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

Mohsen Gol Zardian[1][0000−0002−9878−2033], A Q M Zohuruzzaman[2][0009−0000−2863−1929], Weicong Feng[3][0009−0007−9246−8792], Austin R.J. Downey[1,4][0000−0002−5524−2416], Jie Wei[3][0000−0003−1781−6412], Sadik Khan[2][0000−0002−0150−6105], Paul T. Schrader[5][0000−0002−6017−0509], Erik Blasch[5,6][0000−0001−6894−6108], Erika Ardiles-Cruz[5][0000−0003−3092−4696], Jason Bakos[6][0000−0002−0821−6258], and Jasim Imran[4][0000−0002−3479−3905]

[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, OH, United States of America
[7] Computer Science and Engineering, University of South Carolina, Columbia South Carolina, USA

**Abstract.** This study presents a data-driven framework for localized monitoring of earthworks using Topological Data Analysis (TDA) on repeat LiDAR surveys from the Slope LiDAR Earthwork (SLidE) dataset. A TDA-based homology is used to extract topological features from LiDAR point clouds and track temporal changes in slope morphology. Moving beyond whole-slope analysis, the TDAEED (TDA-Earth DDDAS) partitions the earthwork into overlapping local segments to identify defect-prone regions with greater spatial precision. Within a Dynamic Data-Driven Applications Systems (DDDAS) framework, the localized analysis supports adaptive sampling by iteratively focusing computation and sensing on regions exhibiting abnormal topological behavior. Persistence diagrams and topological descriptors are computed independently for each segment, while a RANdom SAmple Consensus (RANSAC) method reduces computational cost to preserve dominant geometric structure. Compared with whole-slope analysis, the localized workflow improves spatial specificity, directly localizes abnormal regions, and increases computational efficiency for scalable monitoring. Applied to the SLidE dataset, the TDAEED framework reveals nonuniform spatial and temporal evolution of the earthwork surface, highlighting localized regions that may indicate elevated deformation or failure susceptibility. TDAEED demonstrates targeted inspection and long-term geotechnical monitoring within a computationally efficient DDDAS environment.

**Keywords:** infrastructure· 3D LiDAR · Topological Data Analysis (TDA) · Expansive clay · Moisture dynamics · Shrink-swell behavior DDDAS · Dynamic Data Driven Applications Systems · InfoSymbiotic Systems

# 1 Introduction

Earthworks are critical components of the transportation and flood-protection infrastructure, yet their condition can change gradually under environmental and operational loading. Small defects such as cracking, shallow depressions, localized heave, and surface roughening may appear before larger slope instability becomes visible [10]. Early identification of these localized changes is important because maintenance is more effective when distress is detected before it propagates across the embankment [8]. To address embankment monitoring, LIDAR measurements are available to facilitate the Slope LiDAR Embankment(SLidE) challenge [14].

The SLiDE challenge is amplified in expansive clay slopes. In Yazoo clay and similar high-plasticity soils, seasonal wetting and drying can cause repeated swelling, shrinkage, desiccation cracking, and strength reduction. Since these processes are often spatially nonuniform, one region of the slope may deform while another remains comparatively stable. Monitoring methods that describe only the average or whole-slope behavior may therefore miss the local nature of early damage [6]. Conventional inspection, borehole sampling, laboratory testing, and *in-situ* sensors provide useful information, but they are limited in spatial coverage and are difficult to repeat frequently over large areas [11]. LiDAR provides a non-contact alternative by producing dense three-dimensional measurements of slope geometry. The Slope LiDAR earthwork (SLidE) dataset provides repeated LiDAR surveys of a Yazoo clay highway earthwork in Jackson, Mississippi [7]. However, these point clouds contain millions of points, so efficient feature extraction is needed before the data can support monitoring decisions. Topological Data Analysis (TDA), particularly persistent homology, provides a compact way to describe the shape of point-cloud data by tracking connected components and loop-like structures across spatial scales [4,2]. Prior work has shown that TDA can support sensing and infrastructure applications [12]; however, whole-slope TDA produces global descriptors and does not directly identify where abnormal behavior occurs.

This paper presents a computationally efficient adaptive sampling framework for localized TDA-based earthwork monitoring within a DDDAS setting. The proposed workflow partitions the LiDAR point cloud into local segments, computes persistent-homology-based descriptors for each segment, and uses RANSAC-based sampling to reduce computational cost while preserving dominant geometric structure. The goal is to move beyond whole-slope monitoring toward spatially targeted assessment of expansive clay earthworks.

