# Topological Data Analysis in Climate Change Toward PG Programs and Research at UPNG Including Climate Change Research

Lakoa Fitina
Jeffery Ambelye
Jimmy Keake

University of Papua New Guinea

October 12, 2022

# Current Situation

## Program

1. THE division of Mathematics, Statistics and Computer Science (MSCS) at UPNG offers the BSc degree in Mathematics, Statistics and Computer Science
2. There are 4 strands in the Division

# Current Situation

## Program

1. THE division of Mathematics, Statistics and Computer Science (MSCS) at UPNG offers the BSc degree in Mathematics, Statistics and Computer Science
2. There are 4 strands in the Division
   - Pure Mathematics
   - Applied Mathematics
   - Statistics
   - Computer Science

# Current Situation

## Courses

1. Most courses offered are more than 10 years old, some are more than 15 years old.
2. Some changes need to be made, in order to strengthen, as well as modernize the course offering
3. Thus, new courses are being proposed in Pure Mathematics, Applied Mathematics and Computer Science

# Current Situation

## Research

1. Before 2021, there was little to no research
2. In 2021 MSCS started forming research groups, in the following areas:
   - Finite topology and its applications
   - Hypergraphs/graphs and their applications
   - Cryptography
   - Quantum Computing
   - Topological Data Science

# Current Situation

## Research

1. For research to develop properly staff need to be trained in both the theory and the application of that theory
2. Staff need to be brought to a level, where they can ask questions
3. Staff need to be brought to a level where they can contribute to answering those questions
4. Staff need to be introduced to research that is current and that has relevance to their lives
5. One such area is CLIMATE CHANGE, using methods derived from TOPOLOGY

# What is TOPOLOGY ?

## Topology and Data

1. Topology is an area in Pure Mathematics
2. Topology is the study of SHAPE
3. Data has shape
4. Hence Topology would be a good area to start if you want to study the shape of a data set.

# What is TOPOLOGY ?

## Definition

## Definition

Suppose that $X$ is a non empty set. A collection $\tau$ of subsets of $X$ is a **topology** on $X$, if the following axioms are satisfied:


# What is TOPOLOGY ?

## Definition

## Definition

Suppose that $X$ is a non empty set. A collection $\tau$ of subsets of $X$ is a **topology** on $X$, if the following axioms are satisfied:

- $X \in \tau$
- $\emptyset \in \tau$
- If $\{O_i\}$ is an arbitrary collection of subsets of $X$ in $\tau$, then $\cup O_i \in \tau$
- The intersection of any finite number of subsets of $X$ in $\tau$, is also in $\tau$.

If $\tau$ is a topology on $X$ then the pair $(X, \tau)$ is called a **topological space**. Sets in $\tau$ are called **open sets**.

# What is TOPOLOGY ?

## Examples 1

### Example

Some examples of topological spaces include:

1. $X = \{1, 2, 3\}$, $\tau = \{X, \emptyset, \{1, 2\}, \{3\}\}$
2. Suppose $X = \mathbb{R}$ is the set of all real numbers. Define a set $O$ to be 'open', if for each element $p \in X$, there is an open interval $(a, b)$ with $a < p < b$ such that $p \in (a, b) \subset X$. That is $O$ is open if every element of $O$ is inside an open interval that is contained in $O$. Then the set of all such open sets form a topology on $\mathbb{R}$.

# What is TOPOLOGY ?

## Examples 1

A special example of a topological space, is what is called a metric space: these spaces provide topological spaces with the notion of 'distance' between two points in the space.

## Definition

Let $X$ be a non empty set. A **metric** on $X$ is a function

$$d : X \times X \to \mathbb{R}$$

such that:

# What is TOPOLOGY ?

## Examples 1

A special example of a topological space, is what is called a metric space: these spaces provide topological spaces with the notion of 'distance' between two points in the space.

## Definition

Let $X$ be a non empty set. A **metric** on $X$ is a function

$$d : X \times X \to \mathbb{R}$$

such that:

1. $d(x, y) \geq 0$ for any two elements $x, y$ in $X$
2. $d(x, y) = o \iff x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, y) \leq d(x, z) + d(z, y)$

The pair $(X, d)$ is called a **metric space**.

