A Computationally Efficient Adaptive Sampling Framework for Localized TDA of Earthworks in a DDDAS Framework

Mohsen Gol Zardian[1], A Q M Zohuruzzaman[2], Weicong Feng[3], Austin R.J. Downey[1,4], Jie Wei[3], Sadik Khan[2], Paul T. Schrader[5], Erik Blasch[5,6], Erika Ardiles-Cruz[5], Jason Bakos[7], and Jasim Imran[4]

[1]Mechanical Engineering, University of South Carolina, Columbia, South Carolina, USA
[2]Civil and Environmental Engineering, Jackson State University, Jackson, Mississippi, USA
[3]Computer Science, City College of New York, New York City, New York, USA
[4]Civil and Environmental Engineering, University of South Carolina, Columbia, South Carolina, USA
[5]Air Force Research Laboratory, Rome, New York, USA
[6]MOVEJ Analytics, Dayton, Ohio, USA
[7]Computer Science and Engineering, University of South Carolina, Columbia, South Carolina, USA



The City College of New York

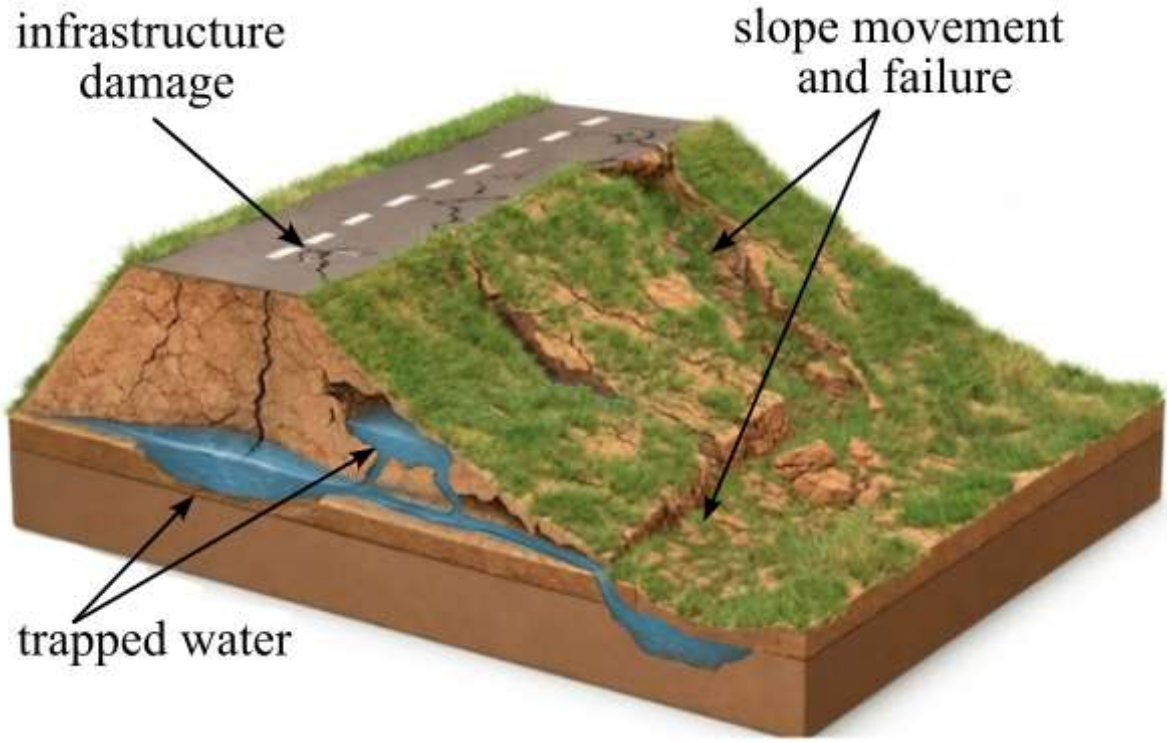






Outline

- Shifting Rainfall Patterns
- Expansive Clay Slopes
- Topological Monitoring of Embankments (Prior Work)
- Localized TDA for Adaptive Sampling (Current Study)



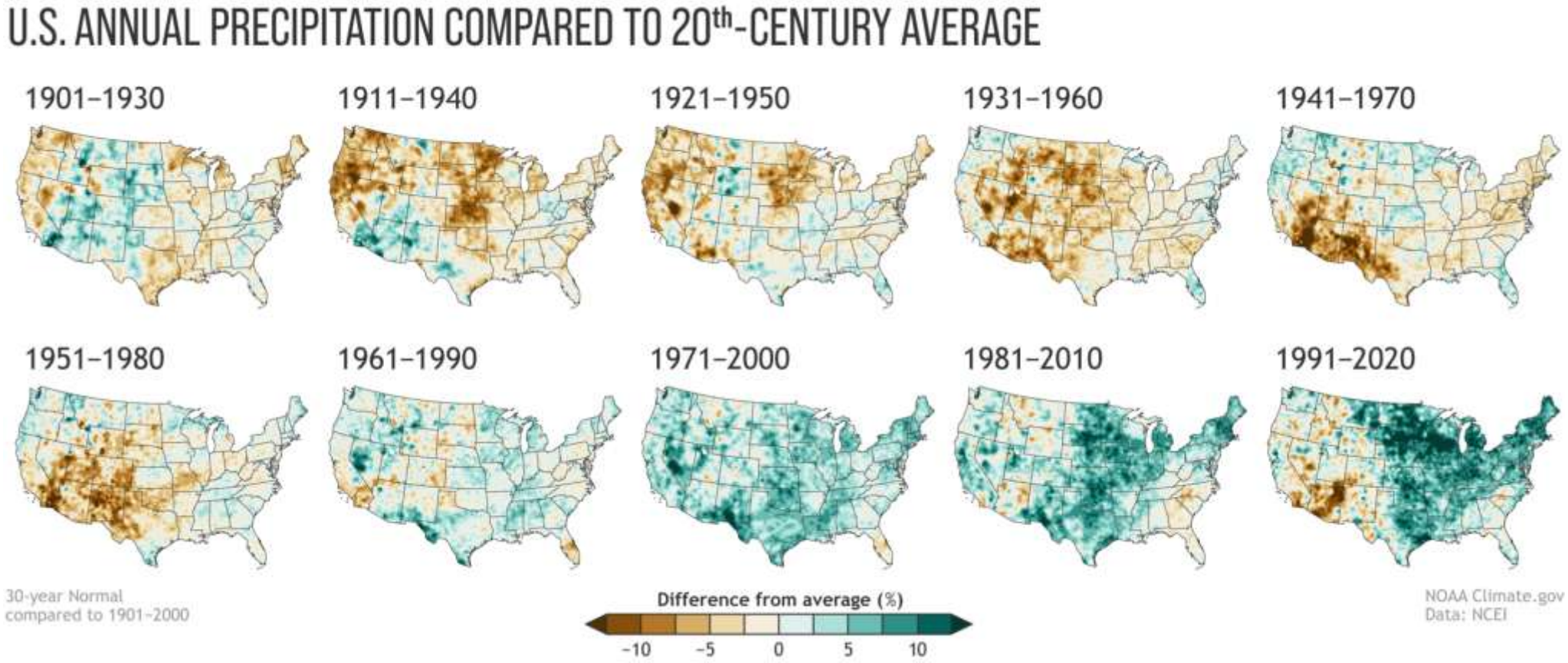
Shifting Rainfall Patterns

The Eastern U.S. Is Getting Wetter

- Annual rainfall over the last 30 years compared with the 20th-century average
- Much of the eastern half of the U.S. gets 2–4+ more inches of rain per year
- The largest gains run through the Midwest, Ohio Valley, Northeast and lower Mississippi Valley
- Much of the West and Southwest is drier over the same period
- Mississippi sits in the wetter half, where expansive clays are common
- More water each year raises the baseline for soil swelling and softening