## 1.1 DDDAS for Intelligent Slope Monitoring

The Dynamic Data-Driven Applications Systems (DDDAS) paradigm links sensing, computation, and decision-making through a feedback process [1]. For earthwork monitoring, knowledge feedback is useful because slope conditions evolve with moisture variation, shrink–swell behavior, cracking, and localized deformation. Instead of collecting and processing data uniformly, a DDDAS framework can use previous observations to decide where additional sensing or computation is most needed, as shown in Figure 1.

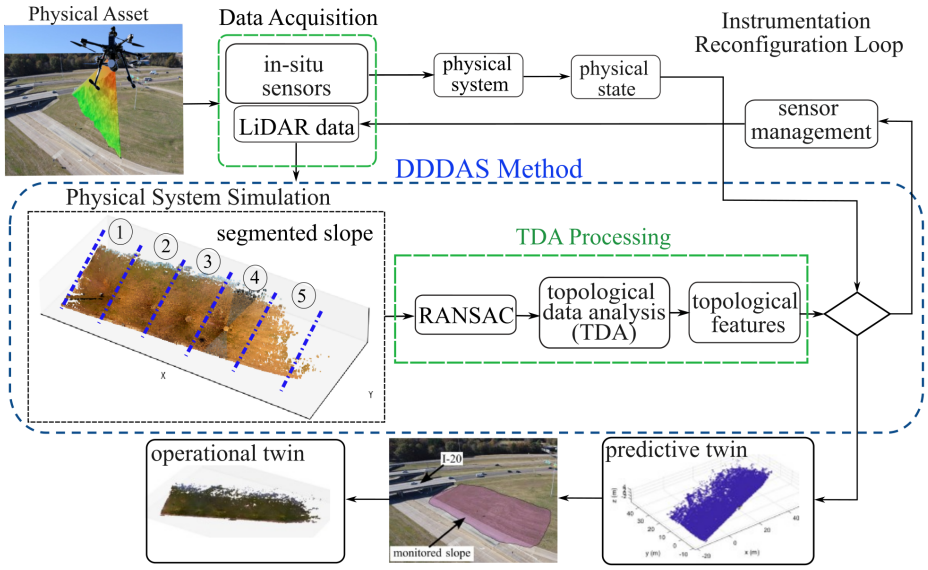



**Fig. 1.** Proposed DDDAS framework for intelligent slope monitoring using localized LiDAR segmentation and TDA-based processing. Repeat LiDAR and in situ observations are used to update the physical state of the earthwork, while the slope surface is partitioned into local segments for region-level topological analysis. TDA based features extracted from each segment support operational and predictive twin updates, sensor management, and adaptive reconfiguration of future monitoring toward regions exhibiting abnormal geometric behavior

Figure 1 illustrates the proposed DDDAS workflow. LiDAR data and available *in-situ* measurements update the physical state of the earthwork, while the point cloud is partitioned into local slope segments. Persistent homology is then applied to each segment to extract topological descriptors, and RANSAC-based sampling reduces computational cost. Regions with abnormal or rapidly changing descriptors can be passed to the sensor-management loop, guiding future LiDAR scans, denser sampling, or targeted field inspection. Stable regions can be assigned lower priority, allowing sensing and computation to focus on candidate defect-prone areas. The proposed framework builds on the DDDAS-based monitoring framework introduced by Downey et al. [3] and shifts from comparisons of full LiDAR scans to spatially resolved analysis within a scan.

The proposed DDDAS-based monitoring process consists of four main components:

1. **Data acquisition and state updating:** Repeat LiDAR surveys and available in situ measurements provide geometric and environmental information about the current condition of the physical earthwork, as shown in Figure 2.
2. **Localized slope segmentation:** The LiDAR point cloud is divided into overlapping local segments so that topological changes can be associated with specific regions of the slope rather than only with the full earthwork.

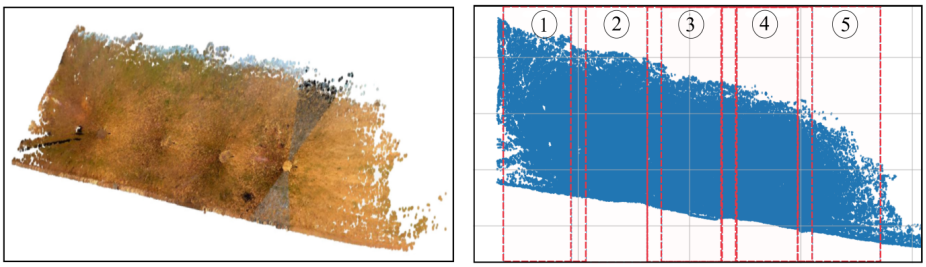



**Fig. 2.** Representative LiDAR point cloud and localized segmentation of the monitored earthwork, showing: (a) Cleaned three-dimensional point cloud of the February 2022 SLidE scan, (b) Two-dimensional profile view of the same point cloud partitioned into five local slope segments, which define the analysis regions used for segment-level TDA feature extraction and adaptive monitoring.

