Characterization of Climate Change in Selected African Countries using Topological Data Analysis

Peguy Kem-Meka Tiotsop Kadzue

(pkadzue@quantumleapafrica.org)
Doctoral Training Program in Data Science in Quantum Leap Africa(QLA)

AIMS-UoL Joint Postgraduate Conference

June 15, 2023

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

UNIVERSITY OF LIVERPOOL

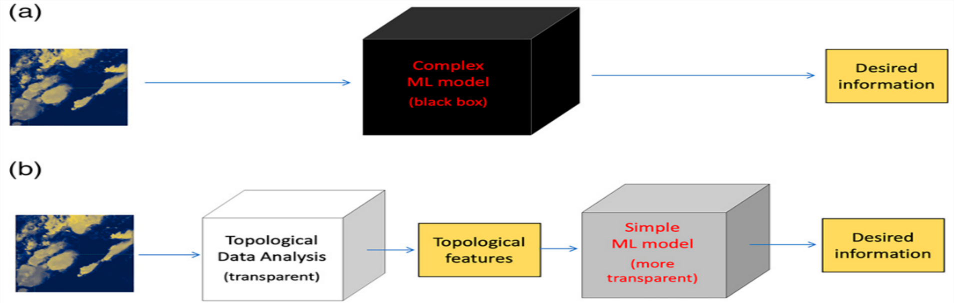
QUANTUM LEAP AFRICA

Outline

Introduction

Aim

- Investigate the application of **Topological Data Analysis (TDA)** as a reliable tool for analyzing (detect and characterize) climate change.

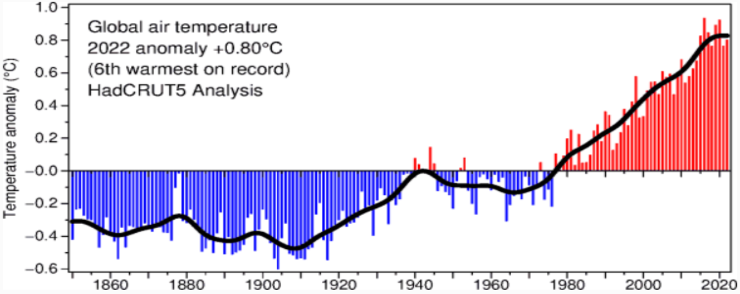


[Source: Hoef et al., 2022]

- **TDA** approach can be more **transparent**, **computationally efficient** and more **expressive.**

What is CLIMATE CHANGE?

- Refers to any significant **change in the measures of climate** lasting for an extended period of time.
- Includes major changes in **temperature**, **precipitation** which are **fundamental measurements** for describing the climate.



Global temperature record - updated February 2023

[Source: Climate Research Unit, UEA, 2023]

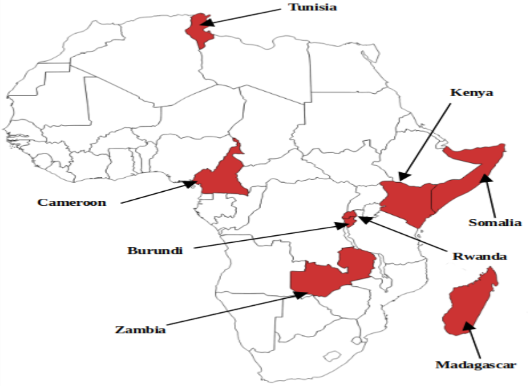
Why are we concerned?

Changes in **temperature** and **precipitation**:

- Can disrupt a wide range of environmental processes impacting agricultural productivity.
- Have led to food insecurity, flood, and droughts in several places in the world.

[Source: MS4CR Women Climate Scientist Publication 2018]

Africa, the most impacted continent!



- Africa is particularly vulnerable to climate change impacts,
- Floods,
- Droughts,
- Storms,
- Main sources of agricultural risk.
- World's hungriest countries 2022[1]: Burundi, Somalia, South Sudan.

[1] "https://www.concernusa.org/story/worlds-hungriest-countries/" 2022.

