homology across multiple embedding dimensions and delays, and extract a suite of geometric and topological features, including Betti numbers, persistence entropy, maximum persistence, and dimensionality reduction of the topological feature space. This framework enables us to address whether marine, terrestrial, and freshwater habitats impose fundamental constraints on population dynamics, or whether common mechanisms generate similar complexity across realms. More broadly, we evaluate what geometric information conventional linear time series methods are not designed to capture and demonstrate how topological invariants provide complementary insights into the organization of ecological attractors.

The contributions of this work are threefold. First, we establish a rigorous mathematical pipeline combining Takens' embedding theorem [13] with Vietoris–Rips persistent homology [15] for large-scale comparative characterization of attractor topology, validated through sensitivity analysis across embedding dimensions $m \in \{2, 3, 4, 5\}$ and series length thresholds. Second, we provide empirical confirmation of the information-theoretic bound $H_k \leq \log_2(\beta_k)$ and demonstrate that the six-dimensional topological feature space organizes along a low-dimensional manifold, results with implications for topological feature selection in dynamical systems analysis. Third, we assess whether ecological system type, taxonomic identity, or environmental realm explain variation in topological complexity, revealing that within-realm heterogeneity vastly exceeds between-realm differentiation. By integrating nonlinear dynamics, empirical dynamic modeling, and TDA, this study provides new insights into the geometry of ecological time series and the mechanisms underlying ecological variability.

The remainder of this paper is organized as follows. Section 2 describes the datasets, embedding procedures, and persistent homology computations. Section 3 presents the results, including

comparisons across realms. Section 4 discusses ecological interpretations and methodological implications. Section 5 concludes with key findings and future directions.

2. Materials and methods

2.1. Data acquisition and preprocessing

Ecological time series data were obtained from BioTIME version 2 (accessed November 2025) [26], a comprehensive global repository of biodiversity time series. The complete database yielded approximately 24,000 study–species combinations after initial parsing. For each study species combination, paired arrays of temporal coordinates t_i and abundance values a_i were constructed, where $i \in \{1, 2, \ldots, n\}$ indexes sequential observations. When multiple species occurred within the same study, each species was retained as a separate time series without aggregation. Although time series from the same study may share environmental context, Kruskal–Wallis tests confirmed that study identity explains less than 0.1% of topological variance ($\epsilon^2 < 0.002$ for all features), indicating that species-specific processes dominate over study-level clustering effects.

A filtering scheme was applied that removed non-numeric entries, infinite values, and non-positive abundances (including zero values). Two constraints were imposed on extracted time series: (1) a minimum of $n \geq 25$ observations to ensure sufficient sampling for phase space reconstruction [13, 14] and (2) an abundance threshold requiring $\sum_{i=1}^{n} a_i \geq 20$ to exclude ecologically marginal populations. Constant time series were filtered out by requiring nonzero temporal variance, $\sigma^2(a) > 0$. Additionally, time series with extreme temporal irregularity (coefficient of variation $R > 2$ for interobservation intervals) were flagged for sensitivity analysis.

The threshold of $n \geq 25$ was selected based on the following considerations. First, Takens' embedding theorem requires $m \geq 2d + 1$ dimensions, where d is the attractor dimension [13, 14]. With $m = 3$, the embedding is theoretically sufficient for attractors of dimension $d \leq 1$ (fixed points and limit cycles), which represent the dominant dynamical regime in ecological time series [1, 7]. For such low-dimensional attractors, the minimum number of reconstructed state vectors required to capture basic topological structure is substantially lower than what is needed for quantitative invariants such as Lyapunov exponents or fractal dimensions [11, 12]. Second, unlike these quantitative measures, the topological invariants computed here, Betti numbers and persistence entropy, characterize the qualitative geometry of the reconstructed point cloud, and the stability theorem [17] guarantees their robustness to noise and sparse sampling. We acknowledge, however, that shorter series may undersample more complex attractors with higher β_1 values, meaning that our results likely provide a conservative estimate of topological complexity. The dataset spans a wide range of lengths: 33.6% of series have $n < 50$, 41.0% have $50 \leq n \leq 200$, and 25.4% have $n > 200$. Across all 500 series, the mean length is 255.2, and the median is 79.5 observations, reflecting a right-skewed distribution driven by a small number of very long series. To verify that the $n \geq 25$ threshold does not compromise the main findings, all analyses were repeated on the subset of 332 series with $n \geq 50$. Results were nearly identical: Three principal components captured 92.5% of topological variance (versus 92.6% for the full dataset), the Spearman correlation between β_1 and H_1 remained $\rho = 0.987$ (versus 0.989), and realm effect sizes remained uniformly negligible ($\epsilon^2 < 0.002$ for all features), confirming that the main conclusions are robust to the choice of minimum series length.

After applying all filtering criteria, 500 valid time series were retained from 212 distinct studies,

spanning marine ($n = 294$, 58.8%), terrestrial ($n = 146$, 29.2%), and freshwater ($n = 60$, 12.0%) ecosystems.

2.2. Time delay embedding and phase space reconstruction

Takens' embedding theorem [13] was employed to reconstruct phase space topology from univariate abundance measurements. For a time series $\{a_1, a_2, \ldots, a_n\}$, the embedding procedure constructs delay vectors in $\mathbb{R}^m$:

$$\mathbf{x}_i = (a_i, a_{i+\tau}, a_{i+2\tau}, \ldots, a_{i+(m-1)\tau}), \quad i \in \{1, 2, \ldots, N\}, \tag{2.1}$$

where m denotes the embedding dimension, τ represents the time delay parameter, and $N = n-(m-1)\tau$ gives the number of reconstructed state vectors [5].

Takens' theorem establishes that for generic observation functions and $m \geq 2d + 1$, where d is the attractor dimension, the embedding preserves topological structure including homology groups and Betti numbers [14].

Prior to embedding, each time series was standardized to unit variance:

$$\tilde{a}_i = \frac{a_i - \bar{a}}{\sigma(a) + \epsilon}, \tag{2.2}$$

where $\epsilon = 10^{-10}$ prevents division by zero. This normalization ensures comparable geometric scales across ecological systems.

The embedding dimension $m = 3$ was selected throughout, providing sufficient dimensionality for attractors of dimension $d \leq 1$ (fixed points and limit cycles) while maintaining computational tractability [1, 7]. The time delay was determined adaptively using $\tau = \max(1, \lfloor n/20 \rfloor)$, balancing decorrelation and data utilization [12].

To evaluate the sensitivity of results to the choice of embedding dimension, all analyses were repeated for $m \in \{2, 4, 5\}$ (Table 1). Results were consistent across all three values: The Spearman correlation between β_1 and H_1 remained $\rho \geq 0.977$, realm effect sizes remained uniformly negligible ($\epsilon^2 < 4 \times 10^{-5}$), and three principal components captured at least 89% of topological variance, confirming that the main conclusions are robust to the choice of embedding dimension.

Table 1. Sensitivity of main findings to embedding dimension m, based on $n = 500$ ecological time series. Realm effect size ϵ^2 quantifies the proportion of variance in β_1 explained by realm membership.

Metric	$m = 2$	$m = 3$	$m = 4$	$m = 5$
Series analyzed	500	500	500	500
PC1 variance explained (%)	58.9	64.5	66.9	67.6
PC1+PC2+PC3 variance explained (%)	89.2	92.6	92.8	93.4
Spearman ρ (β_1–H_1)	0.977	0.989	0.991	0.990
Realm effect size ϵ^2 (β_1)	3.75×10^{-5}	3.82×10^{-5}	3.68×10^{-5}	3.40×10^{-5}

2.3. Persistent homology

Persistent homology extracts multiscale topological features from point cloud data [15, 27]. Given an embedded point cloud $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset \mathbb{R}^m$, the Vietoris–Rips complex $\mathrm{VR}(X, \epsilon)$ was constructed

for scale parameter $\epsilon \geq 0$, consisting of simplices whose vertices have pairwise distances at most ϵ. As ϵ increases, a nested filtration of simplicial complexes is obtained.

For each filtration value, the kth homology group H_k captures k-dimensional topological features: Connected components ($k = 0$) and loops ($k = 1$). Each feature is characterized by birth scale b and death scale d, yielding persistence pairs (b, d) with persistence $\ell = d - b$. The collection of persistence pairs forms the persistence diagram $\mathrm{Dgm}_k(X)$.

Persistence diagrams were computed using Ripser [28]. Point clouds exceeding $N > 1000$ points were randomly subsampled, and a maximum filtration scale of $\epsilon_{\max} = 2.0$ was imposed on normalized point clouds. The stability theorem [17] guarantees robustness to noise:

$$W_\infty(\mathrm{Dgm}_k(X), \mathrm{Dgm}_k(Y)) \leq d_{\mathrm{GH}}(X, Y), \tag{2.3}$$

where W_∞ is the bottleneck distance and d_{GH} the Gromov–Hausdorff distance.

2.4. Topological feature extraction

The kth Betti number β_k quantifies the number of k-dimensional topological features:

$$\beta_k = |\{(b_i, d_i) \in \mathrm{Dgm}_k(X) : d_i < \infty\}|. \tag{2.4}$$

The zeroth Betti number β_0 represents connected components, and β_1 enumerates one-dimensional loops corresponding to periodic or quasiperiodic behavior [19, 20].

From each persistence diagram, summary statistics were extracted:

$$\begin{aligned} \text{Total persistence:} \quad & L_k = \sum_{i=1}^{n_k} (d_i - b_i), \\ \text{Maximum persistence:} \quad & \ell_k^{\max} = \max_i (d_i - b_i), \\ \text{Mean persistence:} \quad & \bar{\ell}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} (d_i - b_i). \end{aligned} \tag{2.5}$$

2.5. Persistence entropy

To quantify distributional complexity, the normalized persistence vector was defined as $p_i = \ell_i / L_k$, forming a probability distribution over topological features. The Shannon entropy was computed as

$$H_k = -\sum_{i=1}^{n_k} p_i \log_2 p_i. \tag{2.6}$$

Low entropy indicates dominance by few features (simple dynamics), whereas high entropy indicates uniform distribution across features (complex, multiscale structure) [2, 6]. The entropy satisfies $0 \leq H_k \leq \log_2 n_k$, complementing Betti numbers by incorporating relative feature importance.