3. **TDA-based feature extraction:** Persistent homology is applied to each segment to compute topological descriptors that capture local geometric structure, while RANSAC-based sampling reduces the computational cost of processing dense point clouds.
4. **Adaptive sensing and decision support:** Segment-level descriptors are used to identify abnormal regions, update the operational and predictive twins, and guide future sensing or inspection toward areas with elevated deformation potential.

## 2 Methodology

This section introduces the SLidE Dataset and the TDA processing technique.

### 2.1 SLidE Dataset

The publicly available Slope LiDAR Earthwork (SLidE) dataset is used as the source of repeat point-cloud observations [7]. The dataset was introduced by Zohuruzzaman et al. [14] and contains LiDAR scans of an earthwork near the I-20E exit toward Terry Road in Jackson, Mississippi [6]. The site is founded on Yazoo clay, a highly plastic expansive soil susceptible to shrink–swell deformation. The monitored slope is approximately 15 ft high, has a grade of 3.5:1 to 4:1 (V:H), and includes both reinforced and as-built sections due to previous shallow landslide remediation.

### 2.2 Topological Data Analysis

Topological Data Analysis (TDA) extracts shape-based information from complex datasets [9]. Persistent homology (PH), one of the main tools in TDA, tracks how topological structures appear and disappear as a filtration scale changes [13]. For LiDAR point clouds, $H_0$ represents connected components and $H_1$ represents loop-like structures. These descriptors provide compact information about point-cloud geometry and are useful when the surface is noisy, irregular, or spatially dense.

Direct PH computation on dense LiDAR data is computationally expensive, with cost scaling as $\mathcal{O}(n^3)$ for $n$ points [5]. Since SLidE scans contain millions of points, full-resolution analysis is impractical, especially when TDA must be repeated over multiple local segments. The proposed workflow reduces this cost by first dividing the earthwork point cloud into local regions and then applying random sampling within each segment. A subset of $m$ points is selected from each local point cloud, reducing the computational cost to $\mathcal{O}(m^3)$ while preserving dominant geometric structure.

To reduce sampling bias, each segment is sampled repeatedly and the median topological response is used as the final descriptor as following the logic of RANdom SAmple Consensus (RANSAC) [14], where multiple random subsets provide a robust estimate. In this study, persistence entropy and related TDA descriptors are computed from $H_0$ and $H_1$ persistence information, with emphasis on selected $H_1$ features for localized slope monitoring. The RANSAC-TDA workflow is summarized in Algorithm 1.

**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**

## 3 Results and Discussion

Four selected $H_1$-based descriptors are used to evaluate spatial changes in the segmented earthwork surface: Wasserstein amplitude, persistence landscape, Betti curve, and heatmap amplitude. These descriptors summarize loop-related behavior in the LiDAR point cloud. This is relevant for expansive clay slopes because wetting, drying, cracking, swelling, and local depressions can create more complex local surface geometry. Since $H_1$ captures loop-like structures, changes in these descriptors can indicate segment-level changes in slope morphology.

Figure 3 shows the normalized values of the four descriptors across five local segments. The response is not uniform along the slope, which confirms that localized TDA can capture spatial variability that would be hidden in a whole-slope descriptor. Segment 3 shows the most distinct behavior, reaching the maximum normalized value for the Wasserstein, Betti, and heatmap descriptors. Although the persistence landscape reaches its maximum at segment 2, segment 3 remains one of the higher-response regions.

