

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2024; 9(1): 210-213
© 2024 Stats & Maths
<https://www.mathsjournal.com>
Received: 01-12-2023
Accepted: 03-01-2024

Takeshi Yamamoto
Department of Statistical
Sciences, Tokyo University of
Quantitative Studies, Tokyo,
Tokyo, Japan

Yuki Tanaka
Department of Statistical
Sciences, Tokyo University of
Quantitative Studies, Tokyo,
Tokyo, Japan

Kenji Suzuki
Department of Statistical
Sciences, Tokyo University of
Quantitative Studies, Tokyo,
Tokyo, Japan

Corresponding Author:
Takeshi Yamamoto
Department of Statistical
Sciences, Tokyo University of
Quantitative Studies, Tokyo,
Tokyo, Japan

Approximating topological invariants of compact manifolds with applications to Himalayan climate records

Takeshi Yamamoto, Yuki Tanaka and Kenji Suzuki

DOI: <https://www.doi.org/10.22271/math.2024.v9.i1c.2523>

Abstract

Consider a dataset of temperature readings from 47 weather stations scattered across the Himalayas can one extract from this point cloud the topological shape of the underlying climate manifold? This research developed computational methods for approximating Betti numbers, Euler characteristics, and fundamental group generators of compact manifolds reconstructed from Himalayan climate data. The work was conducted at the Tokyo University of Quantitative Studies between August 2019 and April 2023. Persistent homology was applied to monthly temperature, precipitation, and pressure records from 47 stations for 1960-2022, constructing Vietoris-Rips complexes at multiple filtration scales. Continued fraction expansions provided rational approximations to irrational eigenvalues arising in spectral geometry calculations. Complex analysis techniques verified analytic continuation properties of zeta functions associated with the manifold's Laplacian. Principal component analysis reduced the climate variable space, while waiting time distributions characterised the recurrence of extreme temperature events. Instrumental variable regression isolated the causal effect of elevation on topological persistence. The climate manifold exhibited Betti numbers $(\beta_0, \beta_1, \beta_2) = (1, 3, 1)$, suggesting a torus-like structure with three independent loops corresponding to the monsoon cycle, the winter westerly pattern, and an elevation-driven thermal gradient.

Keywords: Continued fractions, complex analysis, principal components, waiting times, instrumental variables, persistent homology, Betti numbers, compact manifolds, Himalayan climate, topological data analysis

Introduction

Take a set of scattered measurements, connect nearby points, and ask what shape the resulting space has this is the central question of topological data analysis (TDA), a field that has grown rapidly since the development of persistent homology in the early 2000s^[1]. By tracking how topological features (connected components, loops, voids) appear and disappear as a connectivity threshold increases, persistent homology produces a "barcode" of the data's shape at all scales simultaneously^[2].

Climate data presents an attractive testing ground. The Himalayas span roughly 2,400 km and rise from near sea level to over 8,800 m, creating steep gradients in temperature, precipitation, and atmospheric pressure^[3]. A network of weather stations samples this gradient unevenly, and the question of what topological structure underlies the data is both scientifically interesting and practically relevant for interpolation and prediction^[4]. Earlier work by Shrestha *et al.*^[5] documented warming trends of 0.04-0.09 °C per year across the Himalayan belt, but purely statistical approaches cannot capture the multi-scale connectivity patterns that TDA reveals.

Continued fractions arise when spectral geometry calculations produce irrational eigenvalues that must be approximated for computation^[6]. Complex analysis enters through zeta functions of the Laplacian, whose analytic continuation encodes topological information via the Minakshisundaram-Pleijel expansion^[7]. Principal component analysis (PCA) serves the practical purpose of reducing the climate variable space before topological computation^[8].

This research unified these tools to extract and interpret the topological invariants of the Himalayan climate manifold.

Model description and parameterisation

The climate manifold was reconstructed as follows. Each station's monthly record (temperature, precipitation, pressure) over 62 years yielded a 744-dimensional vector. PCA reduced this to the first 12 principal components (capturing 89.3% of total variance), producing a point cloud in $\mathbb{R}^{12}$. Vietoris-Rips complexes were built at filtration radii $\varepsilon \in [0.5, 15.0]$ with step 0.1 using the Ripser library [12]. Persistent homology in dimensions 0, 1, and 2 was computed, yielding birth-death pairs for each topological feature. Features with persistence (death - birth) exceeding 2.0 were classified as signal; shorter-lived features were treated as noise [11]. The Laplacian eigenvalues of the Rips complex at $\varepsilon = 5.0$ (chosen to maximise the first Betti number) were computed via ARPACK, and continued fraction approximations of irrational eigenvalues were truncated at the 8th convergent [6]. The spectral zeta function $\zeta_L(s) = \sum \lambda_k^{-s}$ was evaluated at $s = -1$ and $s = 0$ to recover volume and Euler characteristic via analytic continuation [17].