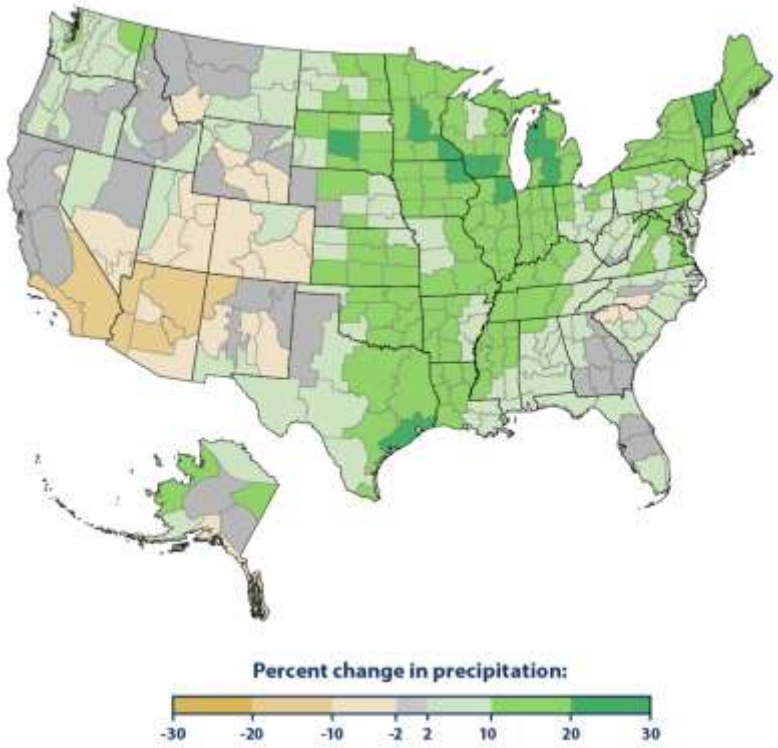


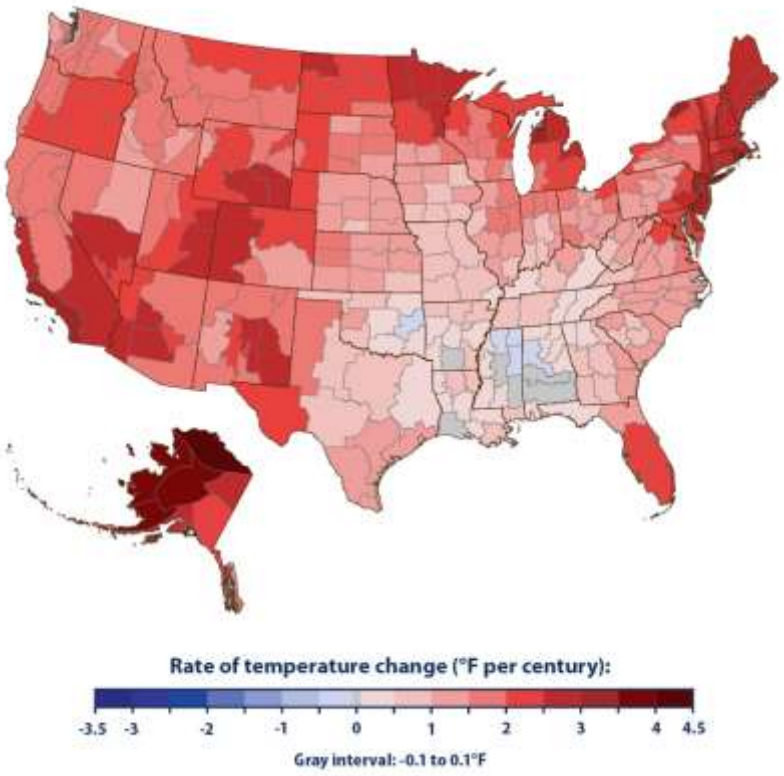
The Shift Is Recent and Growing



Rain and Heat Compound

- Rain wets the soil and heat dries it out
- The swing between wet and dry is what matters for slopes
- Each swing drives shrink–swell, cracking and heave