2.6. Statistical analysis

For inter-realm comparisons, the Kruskal–Wallis H test was employed:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{g} \frac{R_i^2}{n_i} - 3(N+1), \tag{2.7}$$

where N is total sample size, n_i is group i sample size, and R_i is the rank sum for group i. Post-hoc pairwise comparisons used the Dunn test with Bonferroni correction.

Effect sizes were quantified using epsilon-squared: $\epsilon^2 = H/(N^2 - 1)$, interpreted as small (≈ 0.01), medium (≈ 0.06), or large (≈ 0.14).

Both Pearson and Spearman correlation coefficients were computed to quantify associations between topological features. Multiple testing corrections used the Benjamini–Hochberg procedure to control the false discovery rate at $q = 0.05$.

To assess whether topological features are confounded by the time series length, Spearman correlations between series length and each topological feature were computed. Strong positive correlations were observed (Spearman $\rho = 0.684$–0.746 for Betti numbers and entropy measures, $\rho = 0.345$ for maximum persistence). To control for this effect, length-adjusted residuals were obtained by regressing each topological feature on time series length using ordinary least squares, and Kruskal–Wallis tests were repeated on these residuals. Results are reported both before and after the length adjustment in Table 4.

2.7. Dimensionality reduction

Principal component analysis (PCA) was applied to the standardized feature matrix. For feature matrix $\mathbf{X} \in \mathbb{R}^{n\times p}$, PCA identifies orthogonal directions of maximal variance through eigendecomposition of the covariance matrix. The proportion of variance explained by the first q components is

$$R_q^2 = \frac{\sum_{j=1}^{q} \lambda_j}{\sum_{j=1}^{p} \lambda_j}. \tag{2.8}$$

Nonlinear dimensionality reduction was performed using UMAP [29] with $k = 15$ nearest neighbors and t-SNE [30] with perplexity 30, following established conventions for manifold learning.

2.8. Computational implementation

All analyses were implemented in Python 3.9 using Ripser (v0.6.1) for persistent homology [28], scikit-learn (v0.24.2) for dimensionality reduction, and SciPy (v1.7.0) for statistical tests. Random seeds were fixed for reproducibility. Total computation time was approximately two hours for 500 time series on a workstation with an Intel Core i7 processor and 32 GB RAM.

3. Results

3.1. Dataset characteristics

The topological analysis was conducted on $n = 500$ ecological time series from the BioTIME database, spanning marine ($n = 294$, 58.8%), terrestrial ($n = 146$, 29.2%), and freshwater ($n = 60$, 12.0%) ecosystems. Taxonomic composition was dominated by invertebrates (53.8%), followed by fish (16.6%), birds (12.2%), and plants (9.0%). The climatic distribution showed a temperate zone bias (66.0%).

Time series lengths ranged from 25 to 2000 observations across all 500 series (mean = 255.2, median = 79.5), indicating a right-skewed distribution. Phase space reconstruction used embedding dimension $m = 3$ with adaptive time delay $\tau = \max(1, \lfloor n/20 \rfloor)$; see Figure 1.

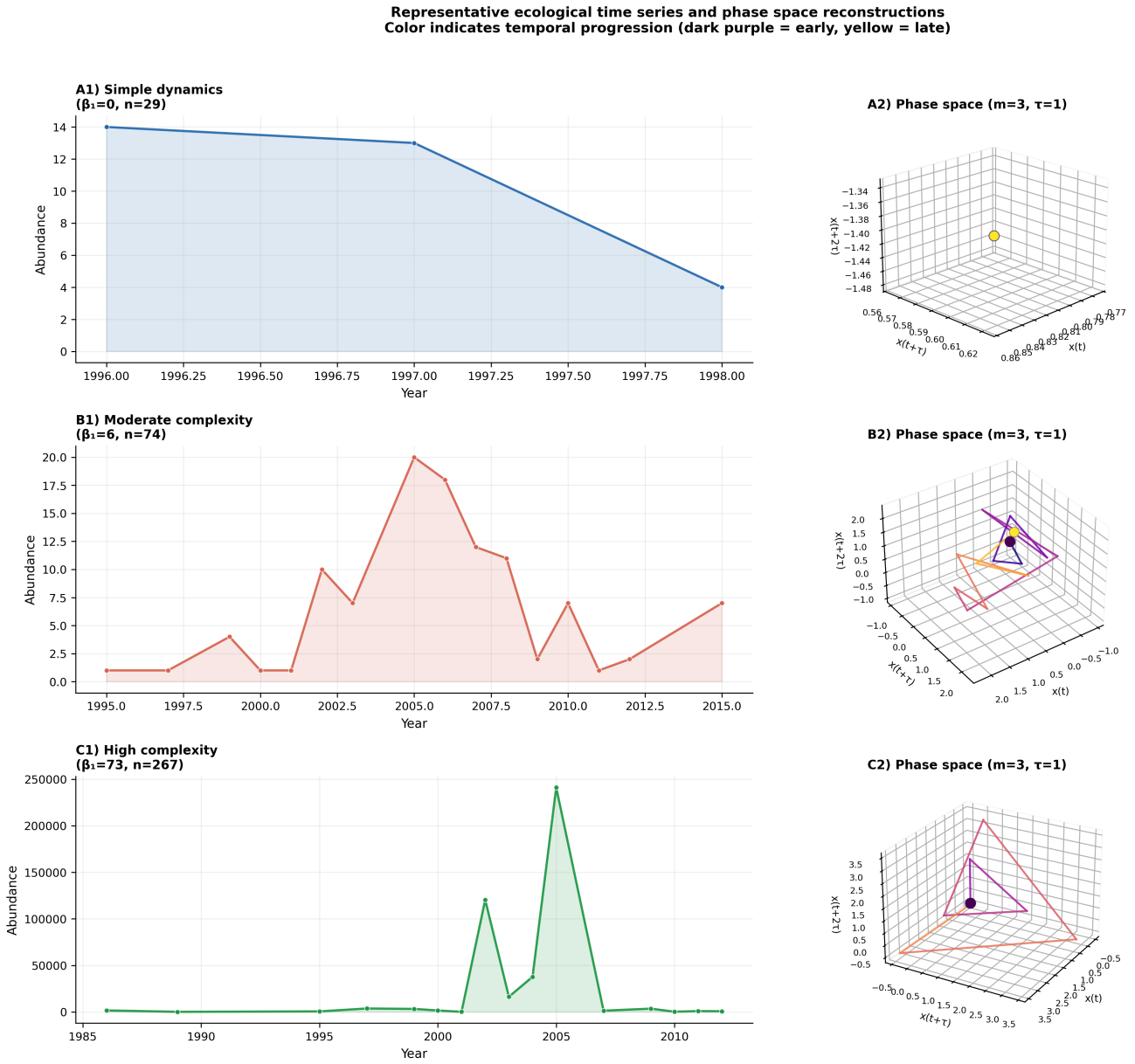

Figure 1. Representative ecological time series (left panels) and their associated three-dimensional phase space reconstructions via time delay embedding (right panels). Each time series is embedded using delay vectors $\mathbf{x}_i = (a_i, a_{i+\tau}, a_{i+2\tau})$ with $m = 3$ and adaptive time delay τ; color indicates temporal progression (dark purple = early, yellow = late). The number of reconstructed state vectors displayed in each phase space plot is $N = n - (m-1)\tau$, which is necessarily smaller than the original series length n: For the three examples shown, $N = 27$ ($n = 29$), $N = 70$ ($n = 74$), and $N = 257$ ($n = 267$), respectively. This reduction is an inherent consequence of the delay embedding procedure and does not reflect data loss. (A) Simple dynamics ($\beta_1 = 0$, $n = 29$): The time series exhibits low variability, and the reconstructed attractor forms a compact, structureless cluster in phase space, indicating the absence of persistent cyclic organization. (B) Moderate complexity ($\beta_1 = 6$, $n = 74$): Recurring population fluctuations produce a discernible loop structure in the reconstructed attractor, corresponding to a small number of persistent topological cycles. (C) High complexity ($\beta_1 = 73$, $n = 267$): Multiscale oscillations generate a dispersed, intricate attractor geometry spanning a large region of phase space, reflecting the presence of many concurrent cyclic processes at different timescales. The first Betti number β_1 quantifies this geometric complexity across all 500 time series analyzed in this study.

Comprehensive descriptive statistics for topological features are presented in Table 2. The zeroth Betti numbers β_0 exhibited mean 118.5 (SD = 201.8) with median 42.0, indicating positive skewness. First Betti numbers β_1 displayed mean 40.8 (SD = 86.4) with median 7.0, and 22.4% of time series exhibited $\beta_1 = 0$. Persistence entropy averaged $H_0 = 3.33$ bits and $H_1 = 1.86$ bits.

Table 2. Descriptive statistics for topological features extracted from $n = 500$ ecological time series.