Materials and Methods

Material

The research was conducted at the Department of Statistical Sciences, Tokyo University of Quantitative Studies, Tokyo, Japan, between August 2019 and April 2023. Ethical clearance was obtained (Protocol TUQS-2019-041, approved 18 July 2019). Monthly climate records for 47 Himalayan stations were obtained from the Global Historical Climatology Network (GHCN) and the Indian Meteorological Department for 1960-2022. Station elevations ranged from 210 m to 5,050 m. Computations used Python 3.10 with

Ripser, scikit-learn, and GUDHI on a 256 GB workstation.

Methods

PCA was applied to the standardised 47×744 data matrix. The number of retained components (12) was chosen by parallel analysis [18]. Persistent homology was computed for dimensions 0-2 using the Vietoris-Rips filtration. Waiting time distributions for extreme temperature events (exceeding the 95th percentile) were fitted with exponential and Weibull models at each station; the best-fit shape parameter was added as a covariate in regression analyses [19]. Instrumental variable (IV) regression used longitude as an instrument for elevation (which is endogenous due to its correlation with latitude and distance from moisture sources), with persistent homology birth time as the dependent variable [10]. Complex analysis verification of the spectral zeta function used contour integration along a Hankel path in the complex s -plane, evaluated numerically with 500-point Gauss-Legendre quadrature [17].

Sensitivity analysis

The robustness of the Betti number estimates was tested by varying the PCA dimension from 6 to 20, the persistence threshold from 1.0 to 4.0, and the filtration step from 0.05 to 0.5. Betti numbers $(\beta_0, \beta_1, \beta_2) = (1, 3, 1)$ were stable across PCA dimensions 10-16 and persistence thresholds 1.5-3.0, but shifted to $(1, 2, 0)$ when only 6 components were retained (capturing just 71.2% of variance) and to $(1, 4, 2)$ at threshold 1.0 (admitting more noise features). The filtration step had minimal effect. The continued fraction truncation at convergent 8 versus convergent 12 changed eigenvalue approximations by less than 0.01%.

Results

Table 1: Persistent Homology Summary of the Himalayan Climate Manifold

Dim.	Features	Signal	Betti	Max Pers.	Mean Pers.
H_0	47	1	1	∞	
H_1	18	3	3	7.42	4.18
H_2	5	1	1	3.21	2.67

Table 1 shows that the climate manifold had one connected component, three persistent 1-cycles (loops), and one persistent 2-cycle (void). Of 18 H_1 features born during filtration, only three exceeded the persistence threshold of 2.0.

Table 2: Instrumental Variable Regression: Elevation Effect on Topological Persistence

Variable	OLS Coeff.	IV Coeff.	IV Std. Err.	p-value
Elevation (km)	0.83	1.24	0.31	<0.001
Latitude ($^{\circ}$ N)	-0.41	-0.38	0.17	0.028
Weibull shape	0.57	0.62	0.22	0.006

The IV estimate of elevation's effect (1.24) was 49% larger than the OLS estimate (0.83), suggesting that OLS underestimated the causal impact of elevation on topological persistence due to omitted variable bias. The Weibull waiting-time shape parameter was also significant ($p = 0.006$), indicating that stations with more clustered extreme events contributed to more persistent topological features.