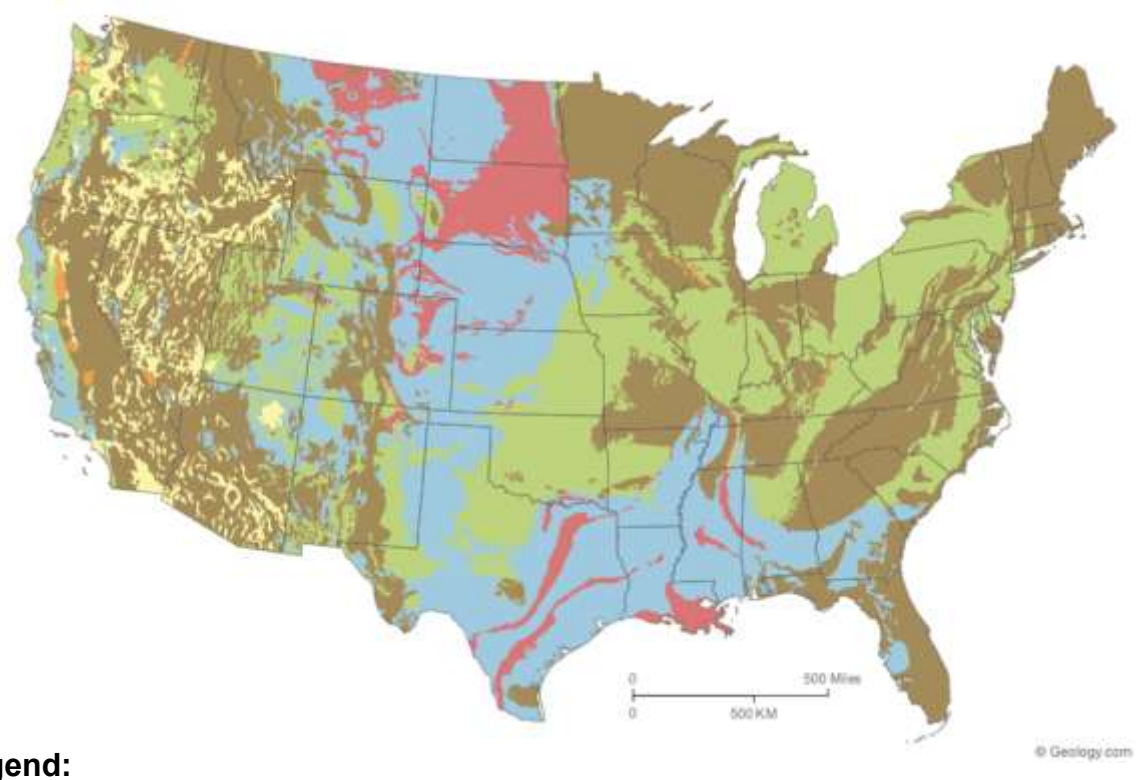




Expansive Clay Slopes

Yazoo Clay Is a Highly Expansive Soil

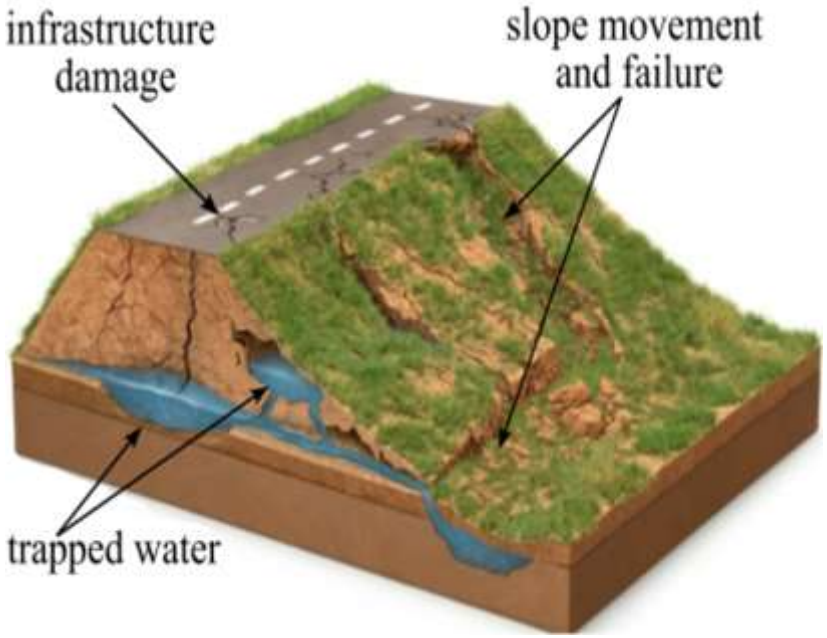
- Expansive clays swell when wet and shrink when dry
- High-swell clays (red) follow the Gulf Coastal Plain and northern Plains
- Yazoo clay of the Jackson Formation runs through central Mississippi
- It can swell up to 300% when wet
- Roads and embankments built on it see repeated heave and settlement

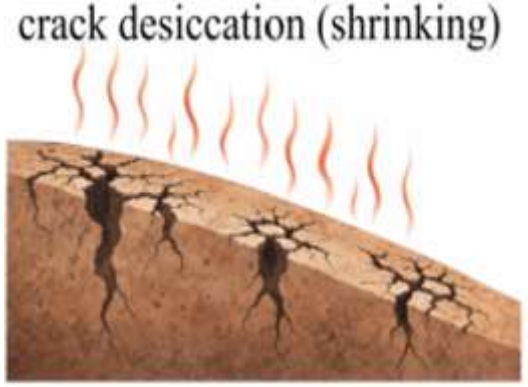


Legend:
Red: Clay having high swelling potential
Blue: Less than 50% of clay contents having high swelling potential
Orange: Clay content having slight to moderate swelling potential
Green: Less than 50% of clay contents having slight to moderate swelling potential
Brown: Little or no swelling clay
Yellow: Insufficient data

Dry Seasons Let the Water In

- Dry spells shrink the clay and open desiccation cracks
- Cracks raise vertical permeability
- Rain flows down the cracks instead of running off
- Water collects in perched zones inside the embankment

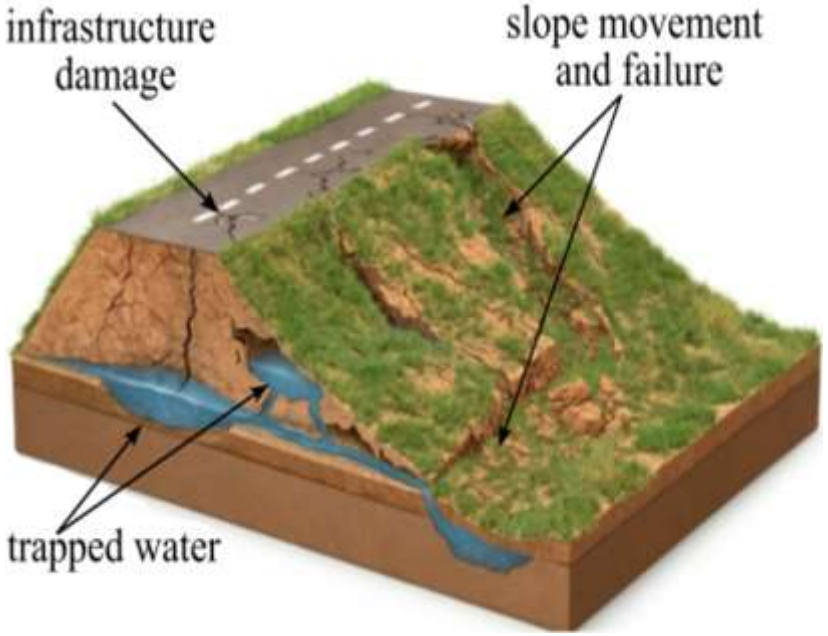




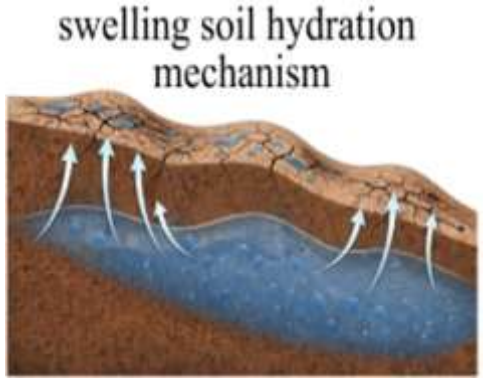


Trapped Water Swells and Weakens the Slope

- Trapped water hydrates the clay and drives swelling
- Swelling heaves the surface and loosens the soil
- Repeated wet–dry cycles reduce near-surface strength
- Shallow slides follow, damaging pavement and shoulders







Shallow Slides on Mississippi Highway Slopes



Clinton,
Mississippi
I-20 West



Madison, Mississippi
I-55 South exit to Sowell Road



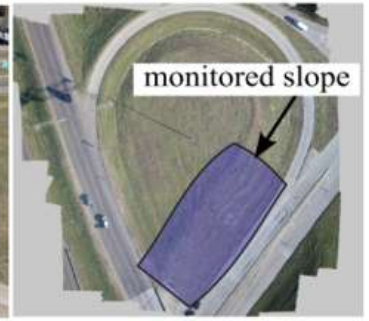
Boundary boxes of the Jackson Formation, including Yazoo clay and its geological equivalents, in Mississippi, Alabama, and Louisiana (after USGS 2010).

