# Predicting Flood Zones Using Topological Data Analysis

Kyle Hess, Marty Roger

Mason Experimental Geometry Lab

George Mason University

5/2/2025

## Introduction: Predicting Flood Risk

Modern flood risk assessment strategies tend to rely heavily on hydrodynamic simulations of water's movement.
We seek to provide qualitatively new insights by instead performing topological analysis on ground contours.
Using Morse Theory and Connection Matrices, we will produce a predictive model that saves compute time over existing methods while providing a new perspective on flooding.

## Mathematical Problem Framing

The data describing a patch of terrain is given as an evenly spaced grid of points (that is, as a 'raster'). We refer to this as a height function, which for some integers $(x, y)$ in $\mathbb{R}^2$ returns a number, representing the height of the terrain at that location.
We then triangulate the domain, forming a simplicial complex, which we use together with the gradient of the terrain to create a directed graph.

## Definition (Multivector Field)

A multivector is a 'locally closed' set; one that is the set difference of two closed sets.
A multivector field is a partition of the underlying complex into locally closed sets.

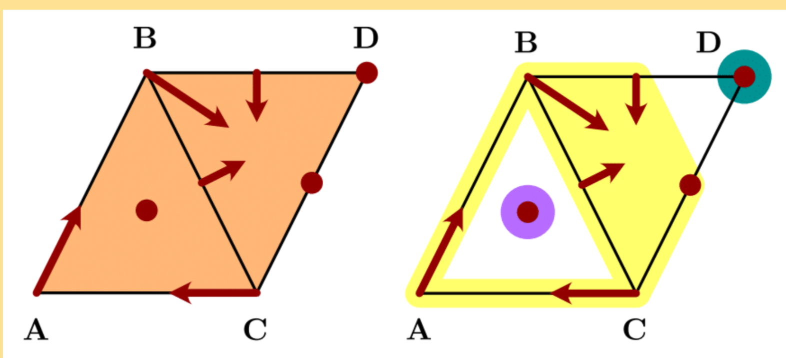


Figure: Decomposition of a multivector fild into morse sets.

Using the ConleyDynamics library, we can create our multivector field, find its Morse decomposition, and compute its connection matrix.

## Definition (Morse Decomposition)

Morse sets represent regions where critical points are guaranteed to exist, like 'hills' or 'valleys'. A Morse decomposition involves finding the morse sets of a multivector field.

## Definition (Connection Matrix)

ConleyDynamics also gives us the connection matrices, which describe flow between morse sets (i.e.– which hills flow to which valleys).

The goal of our analysis is to use the powerful connection matrix to topologically classify and describe flood zones.