Rainfall data sets of 1905 – 2021 are downloaded from

`https://climateknowledgeportal.worldbank.org`

Table 1: Average monthly rainfall (mm) distribution for Kenya, Somalia, Madagascar, and Zambia.

Date	Zambia rainfall	Kenya rainfall	Madagascar rainfall	Somalia rainfall
1905-01-01	223.38	26.67	293.73	2.75
1905-02-01	230.67	9.72	297.54	2.45
1905-03-01	122.23	157.43	175.12	37.31
1905-04-01	24.43	128.27	124.61	59.17
1905-05-01	8.4	90.34	59.14	61.3
1905-06-01	0.21	28.37	36.52	12.09
1905-07-01	0.14	24.16	53.16	16.23
1905-08-01	0.21	33.97	62.2	8.08
1905-09-01	2.95	25.97	37.69	27.27
1905-10-01	14.47	69.94	51.49	39.73
1905-11-01	107.1	112.03	185.03	44.67
1905-12-01	188.26	50.2	238.58	6.77
1906-01-01	235.07	10.83	200.17	2.62

Data Acquisition 2

Temperature data sets of 1905 – 2021 are downloaded from

`https://climateknowledgeportal.worldbank.org`

Table 2: Monthly mean temperature (°C) distribution for Cameroon, Burundi, Rwanda, and Tunisia.

Date	Cameroon temperature	Burundi temperature	Rwanda temperature	Tunisia temperature
1905-01-01	24.12	19.97	18.65	8.22
1905-02-01	25.52	19.91	18.58	9.46
1905-03-01	26.55	19.57	18.3	14.31
1905-04-01	26.64	19.08	17.76	18.27
1905-05-01	25.69	19.61	18.39	20.37
1905-06-01	24.08	19.51	18.53	25.18
1905-07-01	23.22	18.84	17.86	28.09
1905-08-01	23.58	19.94	18.66	28.38
1905-09-01	23.58	20.83	19.03	26.06
1905-10-01	24.18	21.18	19.25	19.39
1905-11-01	24.15	19.85	18.36	15.81
1905-12-01	23.6	19.16	17.78	11
1906-01-01	23.75	19.24	17.8	10.15

Problem Statement: Motivation

How can we **reduce the impacts** of climate change (in agriculture)?

Detecting climate change early

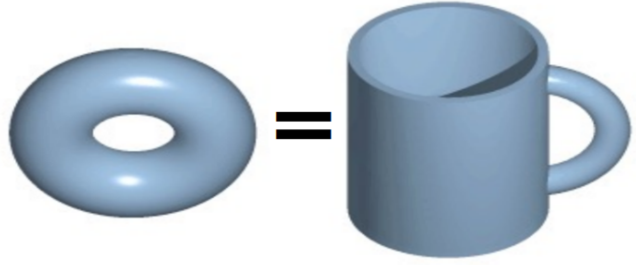
- Allows us to predict changes in the growing seasons.
- Can provide an early warning signal for upcoming dangerous phenomena (droughts, floods, volcanoes, etc).

Research Question

Can **topological features** be a reliable tool for **detecting** climate change?

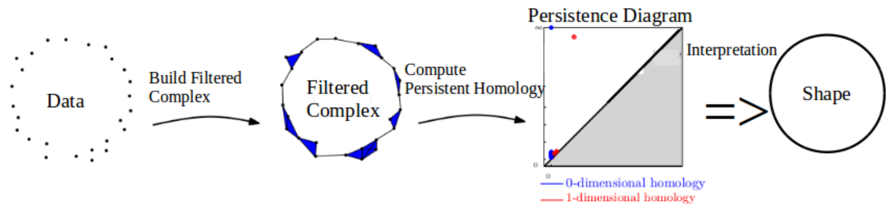
Background

What are Topological Features?



- **Topological features** (connectivity, holes) are properties that are invariant under continuous deformations.
- **Homology groups** are algebraic tools to quantify topological features in data. Data here are abstract mathematical objects.

Persistent homology is an algebraic method for discerning **topological features** (connectivity, holes) in data.