The Terry Road Embankment

- Highway embankment at the I-20 exit to Terry Road, Jackson, MS
- 4.5 m (15 ft) high, graded 3.5H:1V to 4H:1V
- Built on Yazoo clay with a history of shallow slides
- Repaired with steel H-piles, leaving reinforced and as-built sections
- Scanned with terrestrial (Trimble X7) and UAV LiDAR







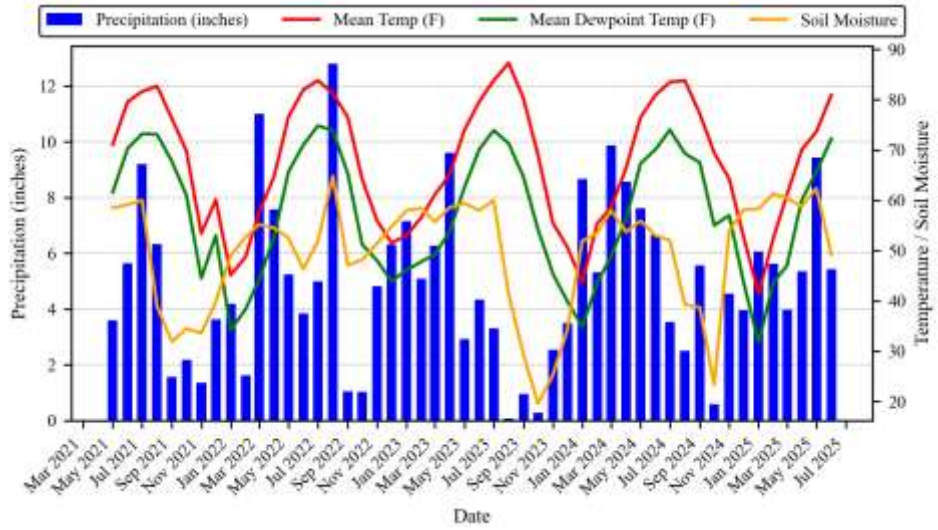






The SLidE Dataset

- Slope LiDAR Embankment (SLidE) dataset, open on GitHub
- Six seasonal scans from June 2021 to September 2023, plus a February 2024 scan
- 5–6 scanner stations and about 20 M raw points per scan
- Registered and merged to about 12 M points per scan
- Paired with monthly rainfall, temperature, and soil moisture records



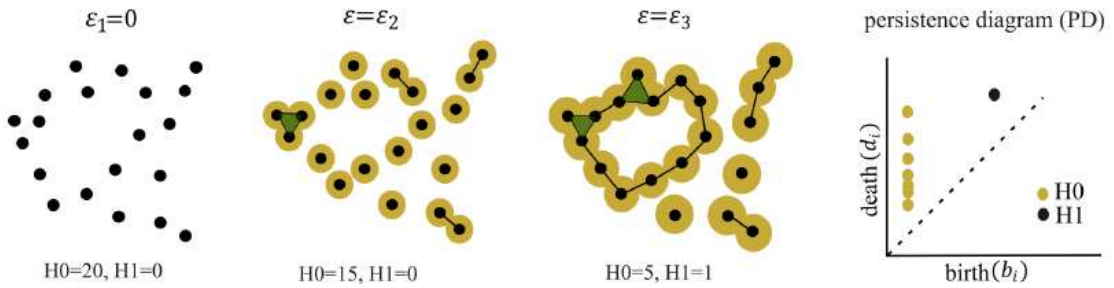


github.com/ARTS-Laboratory/Dataset-SLidE

Topological Monitoring of Embankments (Prior Work)

Persistent Homology Turns Point Clouds into Features

- Growing ε merges points into components (H0) and forms loops (H1)
- The persistence diagram records when each feature is born and dies
- Eight descriptors each for H0 and H1 give 16 features per scan, listed in the table



Persistent homology process: emergence and lifetime of connected components H0 and cycles H1 visualized across increasing scales. The rightmost panel is the persistence diagram (PD).

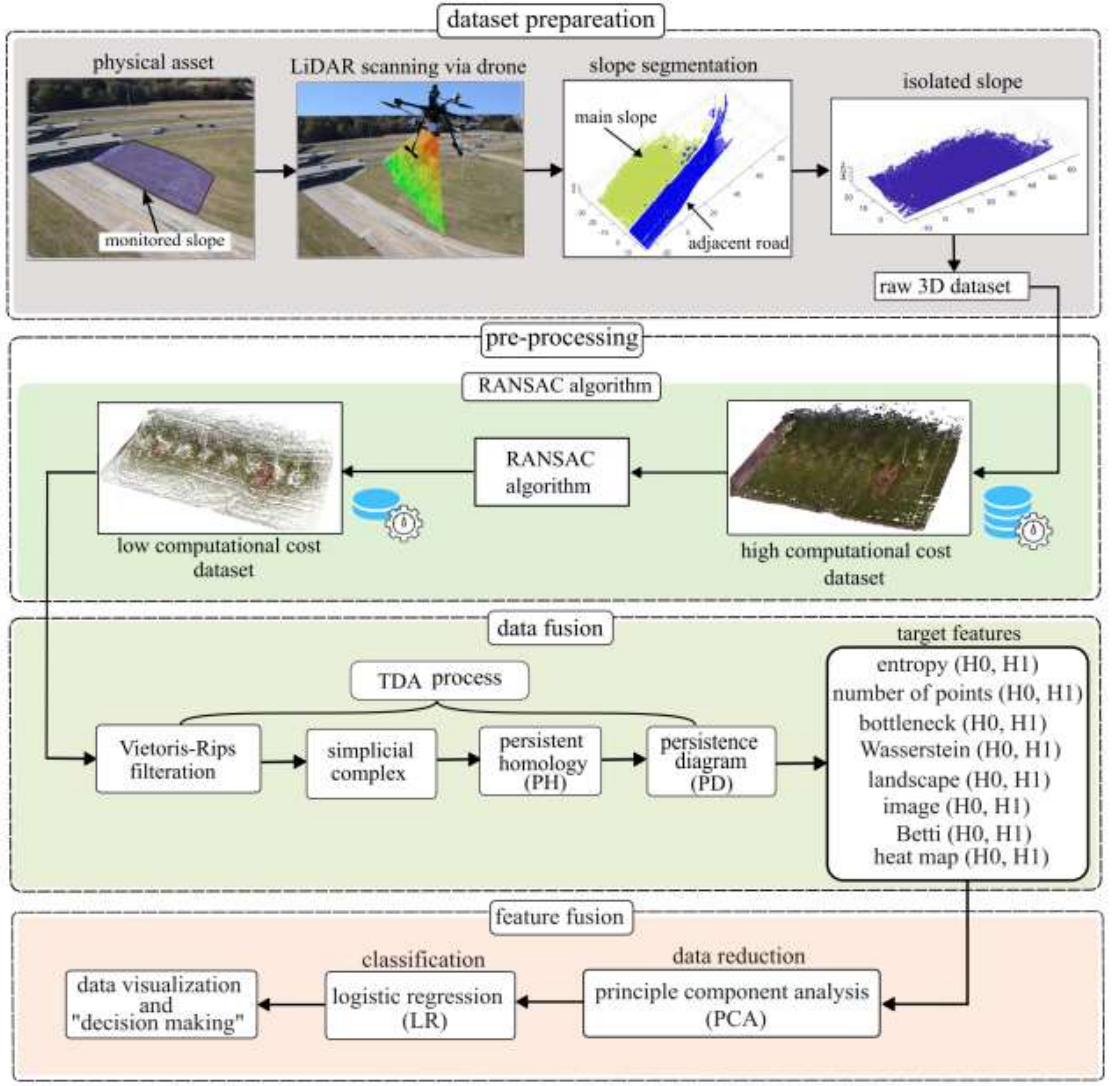
Metric	Dimension	Brief description
persistence entropy	H0,H1	Measures complexity in feature lifetimes.
persistence landscape	H0, H1	Summarizes persistent features as continuous landscapes.
persistence image	H0, H1	Converts diagrams to heatmap-like grids for visualisation.
Betti curve	H0, H1	Records changing counts of H0 and H1 across scales.
heatmap amplitude	H0, H1	Highlights persistent features using heat kernel smoothing.
number of PD points	H0,H1	Counts birth–death pairs per dimension.
bottleneck distance	H0, H1	finds the largest move in the optimal matching to the diagonal.
Wasserstein distance	H0, H1	Computes the norm of point-to-diagonal distances.