Feature	Count	Mean	Std	Min	Q_1	Median	Q_3	Max
Dimension 0 (Connected components)								
β_0	500	118.5	201.8	0.0	20.0	42.0	103.0	997.0
Total persistence L_0	500	31.5	35.7	0.0	10.1	20.0	36.1	214.2
Max persistence $\ell_0^{\max}$	500	1.40	0.60	0.0	1.18	1.61	1.85	2.00
Entropy H_0 (bits)	500	3.33	1.51	0.0	2.75	3.31	4.21	6.54
Dimension 1 (Loops/Cycles)								
β_1	500	40.8	86.4	0.0	1.0	7.0	26.0	511.0
Total persistence L_1	500	3.03	7.64	0.0	0.030	0.511	2.47	87.4
Max persistence $\ell_1^{\max}$	500	0.196	0.181	0.0	0.017	0.159	0.320	0.954
Entropy H_1 (bits)	500	1.86	1.64	0.0	0.0	1.68	2.96	6.04

3.2. Distribution of topological invariants

The distributions of topological invariants are presented in Figure 2. Both Betti numbers exhibited pronounced positive skewness ($\gamma_1 = 2.93$ for β_0; $\gamma_1 = 2.99$ for β_1) with heavy tails. The distribution of β_0 showed modal concentration in $[10, 60]$, whereas β_1 was strongly concentrated near zero, with 81.4% of values below 50.

Persistence entropy H_0 displayed approximately symmetric distribution (skewness $= -0.43$), whereas H_1 showed bimodal structure, with 26.4% of observations at $H_1 = 0$. A remarkably strong monotonic relationship emerged between β_1 and H_1 (Spearman $\rho = 0.989$, Pearson $r = 0.766$), reflecting the information-theoretic constraint $H_1 \leq \log_2(\beta_1)$. The discrepancy between correlation coefficients indicates strongly monotonic but nonlinear association.

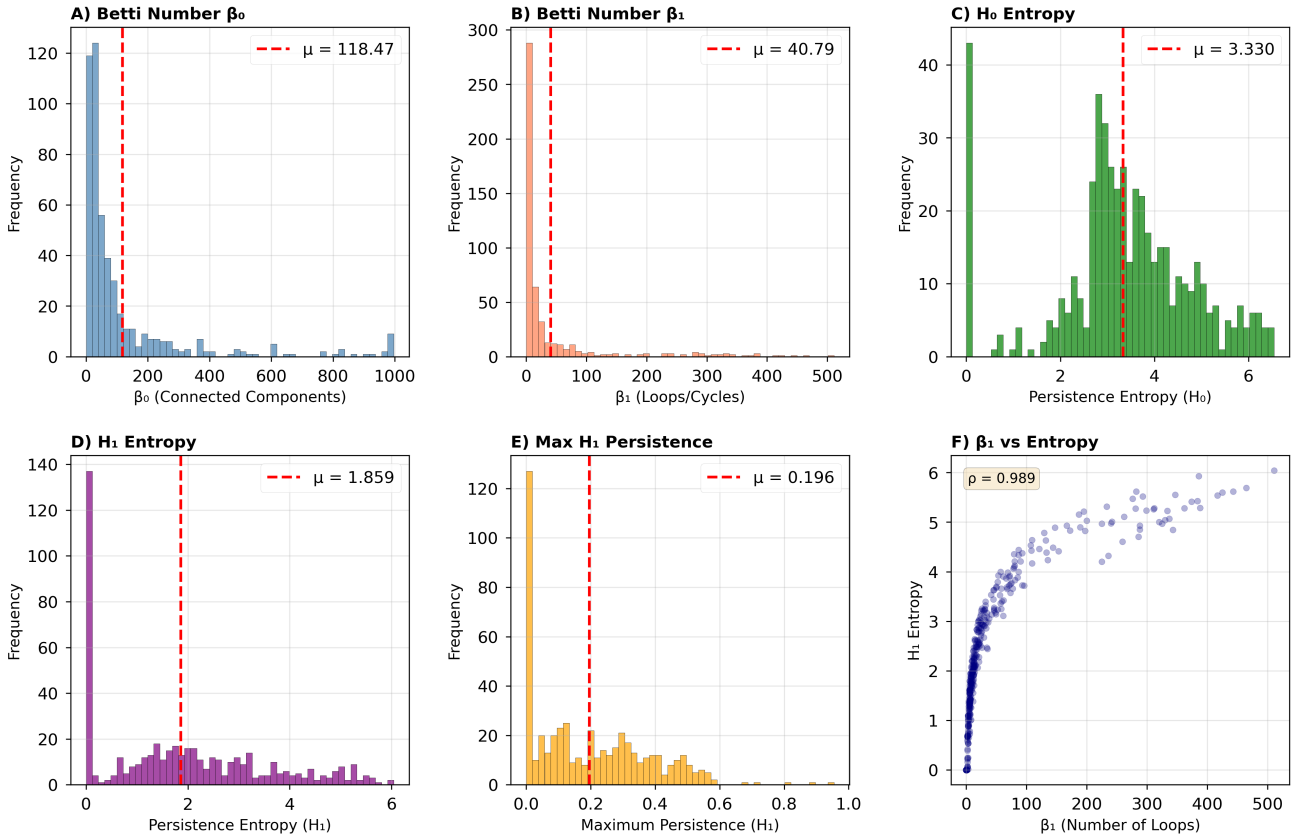

Figure 2. Distribution of topological invariants across $n = 500$ ecological time series. (A) Histogram of β_0 with mean $\mu = 118.47$. (B) Histogram of β_1 with mean $\mu = 40.79$. (C) Distribution of H_0 entropy with mean $\mu = 3.330$ bits. (D) Distribution of H_1 entropy showing bimodal structure with mean $\mu = 1.859$ bits. (E) Distribution of maximum H_1 persistence with mean $\mu = 0.196$. (F) Scatter plot of β_1 versus H_1 showing Spearman correlation $\rho = 0.989$.

3.3. Realm-specific patterns in topological complexity

The distributions of Betti numbers stratified by ecological realm are presented in Figure 3 and Table 3. Marine systems exhibited mean $\beta_0 = 99.6$ and $\beta_1 = 32.2$; terrestrial systems showed $\beta_0 = 128.8$ and $\beta_1 = 50.6$; and freshwater systems displayed the highest values, with $\beta_0 = 186.1$ and $\beta_1 = 59.4$.

The box plots in panels C and D of Figure 3 provide detailed visualizations of realm-specific distributions. For β_0 (panel C), the boxes representing interquartile ranges occupy similar vertical positions for marine and terrestrial systems but shift substantially upward for freshwater systems. The interquartile ranges span from approximately 19 to 90 for marine, 20 to 134 for terrestrial, and 27 to 214 for freshwater systems, reflecting progressively greater dispersion in the upper tail. All three realms exhibit numerous high-leverage outliers extending to values near $\beta_0 = 1000$, with the density of outliers increasing from marine to freshwater systems. The whiskers demonstrate that whereas median values differed modestly across realms, the extreme values were comparable (marine: 991, terrestrial: 989, freshwater: 997), indicating that all realms occasionally generate dynamics with extremely fragmented phase space structure.