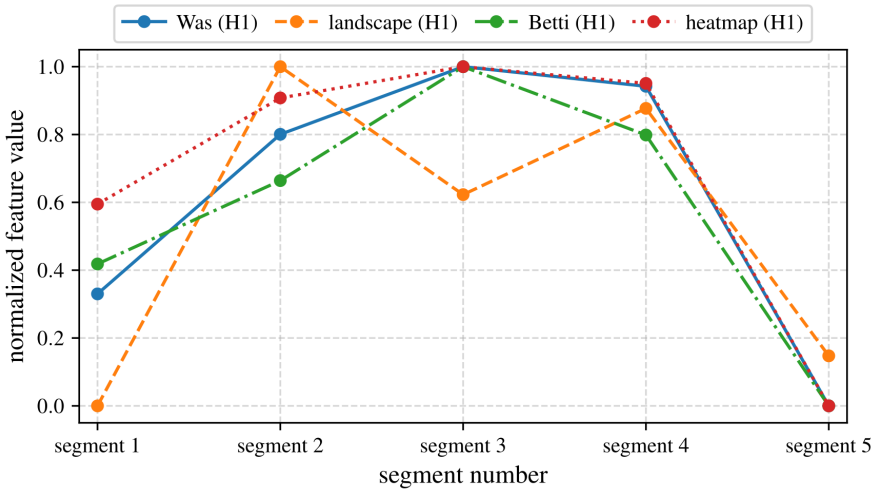



**Fig. 3.** Normalized values of four selected $H_1$-based topological descriptors across five local earthwork segments. The results show nonuniform spatial variation in the topological response, with segment 3 producing the obvious fluctuation normalized value for multiple descriptors.

The individual bar plots in Figure 4 provide a feature-level comparison. Figure 4(a) shows that the Wasserstein descriptor peaks at segment 3, while Figure 4(b) shows the persistence landscape reaching its largest value at segment 2. Figures 4(c) and 4(d) show that the Betti and heatmap descriptors also reach their maximum values at segment 3. Overall, three of the four selected descriptors identify segment 3 as the strongest response region, suggesting localized geometric complexity that may be associated with roughening, deformation, or nonuniform shrink–swell effects.

These results demonstrate the value of segment-level TDA for localized earthwork monitoring. Rather than assigning one topological signature to the entire slope, the proposed workflow evaluates each local segment independently. Within the DDDAS framework, this information can guide future LiDAR scanning, increase sampling density in selected areas, or support targeted field inspection of regions with abnormal topological behavior.

## 4 Conclusion

This study presented a localized TDA-based framework for monitoring earthen earthworks using LiDAR point-cloud data. Instead of analyzing the full slope as a single domain, the proposed method partitions the earthwork into local segments and computes persistent-homology-based descriptors for each region. The selected $H_1$ descriptors showed that segment 3 produced the strongest response in multiple features, indicating that localized segmentation can reveal spatially nonuniform topological behavior that may be hidden in whole-slope analysis.

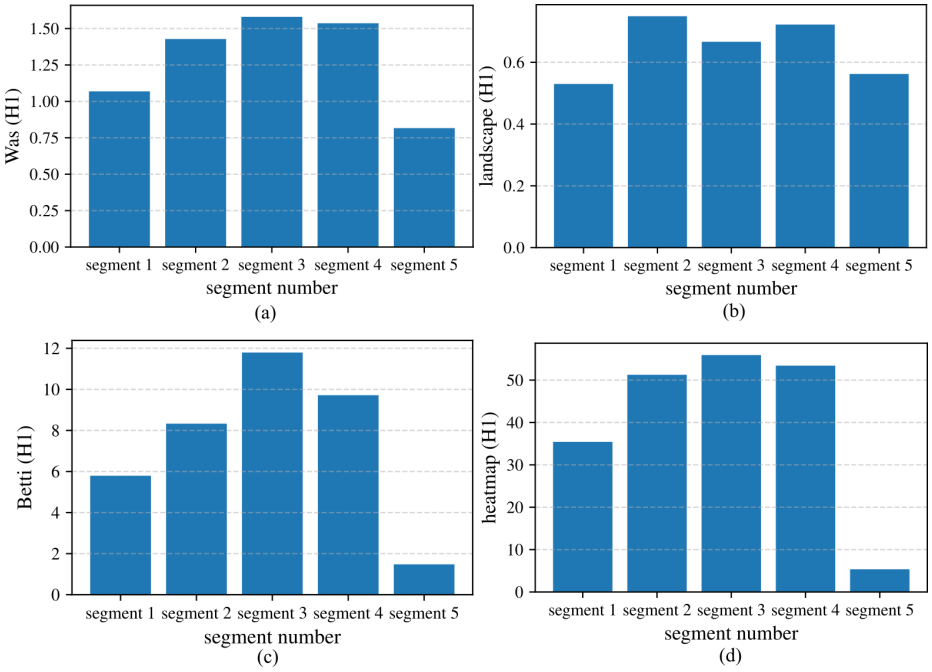



**Fig. 4.** Segment-level comparison of selected $H_1$ topological features, showing: (a) Wasserstein amplitude, (b) persistence landscape, (c) Betti curve, and (d) heatmap amplitude. Segment 3 shows the maximum responses as well as strong dip, indicating a localized region with stronger loop-related topological activity.

The RANSAC-based sampling strategy reduces the computational burden of TDA while preserving the dominant geometric structure of each local segment. Within a DDDAS framework, the resulting segment-level descriptors can support adaptive monitoring by prioritizing abnormal regions for future sensing, denser sampling, or field inspection. Future work will extend the analysis to additional seasonal scans and integrate in situ measurements for further validation.