RANSAC-TDA Feature Extraction

- Isolate the slope from the road/bridge
- Sample m points from the roughly 12 M-point cloud, K times
- Compute persistent homology on each sample
- Use PCA and logistic regression to inform decision-making

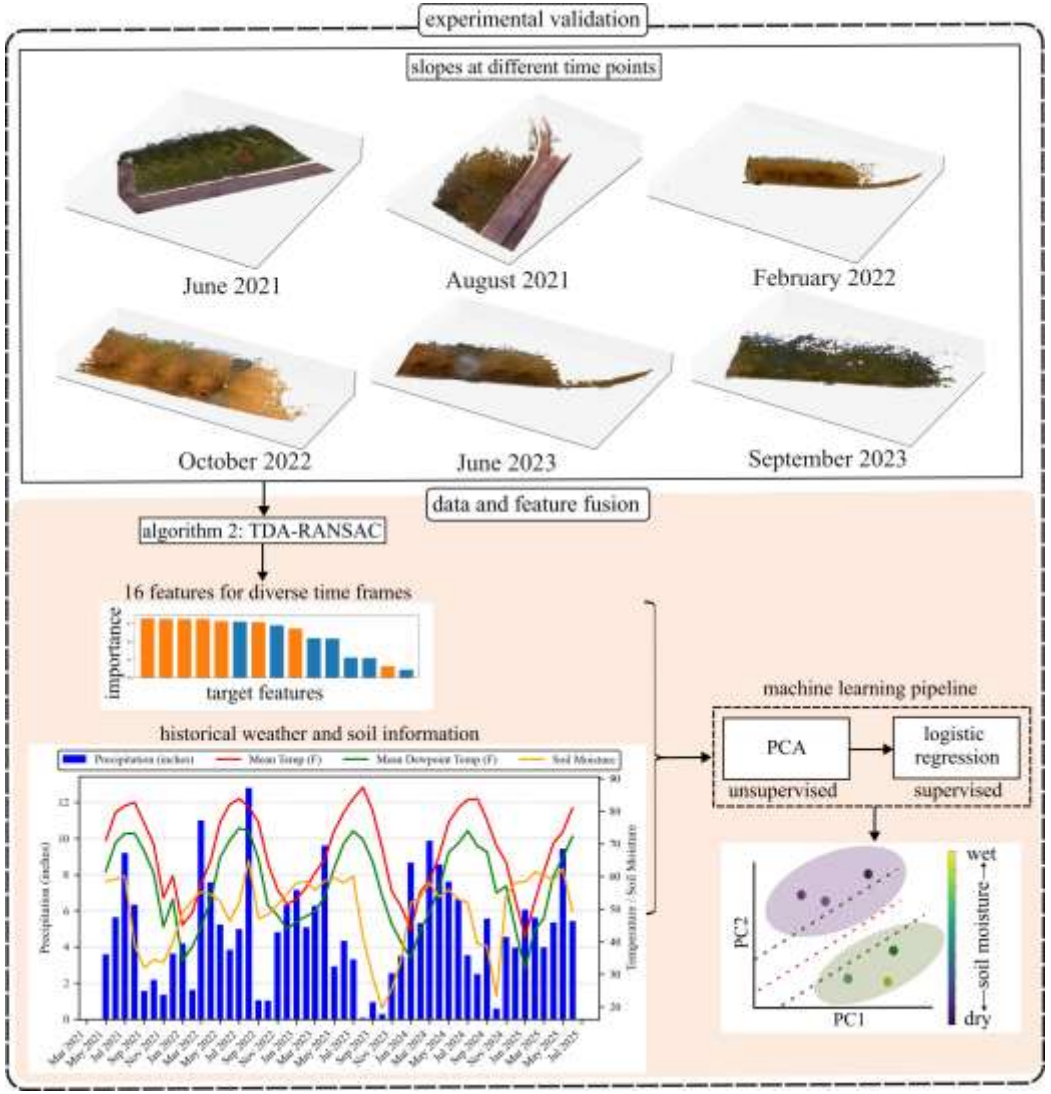
Algorithm 1 RANSAC-TDA Feature Extraction from Point Clouds

procedure RANDOMSAMPLINGCONSENSUSTDA(P)
  Input: $P \in \mathbb{R}^{N\times 3}$, local 3D point cloud
  Output: median TDA feature vector
  Set sample size m and number of repetitions K
  for $i = 1$ to K **do**
    Randomly sample $S_i \subset P$ with m points
    Compute persistent homology on S_i
    Extract selected TDA features from H_0 and H_1
  end for
  Return the component-wise median feature vector
end procedure



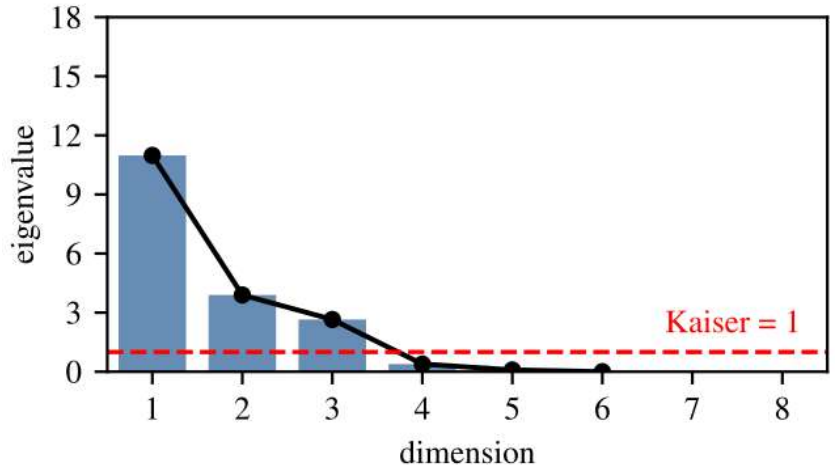
Testing on Six Seasonal Scans

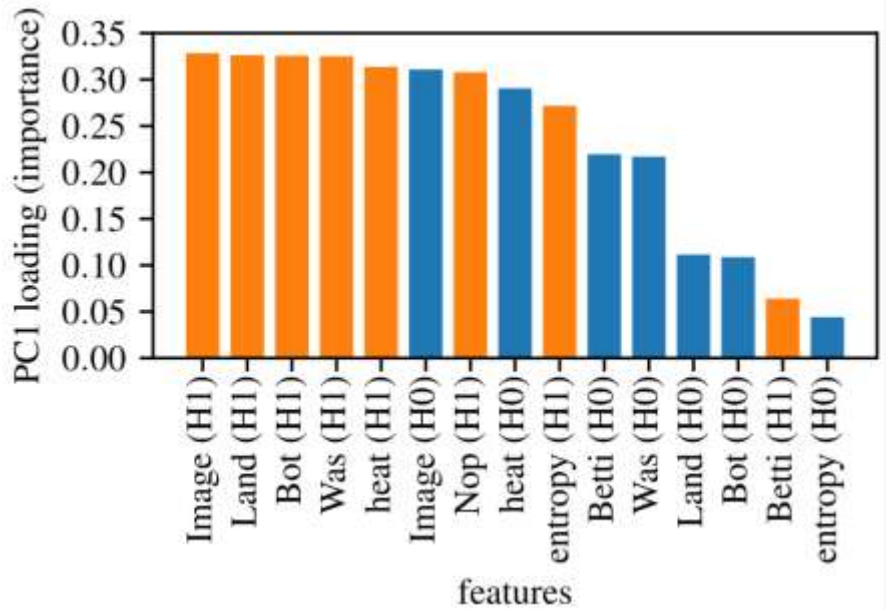
- The same pipeline from the previous slide, run on all six SLidE scans
- Each scan becomes one 16-feature vector
- Soil moisture labels each scan wet or dry
- Moisture is a reference only, not an input



