National Aeronautics and Space Administration

Topological Data Analysis based Deep Learning for Outlier Detection

Principal Investigator: Huikyo Lee (398); Co-Investigators: Michael Garay (329), Krzysztof Gorski (326), Matthew Dixon (Illinois Institute of Technology), Yulia Gel (University of Texas at Dallas)

Program: FY21 R&TD Innovative Spontaneous Concepts

Objectives

Our proposed study will apply TDA to develop new capabilities for identifying and understanding the spatio-temporal variability of anomalous near-surface air temperature events. Our specific objectives are:

1. Combine Zigzag persistent representations of temperature observations from NASA's reanalysis with deep learning (DL) for anomaly detection.

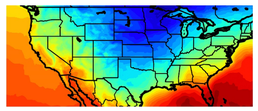
2. Develop novel topological clustering algorithms to group certain spatio-temporal regions in sequences of temperature maps from multiple instruments with similar spatiotemporal structures - with minority clusters representing multi-scale anomalies in space and time.

Background

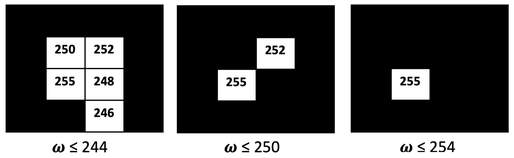
Topology is the study of shape. TDA characterize the shape of n-dimensional point cloud data (i.e. data properties invariant under stretching, bending and rotation).

Approach and Results

Cubical complex filtration

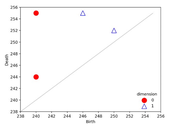


240	242	244	240
240	250	252	242
242	255	248	242
244	244	246	244

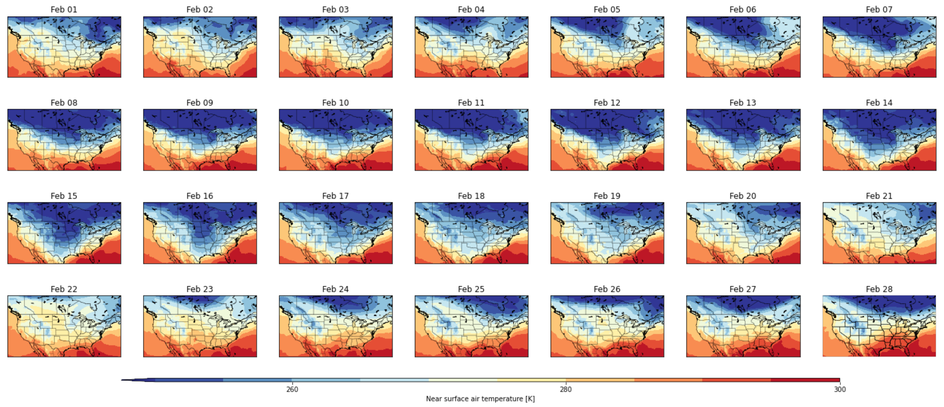


$\omega \leq 244$ $\omega \leq 250$ $\omega \leq 254$

240	242	244	240
240	250	252	242
242	255	248	242
244	244	246	244

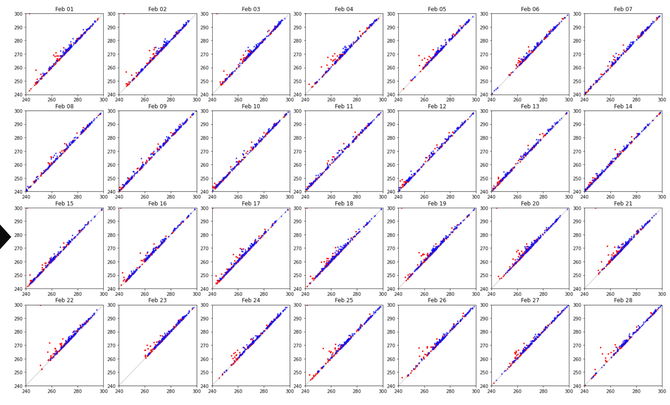


- To study latent shape in temperature maps, we work with the TDA tool known as persistent homology (PH).
  - PH: track birth/deaths and lifespans of topological features at varying temperature thresholds
- The current study captures the topological features of daily temperature maps based on the formation of cubes (cubical complex).



Daily mean temperature from the Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2) in February 2021.

TDA



Persistence diagrams reflecting the lifespan of connected components and holes for the daily temperature maps in February 2021.

Significance/Benefits to JPL and NASA This collaborative project aims at infusing novel statistical and machine learning topological approaches into DL exploration of Earth Science observations from NASA missions.

Publications Ofori-Boateng D, Lee H, Gorski KM, Garay MJ and Gel YR (2021) Application of Topological Data Analysis to Multi-Resolution Matching of Aerosol Optical Depth Maps. Front. Environ. Sci. 9:684716. doi: 10.3389/fenvs.2021.684716

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

www.nasa.gov

PI/Task Mgr Contact
Email: Huikyo.Lee@jpl.nasa.gov

Clearance Number:
RPC/JPL Task Number: R21262