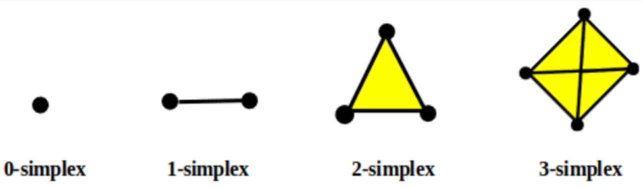


Topological features are k-dimensional homology

- For $k = 0$, it counts the number of **connected components**;
- For $k = 1$, it counts the number of 1**-dimensional holes**;
- For any k, it counts the number of k**-dimensional holes.**

Simplices

A k-simplex is a convex hull of $k+1$ affinely independent points (in $\mathbb{R}^k$).



Simplicial Complexes

A simplicial complex is an object built by gluing together simplices.

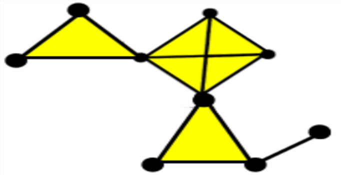
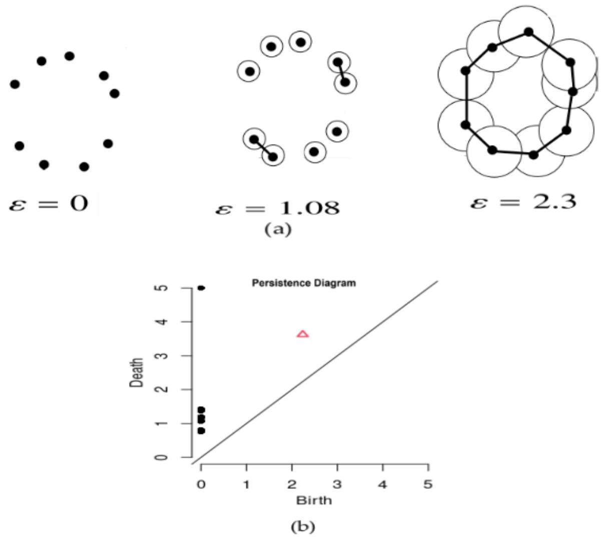


Figure 1: A simplicial complex.

(a) Construction of a Filtered Complex with ε and (b) Persistence Diagram



Extracting the Topological Features

Persistence Landscape

- Persistence Landscape proposed by Peter Bubenik is a real-valued function suitable for ML.
- Intuitively rotate persistence diagram by $\pi/4$, and create a functional representation $\lambda : \mathbb{N} \times \mathbb{R} \to \mathbb{R}$, with $\lambda(k, t) = k \max_p \Lambda_p(t)$ where $\Lambda_p(t)$ are linear functions.

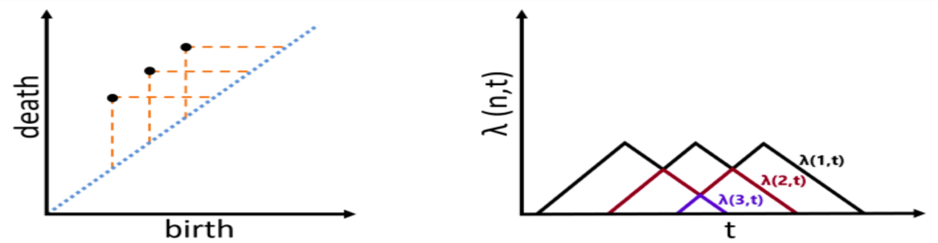
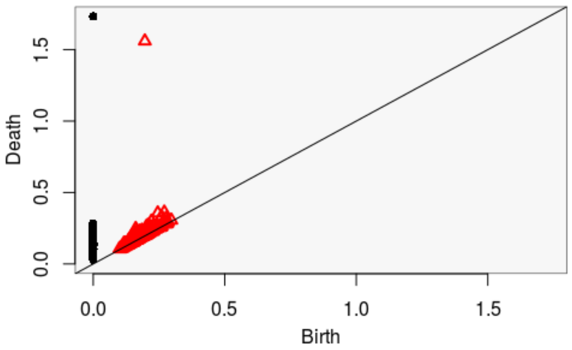


Figure 2: Persistence landscape (right) associated to a persistence diagram (left).