A Few Components Capture Most of the Variation

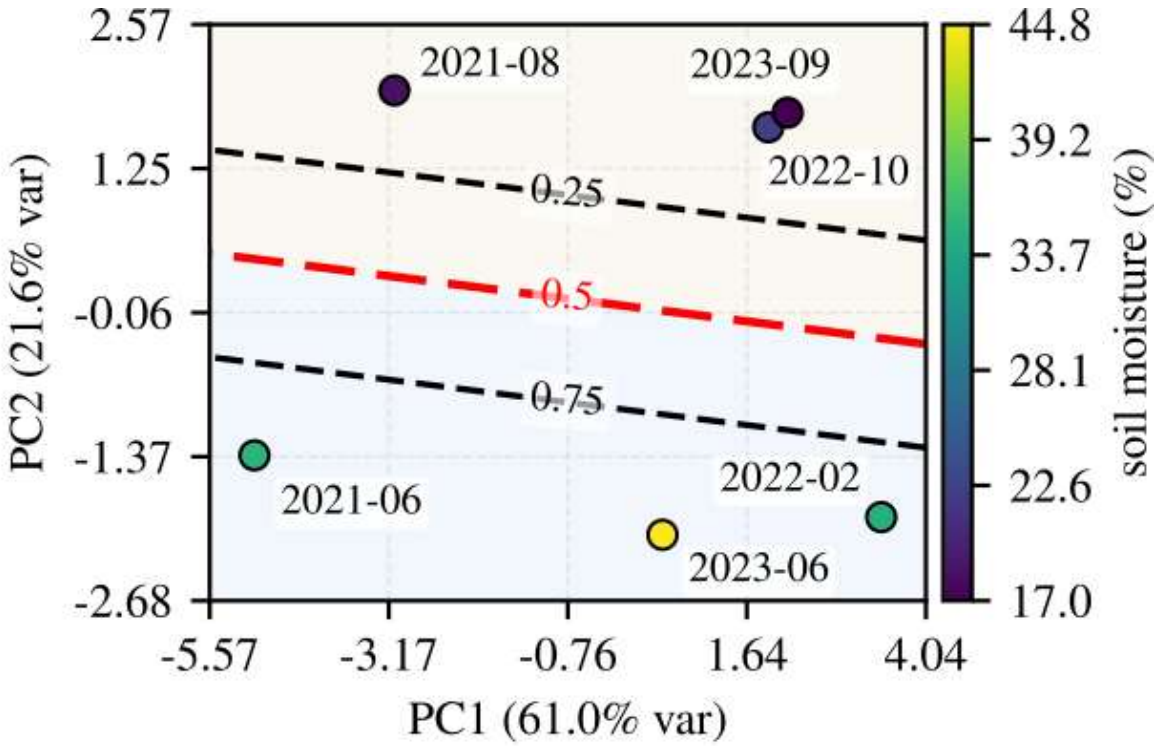
- Single-feature wet–dry accuracy ranges from 0% to 100%
- 16 features compress to 3 components above the Kaiser line
- PC1 and PC2 hold 61.0% and 21.6% of the variance
- The top five PC1 loadings are all H1 features





Fused Features Separate Wet and Dry Scans

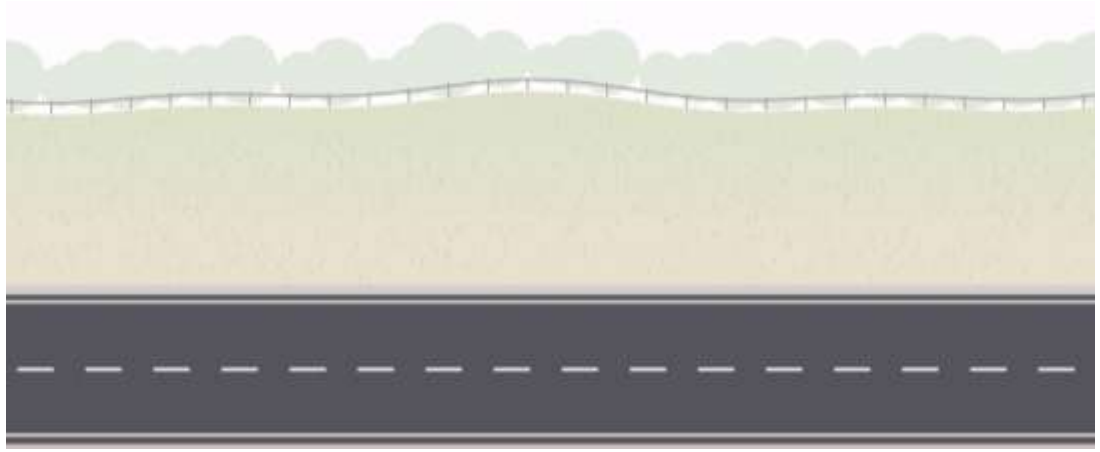
- Each point is one scan, colored by soil moisture
- Dry scans (2021-08, 2022-10, 2023-09) cluster at the top
- Wet scans (2021-06, 2022-02, 2023-06) fall below
- Dashed lines show the logistic-regression probability of a wet scan
- One vector per scan shows when the slope changed, not where



Localized TDA for Adaptive Sampling (Current Study)

Everyday Vehicles as Roadside Scanners

- Data-collection cars already carry roof-mounted LiDAR
- Each pass scans, so DDDAS tracks change between visits
- Localized TDA flags the segment that needs a closer look