## A Familiar Test Case?

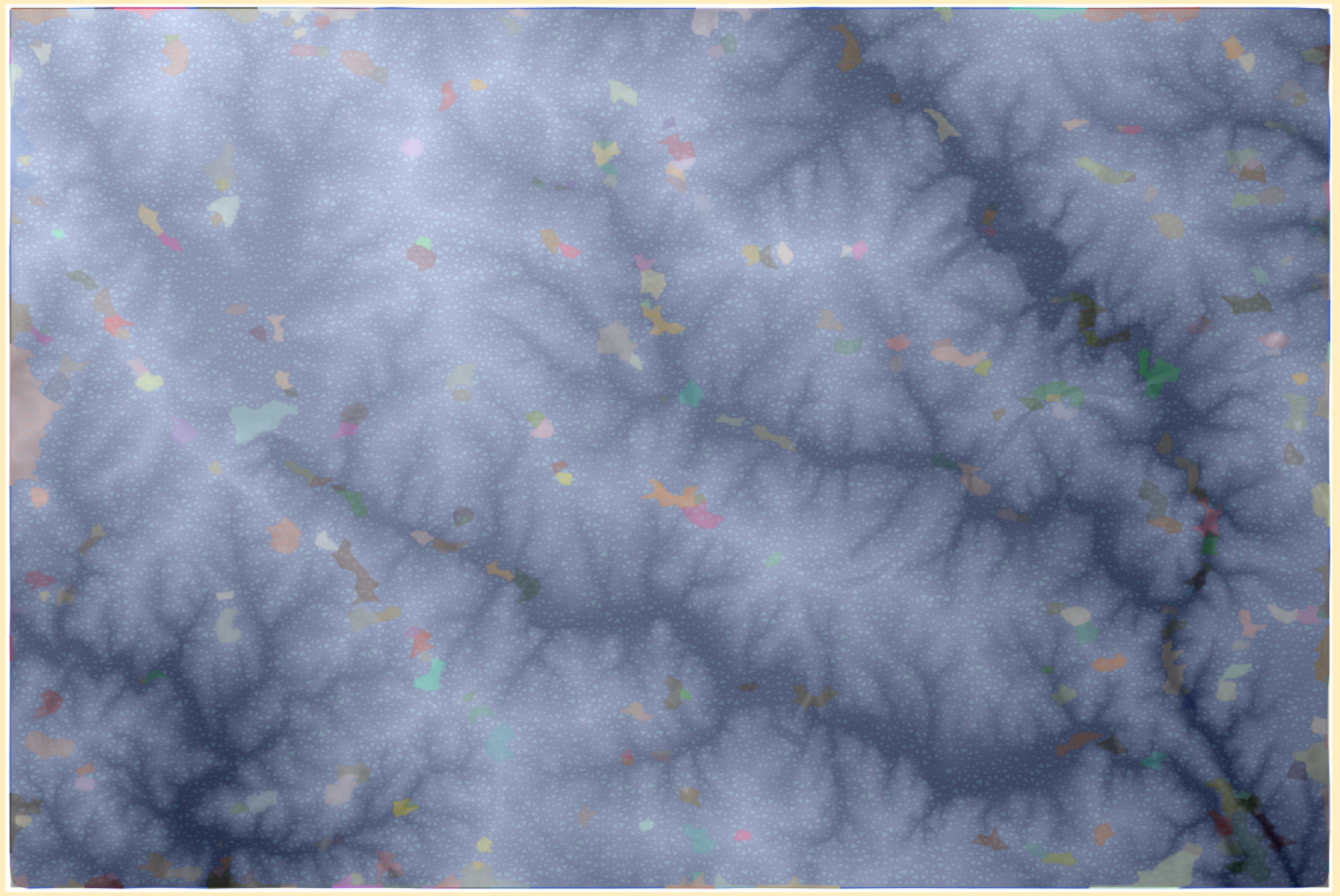

Figure: A topographic image of Burke Lake, overlaid with collection of Morse sets representing critical points in its topology.

## Morse sets and critical points

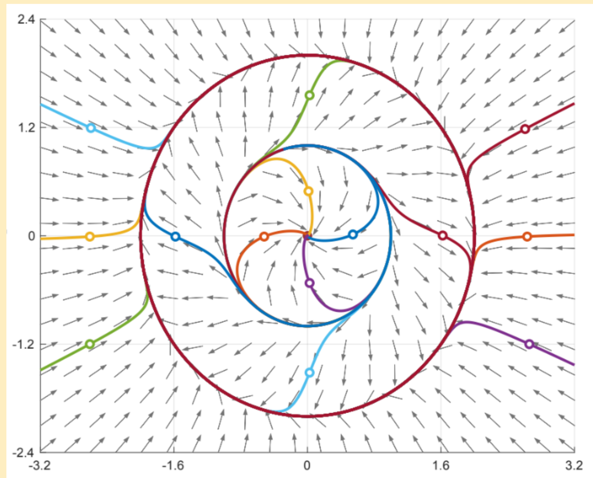

Figure: The phase portrait of a planar ordinary differential equation with two circular periodic orbits and an asymptotically stable origin.