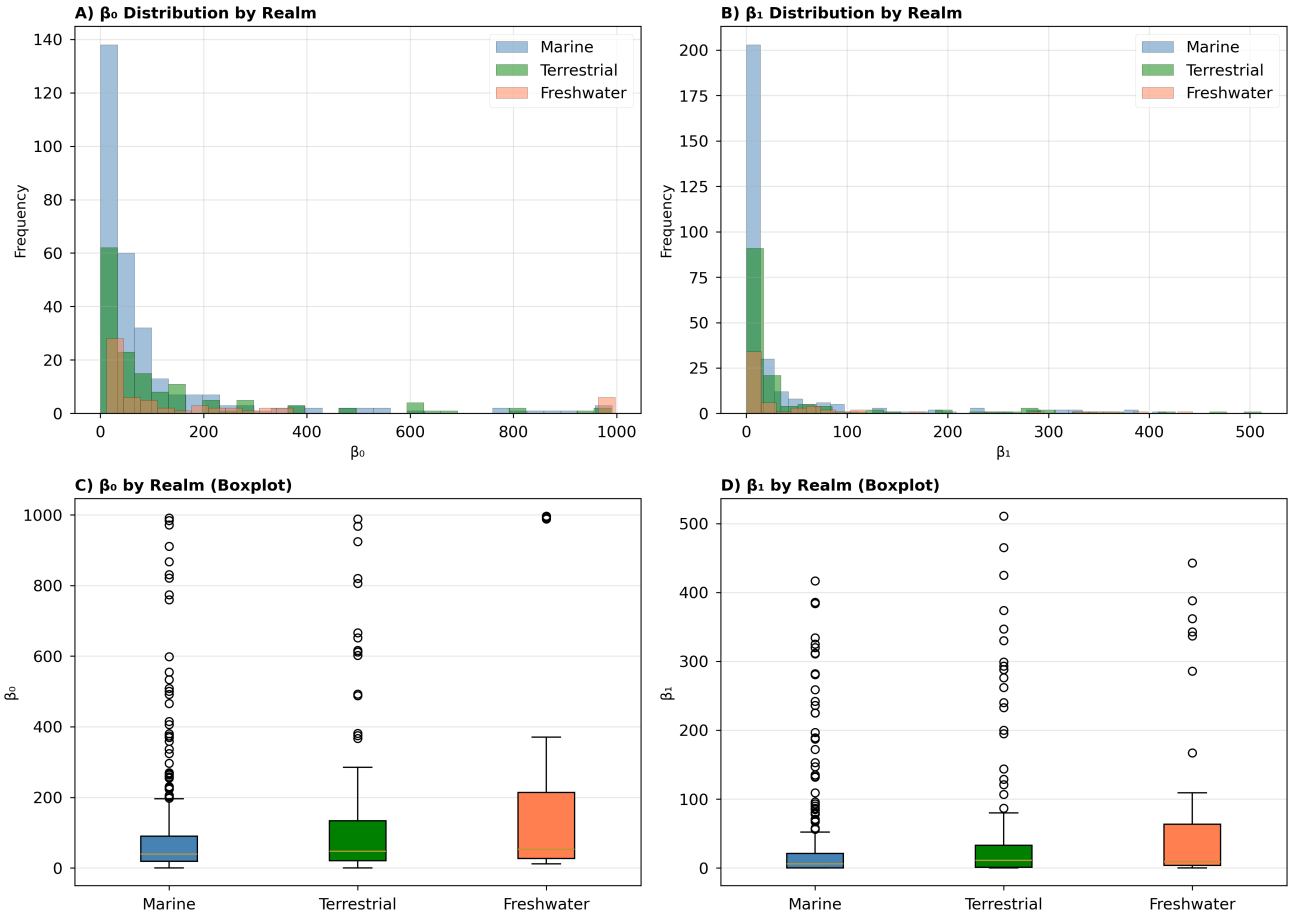

Figure 3. Distribution of Betti numbers across ecological realms. (A) Histograms of β_0 by realm. (B) Histograms of β_1 by realm. (C) Box plots of β_0 showing median values: marine 39.5, terrestrial 47.0, freshwater 52.5. (D) Box plots of β_1 showing median values: marine 6.0, terrestrial 11.0, freshwater 9.0. Kruskal–Wallis tests: β_0 ($H = 9.096$, $p = 0.0106$); β_1 ($H = 9.453$, $p = 0.00886$). High-leverage outliers ($\beta_0 > 500$, $\beta_1 > 200$) correspond to time series with $n \geq 552$ observations (median $n = 1{,}638$), spanning all three realms: freshwater species (BioTIME studies 559, 577, 831, 848), terrestrial species (studies 195, 301, 582, 596, 634, 702, 709), and marine species (studies 45, 119, 163, 176, 180, 189, 190, 196, 200, 213, 310, 374, 438, 595), consistent with the strong positive correlation between series length and Betti numbers (Spearman $\rho = 0.684$–0.746, Section 2.6).

For β_1 (panel D), the boxes occupied relatively narrow vertical ranges due to strong concentration near zero, with medians of 6.0 (marine), 11.0 (terrestrial), and 9.0 (freshwater). The freshwater box extends to substantially higher upper quartile values ($Q_3 = 63.5$) compared to those for marine ($Q_3 = 21.0$) and terrestrial ($Q_3 = 32.8$) realms, indicating broader distribution of moderate loop counts. The most visually striking feature is the extensive field of high-leverage outliers extending beyond $\beta_1 = 500$ for all three realms. The lower quartile approaches zero for marine systems ($Q_1 = 0.25$) but increases progressively for terrestrial ($Q_1 = 1.0$) and freshwater ($Q_1 = 3.75$), quantifying the decreasing prevalence of loop-free dynamics across realms.

Table 3. Betti numbers stratified by ecological realm.

Realm	Feature	Mean	Median	Std	Q_1	Q_3
Marine ($n = 294$)	β_0	99.6	39.5	174.5	19.0	90.0
	β_1	32.2	6.0	71.9	0.3	21.0
Terrestrial ($n = 146$)	β_0	128.8	47.0	205.3	20.3	133.5
	β_1	50.6	11.0	100.7	1.0	32.8
Freshwater ($n = 60$)	β_0	186.1	52.5	288.3	26.8	214.0
	β_1	59.4	9.0	107.9	3.8	63.5
Proportion with $\beta_1 = 0$*: Marine* 25.2%*, Terrestrial* 23.3%*, Freshwater* 6.7%						

Ecologically, higher β_1 in freshwater systems suggests more concurrent periodic processes operating simultaneously, potentially reflecting pronounced seasonal temperature variation in thermally unstratified habitats, multiple hydrological cycles (snowmelt, monsoons, baseflow), and pulsed nutrient dynamics that generate overlapping population oscillations at different timescales. Marine systems, by contrast, benefit from thermal buffering and more stable abiotic conditions that may reduce the diversity of environmental forcing frequencies. However, the negligible effect size ($\epsilon^2 < 0.0001$) indicates that this is a weak tendency rather than a fundamental ecological constraint, with individual marine systems spanning nearly the entire observed range of topological complexity.

Box plot comparisons of additional topological features reveal consistent patterns. For H_0 entropy, freshwater systems exhibit the highest median (3.392 bits) and widest interquartile range (IQR) (IQR = 2.098 bits), with boxes extending from approximately 2.8 to 4.9 bits. Marine systems display the lowest median (3.271 bits) and narrowest IQR (1.358 bits), concentrated between 2.7 and 4.0 bits. Terrestrial systems occupy the intermediate position (median = 3.435 bits, IQR = 1.642 bits). All three realms exhibit outliers at both extremes, with several systems showing H_0 entropy near zero (single dominant feature) and multiple systems exceeding 6 bits.

For H_1 entropy, a distinctive feature is the prominent lower quartile at exactly zero for both marine and terrestrial systems, reflecting substantial proportions of time series lacking persistent loops or possessing single dominant loops. Freshwater systems display a nonzero lower quartile ($Q_1 = 0.970$ bits), consistent with lower prevalence of loop-free dynamics. The boxes exhibit greater vertical extent compared to H_0, with IQR values ranging from 2.663 to 3.216 bits.

Maximum H_1 persistence display the strongest realm differentiation. Terrestrial systems exhibit substantially elevated values (median = 0.258, box from $Q_1 = 0.055$ to $Q_3 = 0.385$), whereas marine systems show markedly lower values (median = 0.115, box from $Q_1 = 0$ to $Q_3 = 0.291$). Notably, the lower quartile for marine systems is exactly zero, indicating that more than 25% of marine time series either lack persistent loops or possess loops with negligible persistence. High-leverage outliers extending beyond 0.8 appear for both marine and terrestrial systems.

Statistical hypothesis testing identified significant realm differences for all features prior to length adjustment (Table 4). After controlling for time series length, Betti number differences between realms remains significant, whereas entropy differences were attenuated to nonsignificance, indicating that length adjustment provides a more conservative and rigorous estimate of realm effects. Importantly, effect sizes remain uniformly small ($\epsilon^2 < 2 \times 10^{-4}$) across all features both before and after adjustment, confirming that realm membership explains a negligible proportion of topological variance regardless of length correction. This reinforces the central finding that within-realm heterogeneity vastly exceeds

between-realm differentiation.

Table 4. Statistical tests for realm comparisons, before and after controlling for time series length. Effect sizes: $\epsilon^2 = H/(N^2 - 1)$.

	Unadjusted			**Length-adjusted**		
Feature	H	p	ϵ^2	H	p	ϵ^2
β_0	9.096	0.0106	3.64×10^{-5}	29.410	4.1×10^{-7}	1.18×10^{-4}
β_1	9.453	0.00886	3.78×10^{-5}	26.553	1.7×10^{-6}	1.06×10^{-4}
H_0 entropy	6.774	0.0338	2.71×10^{-5}	4.333	0.115	1.73×10^{-5}
H_1 entropy	8.515	0.0142	3.41×10^{-5}	5.456	0.065	2.18×10^{-5}
Max H_1 pers.	15.593	0.000411	6.24×10^{-5}	11.748	0.00281	4.70×10^{-5}
Length-adjusted: Residuals from linear regression of each feature on time series length.						

Despite statistical significance, these differences are dwarfed by within-realm variation: Individual marine, terrestrial, and freshwater systems each span three orders of magnitude in Betti numbers, indicating that species-specific life histories, local environmental conditions, and stochastic processes generate far more variability than broad habitat classification.

3.4. Persistence analysis

The persistence analysis reveals hierarchical scale structure in ecological attractors (Figure 4, Table 5). Total persistence in dimension zero ($L_0 = 31.5$) exceeds dimension one ($L_1 = 3.03$) by an order of magnitude. A strong positive correlation between L_0 and L_1 (Spearman $\rho = 0.826$) indicates that topological complexity manifests coherently across dimensions.