**Acknowledgments.** 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.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Blasch, E.P., Darema, F., Bernstein, D.: Introduction to the Dynamic Data Driven Applications Systems (DDDAS) Paradigm, pp. 1–32. Springer International Publishing (2022). https://doi.org/10.1007/978-3-030-74568-4_1

2. Carlsson, G.E.: Topology and data. Bulletin of the American Mathematical Society **46**, 255–308 (2009)
3. Downey, A.R.J., Wei, J., Zohuruzzaman, A.Q.M., Schrader, P.T., Khan, S., Bakos, J., Feng, W., Blasch, E., Ardiles-Cruz, E.: Earthen Embankment Monitoring Using LiDAR Data by Randomized Consensus of Topological Data Analysis, pp. 69–77. Springer Nature Switzerland (Aug 2025). https://doi.org/10.1007/978-3-031-94895-4_7
4. Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. Discrete Comput. Geom. **28**(4), 511–533 (nov 2002)
5. Edelsbrunner, H., Harer, J.: Computational Topology - an Introduction. American Mathematical Society (2010)
6. Khan, M.S., Amini, F., Nobahar, M., et al.: Performance evaluation of highway slopes on Yazoo clay. Tech. rep., Mississippi. Dept. of Transportation (2020)
7. Khan, S., Zohuruzzaman, A., Wambi, D., Downey, A.R.: Dataset-slope-lidar-embankment-(slide) (Apr 2024), https://github.com/ARTS-Laboratory/Dataset-Slope-LiDAR-Embankment-SLidE
8. Laflamme, S., Blasch, E., Ubertini, F., Liu, Z., Wertz, J., Knott, C., Cherry, M., Lindgren, E., Chang, F.K., Kumar, A., Poole, J., Worden, K., Downey, A., Wei, J., Musgrave, P.F., Wong, A.S., Quaranta, G., Rosso, M.M., Marano, G.C., Chen, Y., Ardiles-Cruz, E., Soleimani-Babakamali, M.H., Avci, O., Inman, D.J., Taciroglu, E., Dodson, J., Chen, G., Meng, W., Zhu, C., Liu, Z., Zuo, J., Liu, Q., Khan, S., Hu, C., Hu, Z., Cicirello, A., Cross, E., Chatzi, E., Weng, Y., Yuan, J., Wen, S., Han, L., Metaxas, D., Tronci, E.M., Moaveni, B., Chen, Q., Ng, M.S., Hackl, J., Chen, G., Wei, S., Mitoulis, S.A., Izonin, I., Uva, G., Ruggieri, S., Mao, Z., Kiranyaz, S., Devecioglu, O.C., Gabbouj, M., Mohammadi, J.: Roadmap: Integrating artificial intelligence in structural health monitoring systems. Measurement Science and Technology **37**(10), 103001 (Mar 2026). https://doi.org/10.1088/1361-6501/ae3abb
9. Munkres, J.: Elements of Algebraic Topology. Addison-Wesley Publishing Company, Menlo Park California (1984)
10. Nobahar, M., Khan, M.S., Ivoke, J.: Combined effect of rainfall and shear strength on the stability of highway embankments made of Yazoo clay in Mississippi. Geotechnical and Geological Engineering **38**(3), 2787–2802 (Jan 2020). https://doi.org/10.1007/s10706-020-01187-8
11. Puppala, A.J.: Performance evaluation of infrastructure on problematic expansive soils: Characterization challenges, innovative stabilization designs, and monitoring methods. Journal of Geotechnical and Geoenvironmental Engineering **147**(8) (Aug 2021). https://doi.org/10.1061/(asce)gt.1943-5606.0002518
12. Syzdykbayev, M., Karimi, H.A.: Exploring topological information beyond persistent homology to detect geospatial objects. Remote Sensing **16**(21), 3989 (Oct 2024). https://doi.org/10.3390/rs16213989
13. Vietoris, L.: Uber den hoheren zusammenhang kompakter raume und eine klasse von zussammenhangstruen abbildungen. Mathematical Annuals **97**(1), 454–472 (1927)
14. Zohuruzzaman, A., Wamai, D.P., Feng, W., Khan, S., Downey, A.R.J., Wei, J., Blasch, E., Schrader, P.T.: Highway slope monitoring using 3D laser scanning at different seasons. In: Palaniappan, K., Seetharaman, G. (eds.) Geospatial Informatics XIV. p. 9. SPIE (Jun 2024). https://doi.org/10.1117/12.3016172