# What is TOPOLOGY ?

## Metric in $\mathbb{R}$

1. We can define a metric on $\mathbb{R}$ by $d(x, y) = |x - y|$.
2. That is, $d(x, y)$ is the 'usual' distance between any two real numbers.
3. This distance has all the properties required of a metric, so $(\mathbb{R}, d)$ is in fact a metric space.
4. Having the notion of 'distance' between any two points in a space, turns out to be very important in topological data analysis.

# What is TOPOLOGY ?

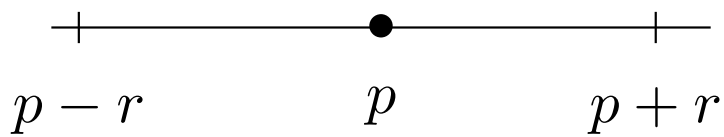
## $\mathbb{R}$ as a Topological Space

### Definition

1. Given a point $p$ in $\mathbb{R}$, we define the open interval with center $p$ and radius $r > 0$, as

$$I_r(p) = \{x \in \mathbb{R} : d(p, x) < r\}$$

2. Note that $I_r(p)$ is the open interval $(p - r, p + r)$.




3. Define a set $O$ of real numbers to be 'open', if for each point $p$ in $O$, there is an open interval $I_r(p)$ whose center is $p$, that lies inside $O$
4. Then the set of all such open sets forms a topology on $\mathbb{R}$.

# What is TOPOLOGY ?

## Generalizing to $\mathbb{R}^n, n = 1, 2, \cdots$

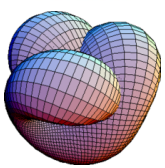
### Definition

1. From $\mathbb{R}$, we can generalize to any $\mathbb{R}^n, n = 1, 2, 3, \cdots$
2. Each of these is a topological space

# What is TOPOLOGY ?

## Some other topological spaces

## Example

| **Object** | **Graphic** |
|---|---|
| line segment |  |
| circle |  |
| simplicial complex |  |
| doughnut |  |
| manifold |  |

# Topological Data Analysis

## TDA

### Definition

Topological Data Analysis (TDA) is an approach to the analysis of data sets using techniques from topology.

1. Metrics can be used to study certain aspects of data sets. Another method that is often used is the idea of a complex.
2. Complexes are useful when it is not easy to define a metric on a space
3. Complexes may also be used to complement the use of metrics

# Topological Data Analysis

## Some simplicial complexes

A **complex** is obtained by joining points (say, in a data set) to form geometric objects.

1-simplex = line segment
2-simplex = triangle
3-simplex = tetrahedron

Each complex above is a topological space.
There are of course, other topological methods that can be used, but we do not have the time to discuss these.

# Topological Data Analysis

## Where TDA has been used in the recent past

1. climate data : classification of *atmospheric rivers*
2. materials science
3. shape analysis
4. image analysis
5. multivariate time series analysis
6. medicine
7. biology
8. genomics
9. chemistry
10. sensor networks
11. identified a subgroup of breast cancers
12. discover phenotype-biomarker associations in traumatic brain injury

# Topological Data Analysis

## TDA on Climate Change

1. Classify atmospheric rivers
2. Understanding the shape of climate attractors : certain regions in a 'climate space'
3. Determining flow patterns
4. Understanding the 'butterfly effect'
5. Predicting when earth's climate will 'tip': 'tipping' is when a temperature threshold is passed, leading to unstoppable change in a climate system

# How to get into TDA research with MSCS

## TDA in MSCS

BSC $\rightarrow$ Honors $\rightarrow$ Masters $\rightarrow$ PhD

Along the way, do courses in Statistics, Group theory, Topology.

# Topology and Statistics

## Two is better than one

Finally, we don't claim that TOPOLOGY will give you all the answers. But we think that a combination of Topology and Statistics will give insight into data that has not previously been realized.

# How to get into TDA research with MSCS

## TDA in MSCS

BSC $\rightarrow$ Honors $\rightarrow$ Masters $\rightarrow$ PhD

Along the way, do courses in Statistics, Group theory, Topology.

# THANKS for Listening

**THANK YOU**