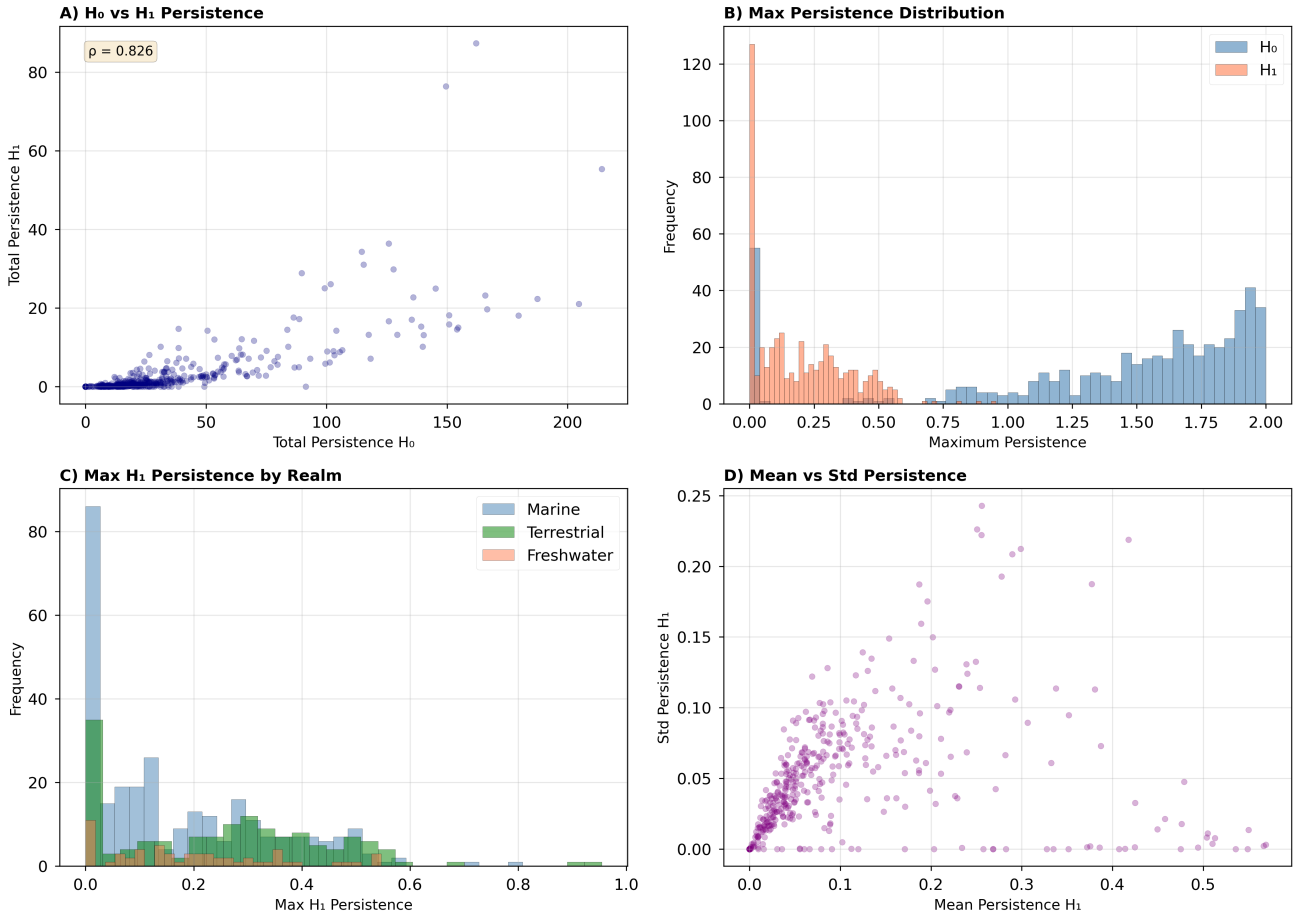

Figure 4. Persistence analysis. (A) Scatter plot of L_0 versus L_1 (Spearman $\rho = 0.826$). (B) Histograms of maximum persistence for H_0 and H_1. (C) Realm-specific distributions of maximum H_1 persistence. (D) Mean versus standard deviation of H_1 persistence ($\rho = 0.683$).

Table 5. Persistence statistics by realm. Standard deviations in parentheses.

Measure	All	Marine	Terrestrial	Freshwater
L_0	31.5 (35.7)	27.0 (31.2)	38.4 (40.9)	36.7 (39.5)
$\ell_0^{\max}$	1.40 (0.60)	1.36 (0.64)	1.42 (0.57)	1.53 (0.47)
L_1	3.03 (7.64)	2.53 (8.00)	4.34 (7.93)	2.26 (3.83)
$\ell_1^{\max}$	0.196 (0.181)	0.169 (0.170)	0.248 (0.200)	0.199 (0.156)

Maximum persistence in dimension zero shows bimodal distribution with concentration near $\ell_0^{\max} \approx 2.0$ (30.2% with $\ell_0^{\max} > 1.8$), and dimension one displays right-skewed distribution with only 6.2%, exceeding $\ell_1^{\max} = 0.5$. Terrestrial systems exhibit highest mean maximum H_1 persistence (0.248), marine systems the lowest (0.169), and freshwater intermediate (0.199).

3.5. Correlation structure

The correlation structure among topological features is presented in Figure 5 and Table 6. Exceptionally strong Spearman correlations emerge between Betti numbers and entropy measures:

β_1–H_1 ($\rho = 0.989$), β_0–H_0 ($\rho = 0.962$), and cross-dimensional β_1–H_0 ($\rho = 0.926$).

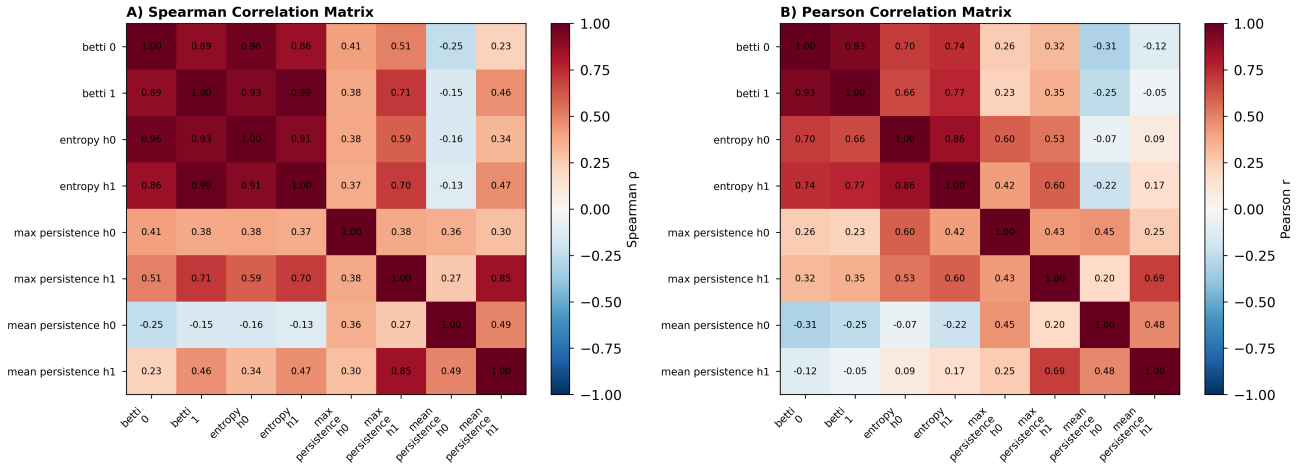

Figure 5. Correlation structure. (A) Spearman correlation matrix showing strong associations among Betti numbers and entropy measures. (B) Pearson correlation matrix showing attenuated magnitudes for entropy-related correlations, indicating nonlinear relationships.

Table 6. Strongest correlations among topological features. $\Delta = r - \rho$ indicates nonlinearity.

Feature 1	**Feature 2**	**Spearman ρ**	**Pearson r**	Δ
β_1	H_1 entropy	0.989	0.766	−0.223
β_0	H_0 entropy	0.962	0.696	−0.266
β_1	H_0 entropy	0.926	0.661	−0.265
β_0	β_1	0.888	0.933	+0.045
$\ell_1^{\max}$	$\bar{\ell}_1$	0.851	0.693	−0.158

Substantial discrepancies between Spearman and Pearson correlations (e.g., $\Delta = -0.266$ for β_0–H_0) indicate strongly monotonic but nonlinear relationships, reflecting the information-theoretic constraint $H_k \leq \log_2(\beta_k)$. Mean persistence in dimension zero ($\bar{\ell}_0$) exhibits weak negative correlations with Betti numbers, reflecting the geometric constraint that more features necessitate smaller typical persistence values.

3.6. Low-dimensional structure

To characterize the overall structure of the topological feature space, we applied dimensionality reduction to the six-dimensional space spanned by the six topological features listed in Table 7 (β_0, β_1, H_0 entropy, H_1 entropy, $\ell_0^{\max}$, and $\ell_1^{\max}$). This analysis revealed that topological variation occurs predominantly along a low-dimensional manifold (Figure 6, Table 7). The first principal component explains 64.9% of variance, with three components capturing 92.6%.

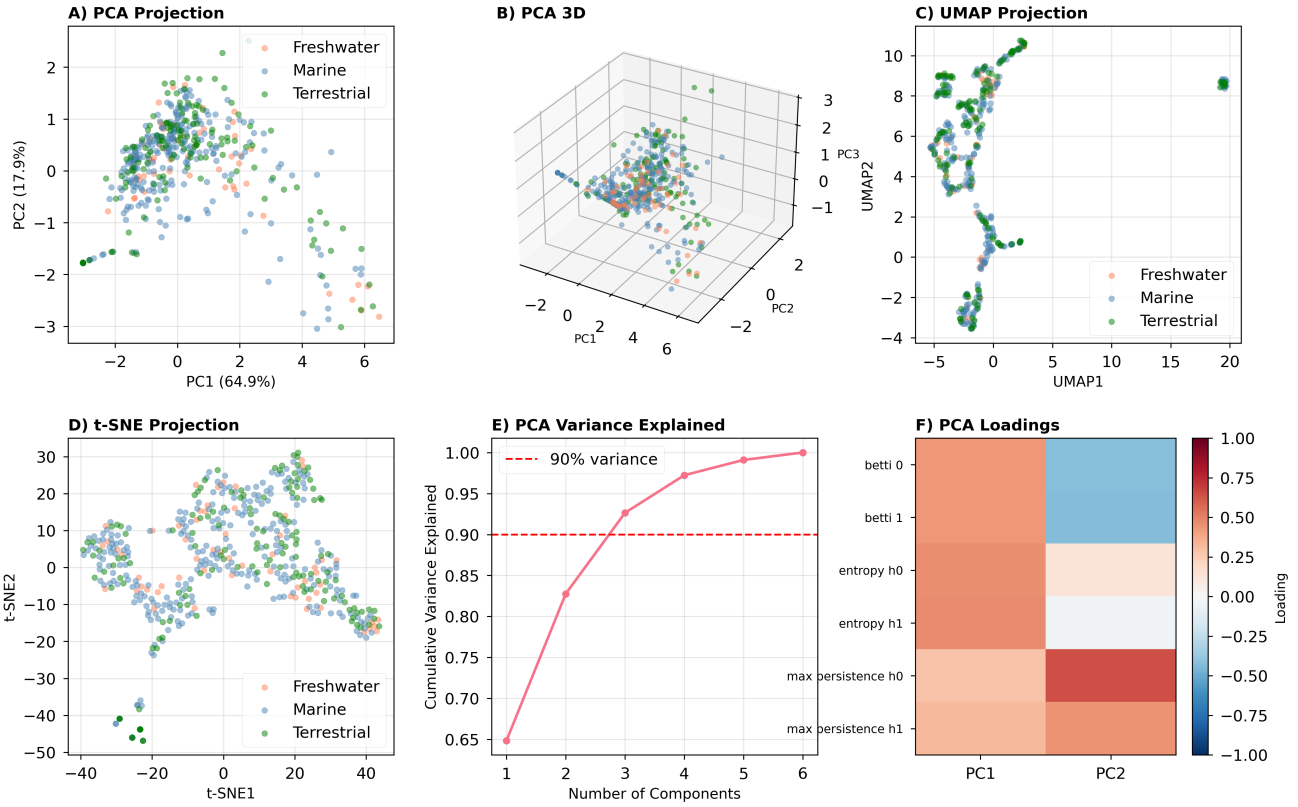

Figure 6. Dimensionality reduction of topological feature space. (A) PCA projection (PC1: 64.9%, PC2: 17.9%) showing substantial realm overlap. (B) Three-dimensional PCA with PC3 (9.9%). (C) UMAP projection showing curved manifold structure. (D) t-SNE projection. (E) Cumulative variance explained. (F) Component loadings heatmap.