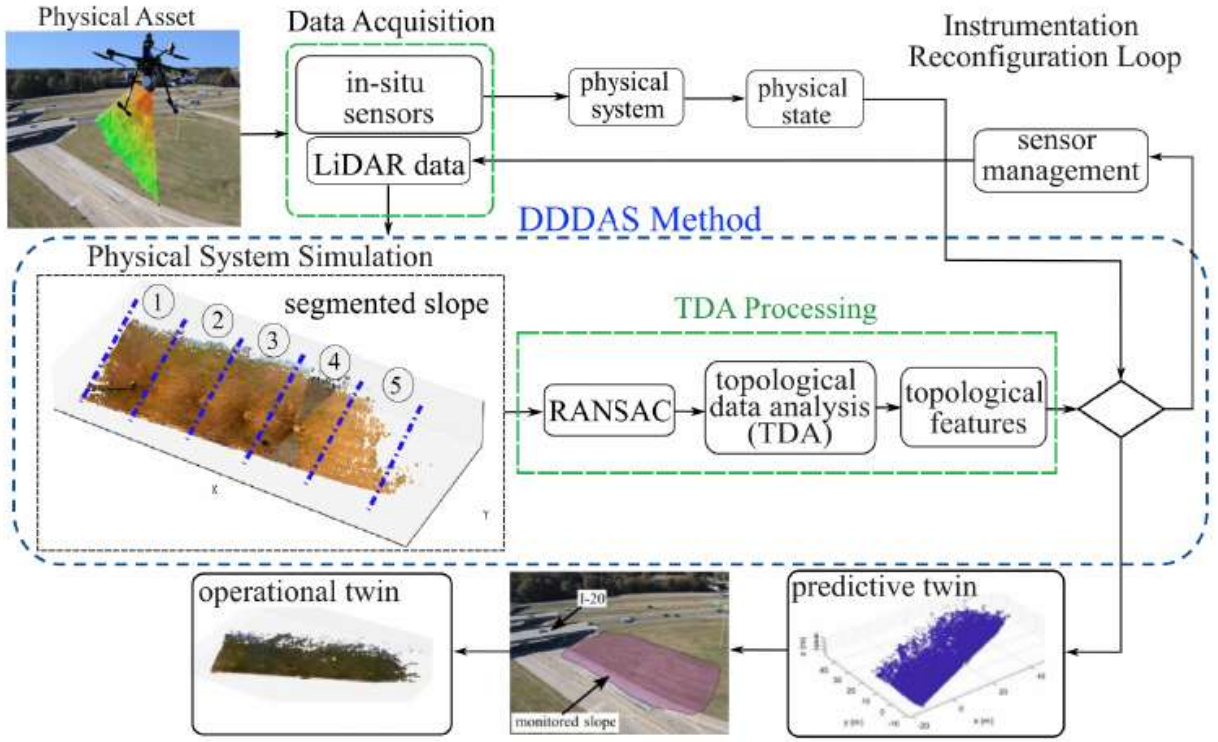




Volvo Cars Data Collection Vehicle, Austin Downey, CC BY-SA 2.0

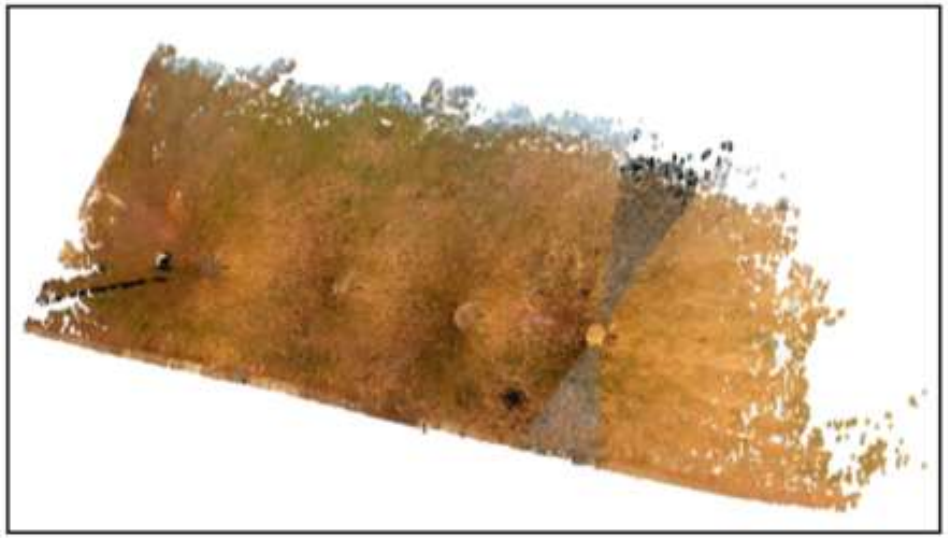
DDDAS Framework for Localized LiDAR-Based Embankment Monitoring

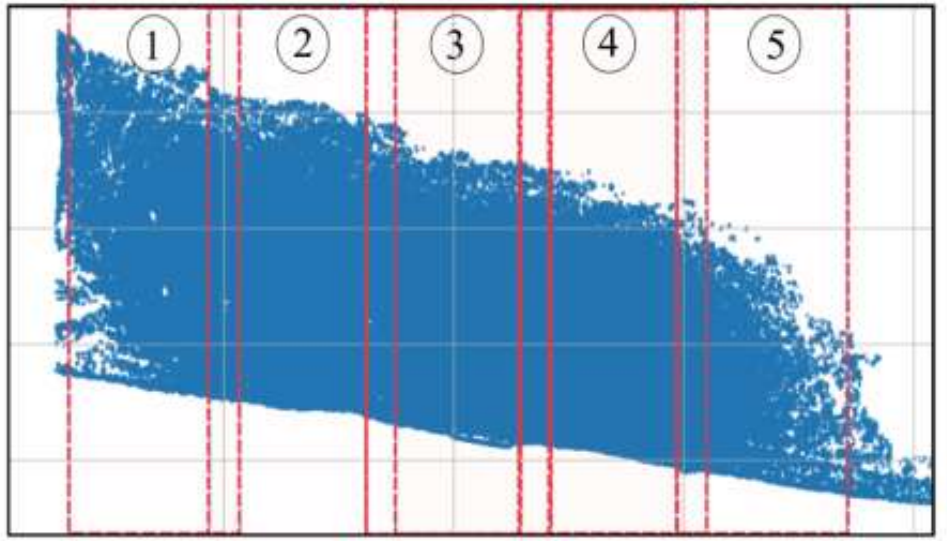
- The proposed framework divides large 3D LiDAR point clouds into overlapping local slope segments for targeted analysis
- Each segment is independently analyzed using RANSAC-based sampling and Topological Data Analysis (TDA) to extract geometric features efficiently
- Localized topological descriptors reveal spatial variations and identify regions exhibiting potentially abnormal surface deformation
- The DDDAS feedback loop supports adaptive monitoring by prioritizing regions of interest for future sensing and inspection