Score of Periodicity



- Use the 1-dimensional homology that persists the most $(b, d) = (0.196, 1.559)$.
- The score

$$S = 1 - \frac{d^2 - b^2}{3} = 0.203;$$

- The time series data is periodic because this score is close to zero.

Topological features have been used in signal processing, image analysis, physics, chemistry, biology, material science, finance, neuroscience, climate science[2] and so on.

Topological Features in Climate Science

- Identify atmospheric river patterns in large climate data.
- Analyze the temperature field over tropical cyclones.
- Preprocessing step to achieve flood early warning systems through critical slowing down.

Research Gap
The usage of **topological features** in detecting and characterizing climate change has not been employed before to the very best of our knowledge.

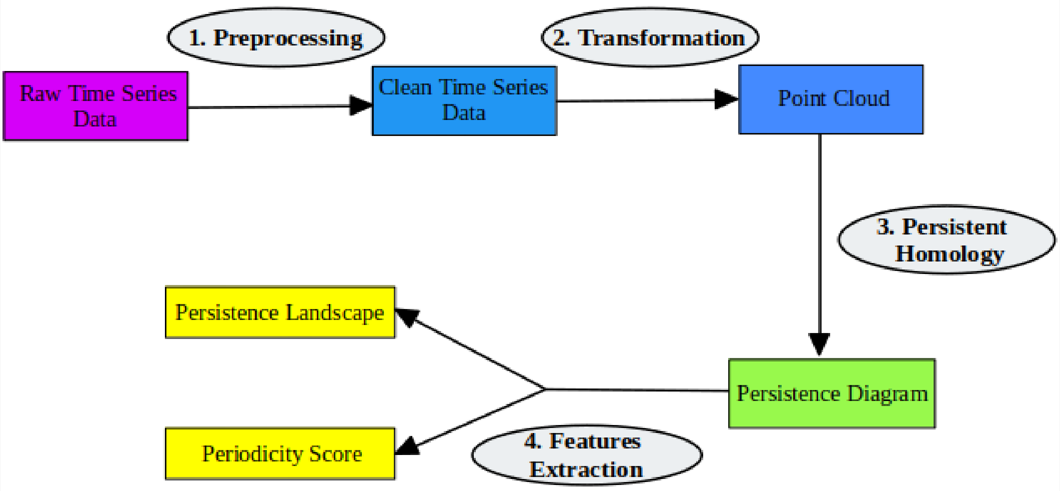
[2] Emrani, Gentimis, and Krim, 2014; Qaiser et al., 2019.

Methodology

Hypothesis

- Climate data is expected to be **periodic** and a move away from periodicity is evidence of change in climate.
- When a country is undergoing strong climate change, then the weather patterns are no longer (annually) **periodic**.
- We use **topological features** through persistent homology to measure this **periodicity**.

Methodology



Data Transformation: Taken's Embedding Theorem

- Our time series $x_1, x_2, \cdots, x_N$ is a 1-dimensional dataset, continuing climate data from 1905 to 2021.
- Let τ be the **time delay** and d be the **embedding dimension**.
- The time series will be converted into vectors

$$v_i = \left[x_i, x_{i+\tau}, \cdots, x_{i+(d-1)\tau}\right]^\intercal \in \mathbb{R}^d \tag{1}$$

$i = 0, \tau, 2\tau, \cdots, 2\pi - (d-1)\tau$.

- The two parameters τ and d require careful selection.
- For this project $d = 15$ and $\tau = \frac{2\pi}{N+d-1}$.