Table 7. PCA summary and component loadings.

Component	Var.	Cum.	Feature	PC1	PC2	PC3
PC1	64.9%	64.9%	β_0	0.434	−0.426	−0.124
PC2	17.9%	82.8%	β_1	0.433	−0.436	−0.028
PC3	9.9%	92.6%	H_0 entropy	0.463	0.131	−0.183
PC4	4.6%	97.2%	H_1 entropy	0.472	−0.033	0.147
PC5	1.9%	99.1%	$\ell_0^{\max}$	0.285	0.641	−0.588
PC6	0.9%	100%	$\ell_1^{\max}$	0.326	0.446	0.764

PC1 loads uniformly and positively on all features (range: 0.285–0.472), representing an overall topological complexity axis. PC2 exhibits bipolar structure, contrasting Betti numbers (negative loadings) against maximum persistence values (positive loadings), reflecting a complexity-versus-prominence trade-off. PC3 captures scale differentiation across dimensions.

Realm centroids in PCA space show modest separation: Freshwater systems exhibit the highest mean PC1 score (0.612), followed by terrestrial (0.284) and marine (−0.266). However, centroid distances (0.33–0.89 standard deviations) are small relative to within-realm variance, with approximately 98% of time series overlapping across realms.

Uniform manifold approximation and projection (UMAP), shown in panel C, constructs a fuzzy topological representation based on local manifold approximations and optimizes a low-dimensional embedding preserving topological structure. The algorithm builds a weighted k-nearest neighbor graph ($k = 15$), where edge weights reflect local distances normalized by the distance to the kth nearest neighbor. The low-dimensional embedding minimizes cross-entropy between high-dimensional and low-dimensional fuzzy set representations through stochastic gradient descent. The UMAP projection reveals substantially greater local structure compared to PCA, organizing time series into a curved, roughly S-shaped or horseshoe-shaped manifold. This characteristic geometry emerges when high-dimensional data lie on a curved submanifold, with the horseshoe representing the topological constraint that low-complexity systems cluster near the origin, whereas high-complexity systems disperse along multiple axes. The embedding displayed several distinct density concentrations, particularly a dense cluster at intermediate coordinates and a more diffuse region extending toward positive UMAP1 values.

The t-distributed stochastic neighbor embedding (t-SNE), shown in panel D, minimizes Kullback–Leibler divergence between probability distributions in high-dimensional and low-dimensional spaces. In high-dimensional space, conditional probabilities are defined using Gaussian kernels with bandwidth determined adaptively to achieve target perplexity ($= 30$). In the embedding space, Student's t-distribution with one degree of freedom (Cauchy distribution) is employed, alleviating the crowding problem inherent in mapping high-dimensional data to low dimensions. The t-SNE projection fragments the topological feature space into several visually distinct clusters arranged in a roughly circular layout. The central region exhibits dense concentration of points from all three realms, and peripheral clusters appeared at various angles around the periphery. The tendency of t-SNE to exaggerate local density variations may amplify genuine multimodal structure or may reflect algorithmic artifacts from stochastic optimization.

Critically, realm colors remain extensively intermixed throughout both UMAP and t-SNE spaces, indicating that nonlinear structure captured by these methods primarily reflects variation in topological complexity independent of realm membership. The extensive realm intermixing within each cluster reinforces the conclusion that topological complexity varies predominantly along axes orthogonal to realm membership, with within-realm diversity exceeding between-realm differentiation.

4. Discussion

The application of persistent homology to 500 ecological time series reveals that topological variation occurs predominantly along a low-dimensional manifold, with three principal components capturing 92.6% of variance despite the six-dimensional feature space. The first component alone explains 64.9% of variance and loads uniformly on all topological features, representing a general complexity axis. This dimensional compression indicates that topological invariants covary in a coordinated manner, suggesting that complexity manifests holistically rather than through independent geometric aspects [10]. The low-dimensional structure implies that ecological attractor topology is governed by a small number of fundamental processes despite the potentially high-dimensional phase space of interacting populations and environmental variables.

Before discussing specific results, we address a potential interpretive concern: The finding that realm membership explains less than 0.01% of topological variance might appear to undermine the

study's contribution. We argue the opposite. The demonstration that a widely assumed structuring variable, habitat type, exerts negligible constraint on population dynamics is itself a substantive result. More importantly, the principal finding of this study is not the absence of realm effects but rather the presence of a common low-dimensional manifold governing ecological attractor topology: that topological complexity, despite spanning three orders of magnitude across 500 time series from three distinct ecosystem types, organizes along three dominant axes is a nontrivial and theoretically informative result. It suggests that ecological dynamics, however diverse in their particulars, are constrained to a restricted region of dynamical space, a finding that cannot be obtained through conventional linear or spectral methods.

Traditional approaches to ecological time series, such as autoregressive modeling, spectral analysis, and variance-based early warning indicators, have proven invaluable for detecting trends, dominant periodicities, and statistical signatures of regime shifts [24, 25]. However, these methods reduce complex temporal dynamics to one-dimensional summary statistics, potentially not capturing the full geometric organization of population trajectories in phase space. We asked whether population dynamics exhibit common patterns across ecosystems or whether marine, terrestrial, and freshwater habitats impose fundamental dynamical constraints. Our topological analysis reveals three key findings with direct implications for ecological understanding.

First, topological methods capture multiscale temporal organization that complements what conventional approaches provide. A system exhibiting $\beta_1 = 50$ persistent loops contains 50 topologically distinct cyclic features across scales, potentially representing overlapping annual breeding cycles, multiyear environmental oscillations (the El Niño–Southern Oscillation, ENSO; and the North Atlantic Oscillation, NAO), predator–prey interactions with different period lengths, and cohort dynamics in age-structured populations. Spectral analysis typically identifies the 2 or 3 strongest periodicities, whereas our approach reveals the complete hierarchy of concurrent temporal processes. For North Atlantic copepods represented in our marine dataset, populations respond simultaneously to seasonal temperature cycles, basin-scale climate indices with 3–7 year periods, and decadal regime shifts [31]. High β_1 values quantify this layered temporal structure that single-frequency methods are not designed to resolve. An ecologist relying solely on dominant spectral peaks might conclude dynamics are driven by annual cycles, missing the influence of multiple interacting timescales that may be critical for understanding population resilience and responses to environmental change.

Second, persistence entropy distinguishes organized from diffuse complexity. Two populations might both exhibit $\beta_1 = 40$, but persistence entropy $H_1 \approx 0$ indicates a single dominant cycle (low complexity despite many features) whereas $H_1 > 4$ reveals many cycles of comparable importance, reflecting genuinely complex multiscale dynamics. This distinction has clear ecological meaning: Low-entropy systems may be highly predictable but vulnerable to disruption of their dominant cycle, and high-entropy systems may exhibit dynamical robustness through functional redundancy across timescales. This parallels concepts from biodiversity theory, where evenness matters as much as richness. The exceptionally strong correlation between β_1 and H_1 ($\rho = 0.989$) reflects fundamental information-theoretic constraints, but the discrepancy between Spearman and Pearson correlations indicates strongly monotonic yet nonlinear relationships that merit further investigation.

Third, and most importantly, our results challenge the assumption that habitat type fundamentally constrains population dynamics, a perspective supported by documented differences in environmental variability structure across marine and terrestrial realms, with marine environments exhibiting redder

noise spectra than terrestrial systems [38, 39]. Although marine systems exhibit modestly lower topological complexity (mean $\beta_1 = 32.2$) compared to freshwater systems ($\beta_1 = 59.4$), potentially reflecting greater thermal buffering in oceans reducing concurrent environmental periodicities, realm membership explains a negligible proportion of topological variance both before and after controlling for time series length ($\epsilon^2 < 2 \times 10^{-4}$). Length adjustment further strengthens this conclusion: Betti number differences between realms remain robust, confirming that the core topological signatures reflect genuine dynamical variation rather than methodological artifacts. Within-realm heterogeneity vastly exceeds between-realm differences: A marine zooplankton species and a terrestrial songbird may exhibit more similar dynamical complexity than two marine species with different life histories. Although direct testing through life history covariates remains beyond the scope of the present study, this interpretation is broadly consistent with ecological research emphasizing species-specific traits such as generation time, body size, and habitat affinity, as primary determinants of population behavior, independent of broad environmental context [7, 34, 36]. This finding supports theoretical frameworks emphasizing shared mechanisms such as density dependence, predator–prey interactions, and environmental stochasticity that operate across ecosystems [1, 7, 34]. Classification schemes based on life history traits such as generation time, body size, and trophic position may be more dynamically relevant than broad habitat categories. The prevalence of chaotic dynamics across diverse ecosystems, as demonstrated experimentally in planktonic communities [32], reinforces this perspective that intrinsic biological interactions can generate complex dynamics independent of environmental context.

Methodologically, persistent homology offers critical advantages for ecological data. The stability theorem [17] guarantees robustness to measurement noise and sampling irregularities that plague field monitoring programs. Where Lyapunov exponent estimation fails on short noisy time series [11], topological invariants remain computable and interpretable. Furthermore, the low-dimensional structure we identified (92.6% of variance captured by three principal components) suggests population dynamics occupy a restricted region of possible dynamical space. This is a quantitative manifestation of the theoretical expectation that ecological interactions impose constraints on long-term behavior [2]. The first principal component, loading uniformly on all features, provides a single scalar measure of overall dynamical complexity that could facilitate comparative analyses across diverse systems.