Slope Segmentation

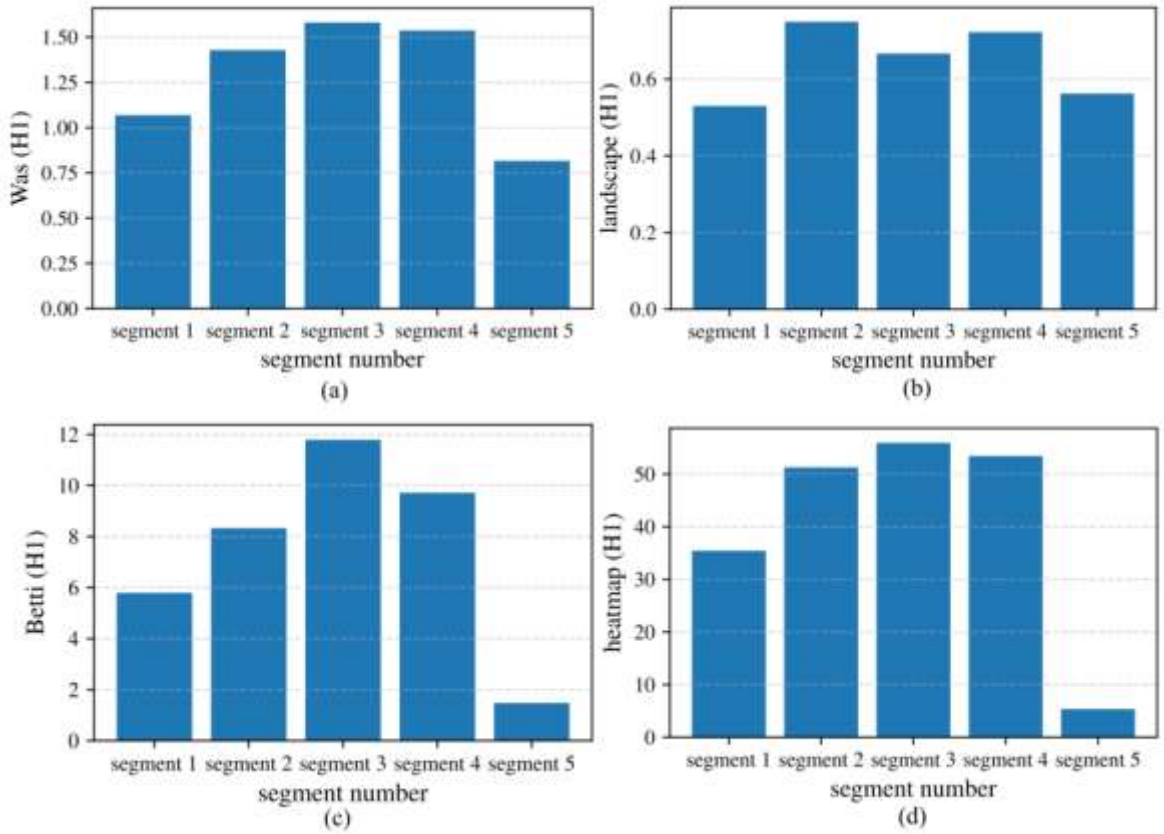
- The cleaned 3D LiDAR point cloud represents the monitored embankment surface
- The embankment is divided into five overlapping local segments to capture spatial variations in surface geometry
- Each segment is independently processed using RANSAC-based sampling and TDA to extract localized topological descriptors





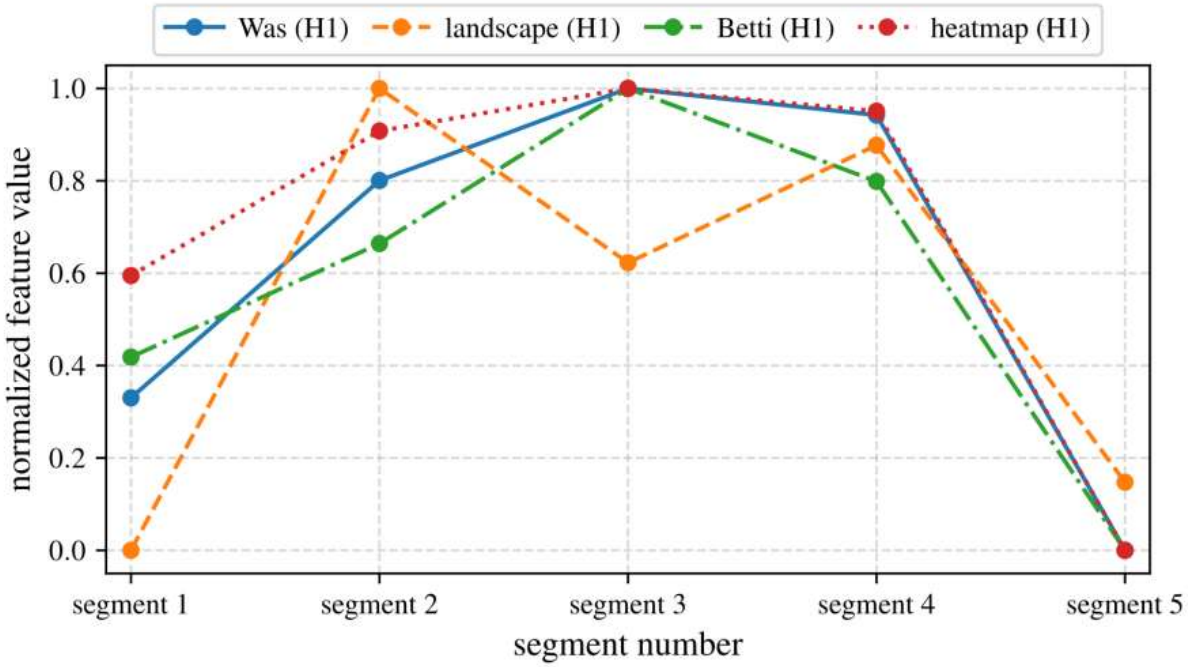
Descriptors Vary Across Segments

- Four H_1-based topological descriptors (Wasserstein, persistence landscape, Betti, and heatmap) are compared across five slope segments
- Segment 3 exhibits the highest Wasserstein, Betti, and heatmap responses, while persistence landscape peaks at segment 2
- Betti H_1 is about 11.8 in segment 3 versus about 1.5 in segment 5, an 8× spread



Normalized Responses Peak Mid-Slope

- Each descriptor is min–max normalized to 0–1 across segments
- Segments 2–4 score at least 0.62 on all four descriptors
- Segment 5 stays at or below 0.15 on all four
- The nonuniform descriptor responses reveal localized geometric variations that may be overlooked in whole-slope analysis



Acknowledgements

This material is based upon work supported by the United States' National Science Foundation grant numbers CCF-1937535, CCF-1956071, CCF-2234921, CCF-2503055, CPS-2237696, IIS-2324052 and CMMI-2152896. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Air Force or the United States' National Science Foundation.






Discussion

SLidE

github.com/ARTS-Laboratory/Dataset-SLidE

Name: Austin Downey
Title: Associate Professor
Email: austindowney@sc.edu

Molinaroli College of Engineering and Computing
UNIVERSITY OF SOUTH CAROLINA

References

- Austin R. J. Downey, Jie Wei, A. Q. M. Zohuruzzaman, Paul T. Schrader, Sadik Khan, Jason Bakos, Weicong Feng, Erik Blasch, and Erika Ardiles-Cruz. Earthen Embankment Monitoring Using LiDAR Data by Randomized Consensus of Topological Data Analysis, pages 69-77. Springer Nature Switzerland, August 2025. doi:10.1007/978-3-031-94895-4_7
- AQM Zohuruzzaman, David P. Wamai, Weicong Feng, Sadik Khan, Austin R. J. Downey, Jie Wei, Erik Blasch, and Paul T. Schrader. Highway slope monitoring using 3D laser scanning at different seasons. In Kannappan Palaniappan and Gunasekaran Seetharaman, editors, *Geospatial Informatics XIV*. SPIE, June 2024. doi:10.1117/12.3016172
- Mohsen Gol Zardian, Weicong Feng, A. Q. M. Zohuruzzaman, Austin R. J. Downey, Jie Wei, Sadik Khan, Paul T. Schrader, Erik Blasch, Erika Ardiles-Cruz, Jason Bakos, and Jasim Imran. Seasonal surface condition measurement of expansive clay embankments based on consensus topological analysis of LiDAR point clouds. SSRN preprint, 2026. doi:10.2139/ssrn.7669874
- Sadik Khan, AQM Zohuruzzaman, David Wambi, Austin R.J. Downey. “Dataset-Slope-LiDAR-embankment-(SLidE),” April 2024. [Online]. Available: URL: https://github.com/ARTS-Laboratory/Dataset-SLidE