Results and Future Work

Detecting Change in Precipitation

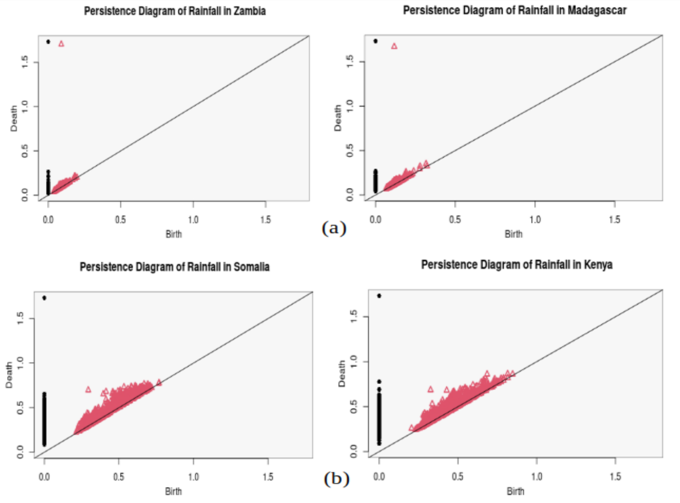


Figure 3: (a) Persistence diagram for countries with no strong change in precipitation (Zambia and Madagascar). (b) Persistence Diagram for countries with a strong climate change (Somalia and Kenya).

Table 3: Score of periodicity for precipitations data

Country	Score	Comment
Madagascar	0.0764	No change in precipitation
Somalia	0.867	Strong climate change
Zambia	0.0351	No change in precipitation
Kenya	0.879	Strong climate change

Detecting Change in Temperature

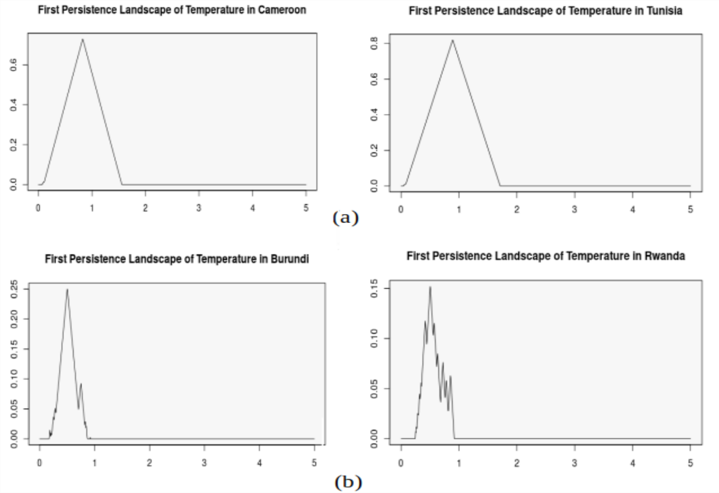


Figure 4: (a) Persistence landscape for countries with no strong change in temperature (Cameroon and Tunisia). (b) Persistence Landscape for countries with a strong climate change (Burundi and Rwanda).

Table 4: Score of periodicity for temperatures data

Country	Score	Comment
Cameroon	0.193	No change in temperature
Rwanda	0.897	Strong climate change
Tunisia	0.026	No change in temperature
Burundi	0.831	Strong climate change

Work in Progress: New Framework

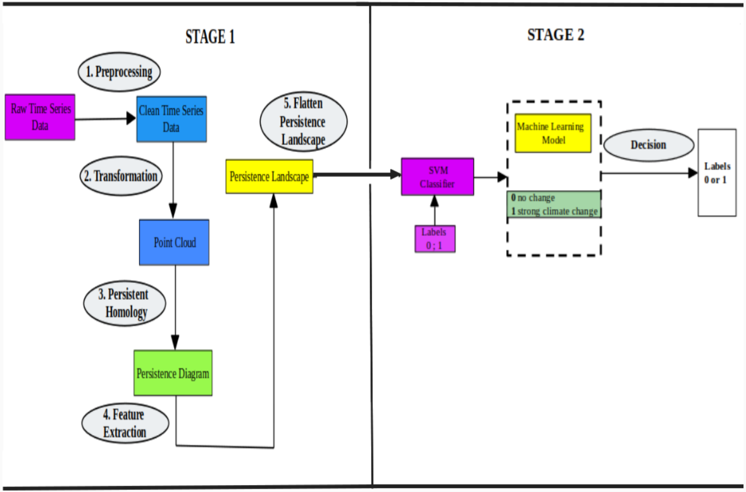


Figure 5: The flowchart of two stages of the climate change model.