Looking forward, topological approaches may enable the detection of regime shifts through geometric reorganization of attractors. [24] demonstrated that critical transitions often exhibit rising variance and autocorrelation, but these are scalar statistics that may not capture structural reorganization of attractors. A shift from limit-cycle dynamics ($\beta_1 \approx 1$) to quasiperiodic or chaotic behavior ($\beta_1 \gg 1$) represents fundamental reorganization potentially preceding ecosystem collapse. Recent work has demonstrated the potential of TDA-based approaches for early warning detection in other complex dynamical systems: Luo et al. [40] introduced a TDA-based framework for predicting infectious disease outbreaks from noisy surveillance data, and Song and Li [41] demonstrated that topological transitions in cryptocurrency systems precede extreme fluctuations in traditional financial markets by up to five days. These results suggest that analogous topological reorganization of reconstructed attractors may provide complementary early-warning signals in ecological systems, a direction that warrants future investigation. Time-varying analysis through sliding windows could track changes in topological complexity as systems approach tipping points [33]. Integration with mechanistic models would further clarify how ecological processes such as competition, predation, resource limitation translate into specific topological signatures, bridging phenomenological pattern

detection with process-based understanding.

Although statistical tests detected significant differences in topological features across marine, terrestrial, and freshwater realms prior to length adjustment (p ¡ 0.05 for all features), effect sizes are vanishingly small ($\epsilon^2 < 0.0001$), with realm membership explaining less than 0.01% of total variance. This illustrates the critical distinction between statistical and practical significance: Realm differences are reliably detectable with large sample sizes but are dwarfed by within-realm heterogeneity from species-specific characteristics, local conditions, and stochastic processes. The 98% overlap in topological distributions across realms indicates that realm membership provides minimal predictive power. This suggests that shared mechanisms such as predator–prey cycles and density dependence [1] may be more important determinants of topological complexity than broad environmental context.

The exceptionally strong correlation between first Betti numbers and persistence entropy ($\rho = 0.989$) reflects an information-theoretic constraint: Shannon entropy increases monotonically with the number of distinguishable elements, subject to $H_k \leq \log_2(\beta_k)$. The discrepancy between Spearman and Pearson correlations ($\Delta = -0.223$) indicates a strongly monotonic but nonlinear relationship. For machine learning applications, this near-redundancy suggests one metric may suffice for prediction, although both provide complementary interpretive perspectives.

Ecological interpretation of topological invariants requires careful consideration. Zeroth Betti numbers quantifying connected components may reflect alternative stable states, environmental forcing creating distinct regimes, sampling artifacts, or chaotic dynamics. First Betti numbers enumerating loops capture cyclic behavior potentially from predator–prey interactions or periodic forcing. For example, a planktivorous fish population might exhibit $\beta_1 = 1$ from a single dominant annual reproductive cycle, whereas a longer-lived predator feeding on that fish might show $\beta_1 = 3$ from overlapping cohorts with two-to-three-year maturation times coupled to the prey's annual cycle. Systems with $\beta_1 > 20$ may reflect complex communities with multiple interacting species exhibiting different periodicities or single populations responding to multiple environmental drivers operating at different timescales such as seasonal forcing, ENSO events, and decadal climate oscillations occurring simultaneously. The 22.4% of systems with $\beta_1 = 0$ indicate that many time series lack detectable cyclic structure. Persistence entropy quantifies distributional complexity, with high entropy indicating multi-scale structure and low entropy suggesting simpler organization with characteristic frequencies.

Methodological considerations include the embedding dimension $m = 3$, which is appropriate for low-dimensional attractors but may miss structure in higher-dimensional chaotic systems. The time delay heuristic ensures consistency but may not optimize individual embeddings. Time series length heterogeneity (25 to 2000 observations) introduces variation in attractor sampling density: Shorter series may undersample complex attractors and provide conservative estimates of topological complexity. However, sensitivity analysis restricting to series with $n \geq 50$ ($n = 332$) yields nearly identical results: three principal components capturing 92.5% of variance, β_1–H_1 correlation of $\rho = 0.987$, and uniformly negligible realm effect sizes ($\epsilon^2 < 0.002$), confirming that the main conclusions are robust to the minimum series length threshold. The stability theorem for persistent homology [17] provides further reassurance that moderate parameter variations preserve topological features. A fundamental limitation of the present study concerns the treatment of chaotic dynamics. The embedding dimension $m = 3$ is theoretically sufficient for attractors of dimension $d \leq 1$ (fixed points and limit cycles), but chaotic attractors typically require substantially higher embedding dimensions ($m \geq 5$) and exhibit fractal geometry that may not be fully captured by the topological invariants

computed here. Although chaotic dynamics have been documented in ecological systems [2, 32], the present framework cannot reliably characterize chaotic attractor geometry. Extending the analysis to higher embedding dimensions and incorporating higher-dimensional homology (H_2) to detect the void structure characteristic of strange attractors represents an important direction for future research.

The findings suggest several implications. First, attractor geometry may serve as a dimensionality-reduction target for complex ecological data, with the first principal component providing a scalar complexity index. Second, the dominance of within-realm variance supports theoretical frameworks emphasizing shared mechanisms across ecosystems. Third, the diversity of topological signatures spanning three orders of magnitude suggests potential for classification based on topological features.

Future research directions include extending to higher embedding dimensions, analyzing time-varying topological features through sliding windows to detect regime shifts [3], integrating with mechanistic models to link topology to ecological processes, developing null models for hypothesis testing, and linking topological complexity to functional outcomes such as stability and resilience. The theoretical framework of Takens' embedding [13] and persistent homology [15, 16] provides a rigorous foundation with stability guarantees [17], though bridging from abstract invariants to concrete mechanisms requires integration with experimental and modeling approaches. The diversity of topological signatures suggests that population dynamics span a vast dynamical space, with variation occurring predominantly along system-specific rather than realm-specific gradients.

5. Conclusions

This study shows that persistent homology provides a mathematically rigorous framework for characterizing population dynamics that reveals multiscale temporal organization not directly captured by conventional linear methods while being robust to the noise and sampling irregularities characteristic of ecological field data. The analysis of 500 time series spanning marine, terrestrial, and freshwater ecosystems from the BioTIME database yields four principal conclusions with implications for both ecological theory and the application of topological data analysis to dynamical systems.

First, this study establishes that the six-dimensional space of topological invariants comprising Betti numbers, persistence statistics, and persistence entropy across two homological dimensions organizes along a low-dimensional manifold that is robust to embedding dimension choice ($m \in \{2, 3, 4, 5\}$) and series length threshold. Three principal components capture 92.6% of total variance, with the first component alone explaining 64.9% through uniform positive loadings across all features. This dimensional compression is a mathematically nontrivial result: It implies that ecological attractor topology, despite arising from diverse nonlinear processes, is constrained to a restricted region of the space of possible topological invariants.

Second, ecological realm membership exerts practically negligible influence on attractor topology, a conclusion that is further strengthened after controlling for time series length. Although Kruskal–Wallis tests identified significant differences across marine, terrestrial, and freshwater systems for all analyzed features ($p < 0.05$), effect sizes were uniformly small ($\epsilon^2 < 0.0001$), with realm explaining less than 0.01% of topological variance. The 98% overlap in topological distributions across realms indicates that within-realm heterogeneity, driven by species-specific life history characteristics, local environmental conditions, and stochastic processes, vastly exceeds between-realm differentiation. This finding supports theoretical frameworks emphasizing shared dynamical mechanisms such as density

dependence and predator–prey interactions that transcend habitat boundaries, and it suggests that life history traits may be more informative predictors of dynamical complexity than broad environmental classification. We emphasize that demonstrating the negligible explanatory power of a widely assumed structuring variable [38, 39] is itself a substantive contribution: It redirects attention from habitat-level comparisons toward species- and community-level determinants of attractor geometry.

Third, the paper provides large-scale empirical confirmation of the information-theoretic bound $H_k \leq \log_2(\beta_k)$ across 500 independently reconstructed attractors. The Spearman correlation $\rho = 0.989$ between β_1 and H_1, combined with the substantial discrepancy between Spearman and Pearson correlations ($\Delta = -0.223$), reveals strongly monotonic but nonlinear geometric structure in the relationship between cardinality-based and distributional topological descriptors. This result has direct implications for topological feature selection in machine learning applications of TDA to dynamical systems.

Fourth, ecological time series span an enormous range of topological complexity, with Betti numbers varying across three orders of magnitude from systems with zero persistent loops to those exhibiting hundreds. This diversity of topological signatures, combined with the low-dimensional manifold structure of the feature space, suggests that population dynamics occupy a vast but structured dynamical space amenable to systematic classification and comparison.

The methodological framework established here, combining Takens' embedding theorem for phase space reconstruction with persistent homology for multiscale topological characterization, provides a principled approach for extracting geometric information from ecological time series that complements traditional linear and spectral methods. The stability guarantees of persistent homology ensure robustness to the measurement noise and sampling irregularities characteristic of field monitoring data. Future applications may extend this framework to detect regime shifts through time-varying topological analysis, link topological invariants to mechanistic ecological models, and develop null models for hypothesis testing regarding attractor geometry. The predominance of system-specific over realm-specific topological variation observed here suggests that the search for general principles governing ecological dynamics may be most fruitful when focused on shared mechanisms transcending broad environmental classifications.

Use of Generative-AI tools declaration

The author declares he has not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The author declares that no external funding was received for this research. The author thanks the BioTIME database contributors for making the ecological time series data publicly available.

Conflict of interest

The author declares no conflict of interest.