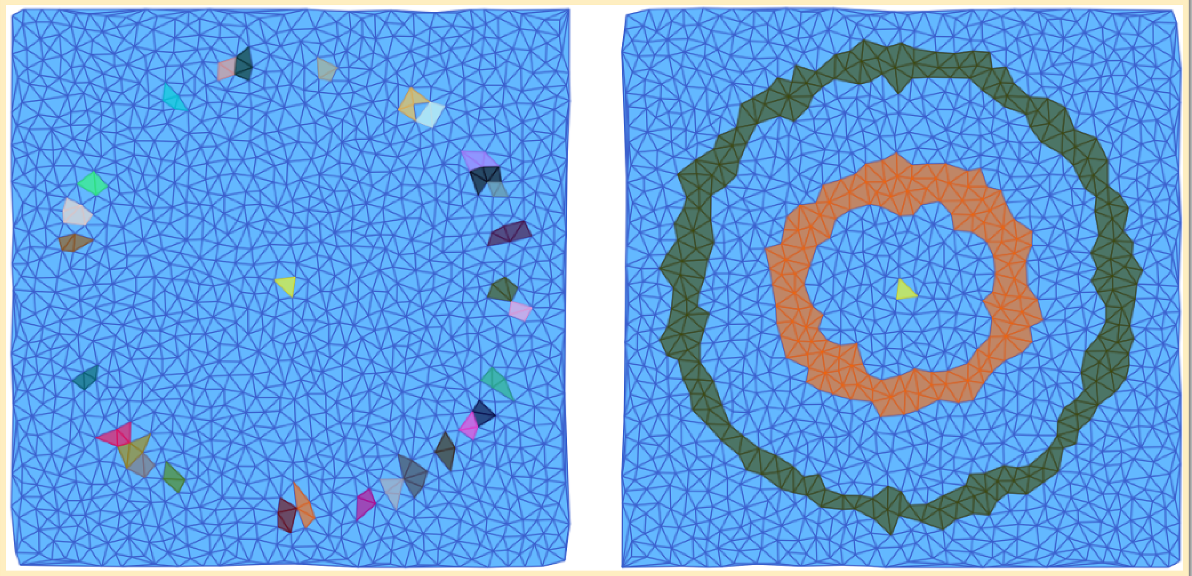

Figure: Left: A multivector field over the above ODE. Highlighted are multivectors of size 7 or greater, while the average size is 2-4.
Right: Its Morse decomposition containing three morse sets, one containing each unique equilibrium solution.

## What is GIS?

Geographic Information Systems is a field of study concerning the representation, analysis, and visualization of the planet. Problems long solved in this field include accounting for the irregular curvature of the earth, projecting between different images of the globe, and efficient formats for files and data.

Figure: Rasterized height data, visualized in grayscale, with brightness corresponding to elevation.

## Adapting to working with GIS tools

Our definitions and methods of construction rely on a planar coordinate system to represent the subject area. However, that data is point-like, and must be projected into space before it is meaningful.
For this task, a model called a 'datum' represents the size and shape of the earth, enabling projection of geographic coordinates into $\mathbb{R}^2$ for analysis, then vice versa.

## Difficulties of the real world and managing the data

One problem we are facing is false positives as a result of roughness in the data. Too much local variation in the height gradient will result in trivial local minima between high and low points. Here, Morse sets will be found where the topography is not actually large enough in scale to impact flooding.
Another issue is the process that constructs the multivector partition can create multivectors that are too large, complicating fine-grained analysis.

## Future work

The next crucial step of our analysis is finding a way to tune the sensitivity of our findings. Altering the parameters of the triangulation, performing data smoothing, and adding search and refinement heuristics will all affect which topological features are highlighted as Morse sets.
More work is needed to verify the accuracy and correctness of our results. Right now, it is difficult to tell the degree to which our model agrees with the reality. We are looking to do a case study where the result of a flood is known, and can be compared to our analysis. However, it is difficult to capture the flow of a real flood accurately and precisely for comparison.
Finally, it would be very interesting to associate some aspects of more traditional fluid simulation to this model. With relatively minor modifications, it would be possible to quantify the area upstream of a given point, which would provide some interesting comparisons with our current model.

## Acknowledgments

We would like to thank Anton Lukyanenko and Swan Klein and all others who run and support MEGL.
Additionally, we would like to thank our graduate mentor, Frank Pryor, and our faculty mentors, Dr. Sander and Dr. Wanner.

## References

- Apel, H., Thieken, A. H., Merz, B., & Bloschl, G. (2024).Flood Risk Assessment and Associated Uncertainty.
- Mrozek, M., & Wanner, T. (2025). Connection matrices in combinatorial topological dynamics.
- Plate, E. J. (2002).Flood risk and flood management. *Journal of Hydrology, 267*(1), 2–11.
- Tsakiris, G. (2014).Flood risk assessment: Concepts, modelling, applications. *Natural Hazards and Earth System Sciences, 14*(5), 1361–1369.