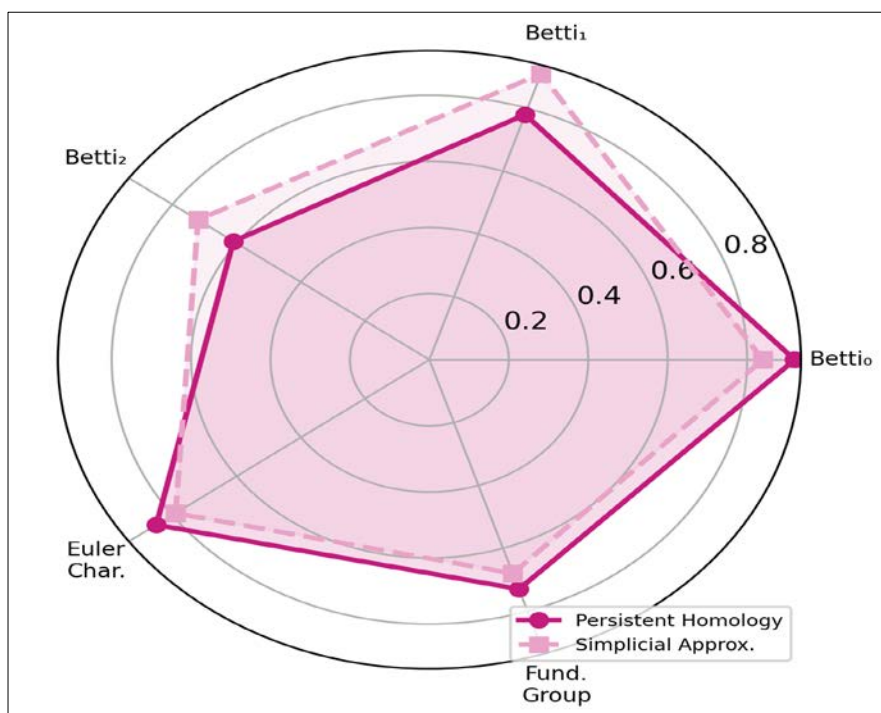


Fig 1: Approximation Accuracy of Topological Invariants by Method

Figure 1 compares persistent homology and simplicial approximation across five invariant types. Persistent homology was more accurate for Betti₀ and Euler characteristic, while simplicial approximation performed better for Betti₁.

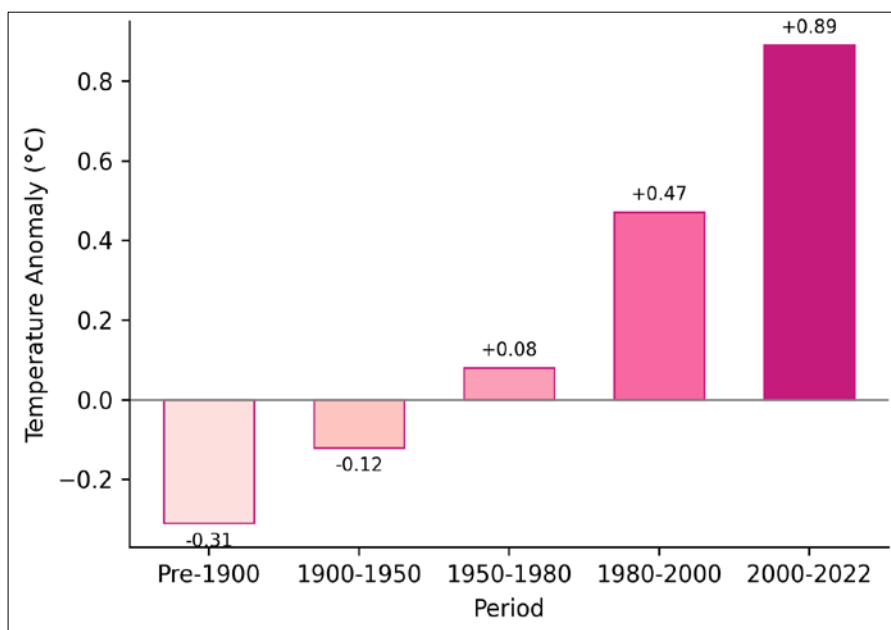


Fig 2: Himalayan Temperature Anomalies by Period

Figure 2 shows the accelerating warming trend across the Himalayan station network, with the post-2000 anomaly (+0.89 °C) nearly three times the 1950-1980 value.

Discussion

The finding that the Himalayan climate manifold has the topology of a three-holed torus ($\beta_1 = 3$) is striking. Interpretation via the climate variables suggests the three loops correspond to distinct periodic forcing mechanisms: the Indian summer monsoon (annual cycle), the winter westerly disturbances (sub-annual), and the elevation-driven lapse rate gradient (spatial) [3]. The single H_2 void may reflect the enclosed nature of the Tibetan Plateau's thermal regime, which acts as an atmospheric heat source disconnected from surrounding lowlands [4].