Strategy

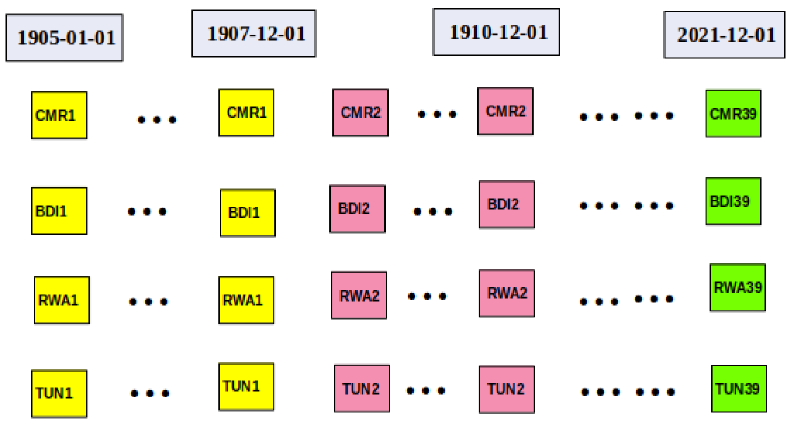


Figure 6: Subdivision of temporal data of temperature over 1905-2021.

Preliminary Result

Table 1: Results for Training Set.

ACC.	Sen	Specificity	Precis.(0)	Precis.(1)	recall	F1-score
0.99	0.98	1.00	1.00	0.99	0.99	0.99

Table 2: Results for Testing Set.

ACC.	Sen	Specificity	Precis.(0)	Precis.(1)	recall	F1-score
0.98	0.98	0.97	0.97	0.99	0.99	0.98

Challenge: How to Label the Dataset?

X_train	Y_train
CMR1 ... CMR1 ⋮ ⋮ CMR39 ... CMR39	1 ⋮ 1
BDI1 ... BDI1 ⋮ ⋮ BDI39 ... BDI39	0 ⋮ 0
RWA1 ... RWA1 ⋮ ⋮ RWA39 ... RWA39	0 ⋮ 0
TUN1 ... TUN1 ⋮ ⋮ TUN39 ... TUN39	1 ⋮ 1

- Model which to learn f such that:

$$Y\_train = f(X\_train) \tag{2}$$

- New input data X, the models can predict output variable

$$Y = f(X) \tag{3}$$

Conclusions

Conclusion

- **Topological features** offers a new perspective on deciphering the non-linear connection in climate science.
- In climate science, we proposed a new and innovative framework for **detecting climate change**.
- We proposed a new technique for **feature extraction** of weather patterns in climate data.
- The proposed method is **coordinate-free** and **threshold-free.**
- In Agriculture, the proposed model can be used as an **early warning system** for detecting local **changes in growing seasons.**
- For policy makers, the proposed model may be used to **prevent disasters** (droughts, storms, volcanoes, etc).

Outcomes: Papers

1. Peguy K. T. K., Adrien N. M., Armandine S. K. M., Rosly G. M.: Topological Characterization of Climate Change in Selected African Countries using Persistent Homology
   Paper draft
2. Peguy K. T. K., Armandine S. K. M.: Persistent Homology and Machine Learning for Detecting Climate Change
   Paper draft in progress

Acknowledgements

- The University of Liverpool for given us the opportunity to present our work.
- International Development Research Centre (IDRC).
- Global Affairs Canada (GAC).
- Mathematical Sciences for Climate Change Resilience (MS4CR).
- African Institute for Mathematical Sciences (AIMS NEI).
- Prof. Henry Adams for useful discussions during the preparation of this project.

References

- Emrani, Saba, Thanos Gentimis, and Hamid Krim (2014). “Persistent homology of delay embeddings and its application to wheeze detection”. In: *IEEE Signal Processing Letters* 21.4, pp. 459–463.
- Hoef, Lander Ver et al. (2022). “A primer on topological data analysis to support image analysis tasks in environmental science”. In: *arXiv preprint arXiv:2207.10552*.
- “https://www.concernusa.org/story/worlds-hungriest-countries/” (2022). In.
- Qaiser, Talha et al. (2019). “Fast and accurate tumor segmentation of histology images using persistent homology and deep convolutional features”. In: *Medical image analysis* 55, pp. 1–14.

"Data has shape and shape matters." – Gunnar Carlsson

pkadzue@quantumleapafrica.org