References

1. R. M. May, Simple mathematical models with very complicated dynamics, *Nature*, **261** (1976), 459–467. https://doi.org/10.1038/261459a0
2. A. Hastings, C. L. Hom, S. Ellner, P. Turchin, H. C. J. Godfray, Chaos in ecology: Is mother nature a strange attractor?, *Annu. Rev. Ecol. Syst.*, **24** (1993), 1–33. https://doi.org/10.1146/annurev.es.24.110193.000245
3. M. Scheffer, S. Carpenter, J. A. Foley, C. Folke, B. Walker, Catastrophic shifts in ecosystems, *Nature*, **413** (2001), 591–596. https://doi.org/10.1038/35098000
4. E. N. Lorenz, Deterministic nonperiodic flow, *J. Atmos. Sci.*, **20** (1963), 130–141. https://doi.org/10.1175/1520-0469(1963)020¡0130:DNF¿2.0.CO;2
5. N. H. Packard, J. P. Crutchfield, J. D. Farmer, R. S. Shaw, Geometry from a time series, *Phys. Rev. Lett.*, **45** (1980), 712–716. https://doi.org/10.1103/PhysRevLett.45.712
6. S. Ellner, P. Turchin, Chaos in a noisy world: New methods and evidence from time series analysis, *Am. Nat.*, **145** (1995), 343–375. https://doi.org/10.1086/285547
7. P. Turchin, *Complex population dynamics: A theoretical/empirical synthesis*, Princeton University Press, Princeton, NJ, USA, 2003.
8. G. Sugihara, R. May, H. Ye, C. H. Hsieh, E. Deyle, M. Fogarty, et al., Detecting causality in complex ecosystems, *Science*, **338** (2012), 496–500. https://doi.org/10.1126/science.1227079
9. E. R. Deyle, R. M. May, S. B. Munch, G. Sugihara, Tracking and forecasting ecosystem interactions in real time, *Proc. Natl. Acad. Sci. USA*, **113** (2016), E4736–E4744. https://doi.org/10.1073/pnas.1517161113
10. S. H. Strogatz, *Nonlinear dynamics and chaos*, Westview Press, Boulder, CO, USA, 1994.
11. M. T. Rosenstein, J. J. Collins, C. J. De Luca, A practical method for calculating largest Lyapunov exponents from small data sets, *Physica D*, **65** (1993), 117–134. https://doi.org/10.1016/0167-2789(93)90009-P
12. H. Kantz, A robust method to estimate the maximal Lyapunov exponent of a time series, *Phys. Lett. A*, **185** (1994), 77–87. https://doi.org/10.1016/0375-9601(94)90991-1
13. F. Takens, *Detecting strange attractors in turbulence*, in Dynamical Systems and Turbulence, Warwick 1980, Springer, Berlin, Heidelberg, 1981, 366–381. https://doi.org/10.1007/BFb0091924
14. T. Sauer, J. A. Yorke, M. Casdagli, Embedology, *J. Stat. Phys.*, **65** (1991), 579–616. https://doi.org/10.1007/BF01053745
15. H. Edelsbrunner, D. Letscher, A. Zomorodian, Topological persistence and simplification, *Discrete Comput. Geom.*, **28** (2002), 511–533. https://doi.org/10.1007/s00454-002-2885-2
16. A. Zomorodian, G. Carlsson, Computing persistent homology, *Discrete Comput. Geom.*, **33** (2005), 249–274. https://doi.org/10.1007/s00454-004-1146-y
17. D. Cohen-Steiner, H. Edelsbrunner, J. Harer, Stability of persistence diagrams, *Discrete Comput. Geom.*, **37** (2007), 103–120. https://doi.org/10.1007/s00454-006-1276-5

18. P. Bubenik, Statistical topological data analysis using persistence landscapes, *J. Mach. Learn. Res.*, **16** (2015), 77–102.

19. J. A. Perea, J. Harer, Sliding windows and persistence: An application of topological methods to signal analysis, *Found. Comput. Math.*, **15** (2015), 799–838. https://doi.org/10.1007/s10208-014-9206-z

20. J. R. Tempelman, F. A. Khasawneh, A look into chaos detection through topological data analysis, *Physica D*, **406** (2020), 132446. https://doi.org/10.1016/j.physd.2020.132446

21. L. M. Seversky, S. Davis, M. Berger, On time-series topological data analysis, in *CVPR Workshops*, 2016. https://doi.org/10.1109/CVPRW.2016.19

22. Y. Hiraoka, T. Nakamura, A. Hirata, E. G. Escolar, K. Matsue, Y. Nishiura, Hierarchical structures of amorphous solids characterized by persistent homology, *Proc. Natl. Acad. Sci. USA*, **113** (2016), 7035–7040. https://doi.org/10.1073/pnas.1520877113

23. C. Giusti, E. Pastalkova, C. Curto, V. Itskov, Clique topology reveals intrinsic geometric structure in neural correlations, *Proc. Natl. Acad. Sci. USA*, **112** (2015), 13455–13460. https://doi.org/10.1073/pnas.1506407112

24. M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, et al., Early-warning signals for critical transitions, *Nature*, **461** (2009), 53–59. https://doi.org/10.1038/nature08227

25. V. Dakos, S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, et al., Methods for detecting early warnings of critical transitions, *PLoS One*, **7** (2012), e41010. https://doi.org/10.1371/journal.pone.0041010

26. M. Dornelas, L. H. Antão, F. Moyes, A. E. Bates, A. E. Magurran, D. Adam, et al., BioTIME: A database of biodiversity time series, *Glob. Ecol. Biogeogr.*, **27** (2018), 760–786. https://doi.org/10.1111/geb.12729

27. G. Carlsson, Topology and data, *Bull. Am. Math. Soc.*, **46** (2009), 255–308. https://doi.org/10.1090/S0273-0979-09-01249-X

28. U. Bauer, Ripser: Efficient computation of Vietoris–Rips persistence barcodes, *J. Appl. Comput. Topol.*, **5** (2021), 391–423. https://doi.org/10.1007/s41468-021-00071-5

29. L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: Uniform manifold approximation and projection, *J. Open Source Softw.*, **3** (2018), 861. https://doi.org/10.21105/joss.00861

30. L. van der Maaten, G. Hinton, Visualizing data using t-SNE, *J. Mach. Learn. Res.*, **9** (2008), 2579–2605.

31. G. Beaugrand, P. C. Reid, F. Ibañez, J. A. Lindley, M. Edwards, Reorganization of north atlantic marine copepod biodiversity and climate, *Science*, **296** (2002), 1692–1694. https://doi.org/10.1126/science.1071329

32. E. Beninçà, J. Huisman, R. Heerkloss, K. D. Jöhnk, P. Branco, E. H. Van Nes, et al., Chaos in a long-term experiment with a plankton community, *Nature*, **451** (2008), 822–825. https://doi.org/10.1038/nature06512

33. O. N. Bjørnstad, B. T. Grenfell, Noisy clockwork: Time series analysis of population fluctuations in animals, *Science*, **293** (2001), 638–643. https://doi.org/10.1126/science.1062226

34. B. E. Kendall, C. J. Briggs, W. W. Murdoch, P. Turchin, S. P. Ellner, E. McCauley, et al., Why do populations cycle? A synthesis of statistical and mechanistic modeling approaches, *Ecology*, **80** (1999), 1789–1805. https://doi.org/10.1890/0012-9658(1999)080[1789:WDPCAS]2.0.CO;2

35. A. Jin, P. Wang, G. Zhang, H. Shi, H. Li, Ecological quality and spatial structure dynamics under future scenarios: A topological perspective from the Yellow River Basin, *J. Clean. Prod.*, **522** (2025), 146346. https://doi.org/10.1016/j.jclepro.2025.146346

36. A. Jin, H. Shi, Y. Ma, P. Wang, G. Zhang, H. Li, Optimizing multi-species avian migration networks using an integrated habitat suitability and connectivity framework: A case study of the Changzhutan Metropolitan Area, *Environ. Impact Assess. Rev.*, **120** (2026), 108442. https://doi.org/10.1016/j.eiar.2026.108442

37. K. Strommen, M. Chantry, J. Dorrington, N. Otter, A topological perspective on weather regimes, *Clim. Dyn.*, **60** (2023), 1415–1445. https://doi.org/10.1007/s00382-022-06395-x

38. J. H. Steele, A comparison of terrestrial and marine ecological systems, *Nature*, **313** (1985), 355–358. https://doi.org/10.1038/313355a0

39. D. A. Vasseur, P. Yodzis, The color of environmental noise, *Ecology*, **85** (2004), 1146–1152. https://doi.org/10.1890/02-3122

40. X. Luo, Q. H. Liu, X. Lu, Z. Wang, Z. Jin, L. Li, et al., TDA-PIDO: A topological data analysis approach for early warning of infectious disease outbreaks, *Innovation*, 2026, 101396. https://doi.org/10.1016/j.xinn.2026.101396

41. S. Song, H. Li, Can topological transitions in cryptocurrency systems serve as early warning signals for extreme fluctuations in traditional markets?, *Physica A*, **657** (2025), 130194. https://doi.org/10.1016/j.physa.2024.130194

42. İ. Güzel, E. Munch, F. A. Khasawneh, Detecting bifurcations in dynamical systems with CROCKER plots, *Chaos*, **32** (2022), 093111. https://doi.org/10.1063/5.0102421

43. A. Myers, D. Muñoz, F. A. Khasawneh, E. Munch, Temporal network analysis using zigzag persistence, *EPJ Data Sci.*, **12** (2023), 6. https://doi.org/10.1140/epjds/s13688-023-00379-5

44. W. H. Shah, S. R. Fatima, G. Huerta-Cuellar, J. H. García-López, C. G. Mata Ramirez, R. Jaimes-Reátegui, Topological data analysis approach to time series and shape analysis of dynamical system, *Chaos*, **35** (2025), 063129. https://doi.org/10.1063/5.0268340

AIMS Press