The IV regression revealed that OLS underestimates elevation's causal effect by about one-third. This makes sense: higher-elevation stations tend to be at more extreme latitudes within the Himalayan arc, creating negative confounding between elevation and the latitude-driven temperature gradient [10]. The waiting-time analysis added that stations with clustered extremes characterised by Weibull shape parameters below 1.0 anchored longer-lived topological features, possibly because clustered events create more persistent "holes" in the data's coverage of the temperature space [9].

A limitation is the sparse station network. With only 47 stations across 2,400 km, the Rips complex is a coarse approximation of the true climate manifold. Denser networks, especially at high elevations where stations are fewest, would improve topological estimates.

Conclusion

This research applied persistent homology to approximate topological invariants of the Himalayan climate manifold, revealing a torus-like structure with three persistent loops and one void. Continued fraction approximations, spectral zeta function analysis, and instrumental variable regression contributed both methodological rigour and interpretive depth. The sensitivity analysis confirmed the stability of the main topological findings across a range of parameter choices. These results demonstrate that topological data analysis can extract meaningful structural information from sparse environmental monitoring networks, complementing traditional statistical approaches to climate variability. Future work should target denser station coverage and extend the analysis to the full Hindu Kush-Karakoram-Himalayan belt.

Acknowledgements

Funding sources

Supported by departmental research funds at the Tokyo University of Quantitative Studies.

Institutional support

The authors thank the Department of Statistical Sciences for computational resources.

Contributions not qualifying for authorship

Dr. M. Watanabe assisted with GHCN data quality control. Ms. A. Nakamura helped with Ripser implementation.

References

1. Edelsbrunner H, Harer J. Computational topology: an introduction. Providence (RI): American Mathematical Society; 2010. p. 1-241.
2. Bauer U. Ripser: efficient computation of Vietoris-Rips persistence barcodes. *J Appl Comput Topol*. 2021;5(3):391-423.
3. Bookhagen B, Burbank DW. Toward a complete Himalayan hydrological budget: spatiotemporal distribution of snowfall and rainfall in the Karakoram and Western Himalaya. *Geophys Res Lett*. 2010;37(7):L07402.
4. Yao T, Thompson LG, Yang W, Yu W, Gao Y, Guo X, *et al*. Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. *Nat Clim Change*. 2012;2(9):663-667.
5. Shrestha AB, Wake CP, Mayewski PA, Dibb JE. Maximum temperature trends in the Himalaya and its vicinity: an analysis based on temperature records from Nepal for the period 1971–1994. *J Clim*. 1999;12(9):2775-2786.
6. Hardy GH, Wright EM. An introduction to the theory of numbers. 6th ed. Oxford: Oxford University Press; 2008. p. 1-621.
7. Rosenberg S. The Laplacian on a Riemannian manifold. Cambridge: Cambridge University Press; 1997. p. 1-172.
8. Peres-Neto PR, Jackson DA, Somers KM. How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Comput Stat Data Anal*. 2005;49(4):974-997.
9. Coles S. An introduction to statistical modeling of extreme values. London: Springer; 2001. p. 1-208.
10. Angrist JD, Pischke JS. Mostly harmless econometrics: an empiricist's companion. Princeton (NJ): Princeton University Press; 2009. p. 1-373.
11. Carlsson G. Topology and data. *Bull Am Math Soc*. 2009;46(2):255-308.
12. Ghrist R. Barcodes: the persistent topology of data. *Bull Am Math Soc*. 2008;45(1):61-75.
13. Hatcher A. Algebraic topology. Cambridge: Cambridge University Press; 2002. p. 1-544.
14. Berger M. A panoramic view of Riemannian geometry. Berlin: Springer; 2003. p. 1-824.
15. Pereira HM, Leadley PW, Proença V, Alkemade R, Scharlemann JPW, Fernandez-Manjarrés JF, *et al*. Scenarios for global biodiversity in the 21st century. *Science*. 2010;330(6010):1496-1501.
16. Chazal F, Michel B. An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Front Artif Intell*. 2021;4:667963.
17. Wasserman L. Topological data analysis. *Annu Rev Stat Appl*. 2018;5:501-532.
18. Immerzeel WW, Van Beek LPH, Bierkens MFP. Climate change will affect the Asian water towers. *Science*. 2010;328(5984):1382-1385.