Interactively Visualizing Future Fires using Deep Learning

Aditi Kannan* Olivia Heng†

Brown University

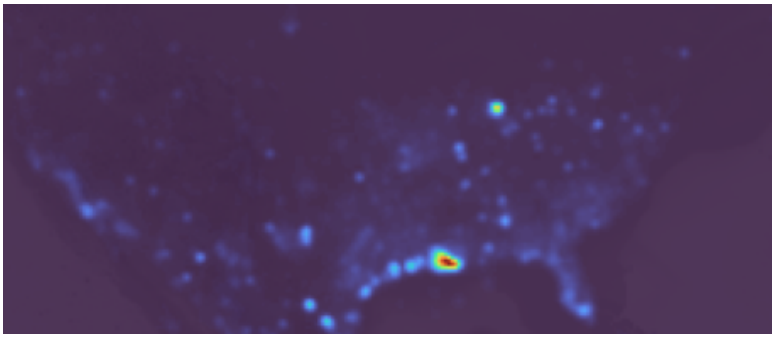

Figure 1: A result of the DL model results for a user-given wind speed and oxygen level.

ABSTRACT

We present a lightweight visual analysis pipeline that integrates NASA VIIRS/MODIS active fire detections with ERA5 atmospheric variables to produce 2D spatiotemporal visualizations of potential fire spread in the United States. Contrary to our original proposal for a full 3D oxygen–climate–ecology video-generation framework, our final system focuses on a tractable subset: a data-driven 2D fire-spread predictor conditioned on wind speed and near-surface oxygen-related ERA5 variables, paired with a heatmap-based visualization tool. Using a small deep learning model for 2D video generation, we produce short predictive animations of fire movement across time based on preceding observations. These results are overlaid on geographic maps and explored in ParaView to support interpretation. Feedback from our scientific collaborator confirms that the visualization tool is useful for qualitatively assessing atmospheric contributions to spread dynamics. This report describes the implemented pipeline, its evaluation, and limitations, as well as directions for extending the system to the full oxygen-integrated 3D framework originally proposed.

Keywords: Deep learning, wildfire, evaluation.

1 INTRODUCTION

Wildfire behavior depends strongly on atmospheric conditions, yet most operational visualizations rely on static maps or retrospective satellite composites. Our original project aimed to build a comprehensive 3D video-generation model that integrates oxygen flux, climate variables, soil data, and ecological structure. Within the six-week timeframe, we implemented a simplified but functional subset: a 2D video-generation pipeline using NASA FIRMS fire detections and ERA5 reanalysis data to qualitatively predict fire propagation patterns influenced by wind speed and oxygen-related atmospheric variables.

*e-mail: aditi_kannan@brown.edu

†e-mail: olivia_heng@brown.edu

Our goals were (1) to build a minimal predictive model capable of generating short forward simulations, and (2) to produce an interpretable visualization tool enabling our environmental-science collaborator to explore the predicted spatiotemporal fire behavior.

2 EXPOSITION

Wildfires are an increasingly urgent environmental hazard, with recent events in the western United States highlighting the need for better predictive tools. Our project explores whether a deep-learning–based visualization pipeline can help researchers understand and anticipate fire movement using atmospheric and remote-sensing data. While our goal was not to develop a state-of-the-art forecasting model, we aimed to build a workflow that integrates real environmental data, trains a predictive model, and generates exploratory visualizations that support expert analysis.

We partnered with a domain expert studying wildfire dynamics and sought to understand what types of visual outputs would be most helpful for her analysis. We also examined previous visualization approaches by replicating baseline techniques in ParaView, which helped us understand the strengths and limitations of existing workflows. Our final tool, though limited in scope, provides a proof-of-concept interface for visualizing predicted fire spread using wind speed, oxygen levels, and active fire detections.

2.1 Method: Data Collection and Preprocessing

We downloaded VIIRS/MODIS active fire data from NASA FIRMS and ERA5 pressure-level atmospheric variables. All datasets were regridded to a common resolution and temporally aligned at 3–6 hour intervals. From ERA5, we extracted wind speed fields and oxygen-related variables (e.g., vertical wind, pressure gradients) to serve as model inputs.

2.2 Model and Datasets

We implemented a 2D deep learning model to generate short predictive sequences of wildfire movement. The model takes as input:

- the most recent frames of observed fire detections, and
- corresponding ERA5 atmospheric variables.

It outputs a 2D predicted video of fire-spread likelihood. This approach is data-driven rather than physically grounded, but it produces plausible directional propagation patterns influenced by wind and oxygen-related features.

We worked with two primary datasets: (1) NASA FIRMS VIIRS and MODIS active fire detections, accessed through the NASA FIRMS API. These datasets provide geolocated thermal anomalies used as ground-truth fire points. (2) ERA5 reanalysis data, sourced through the Copernicus Climate Data Store, providing atmospheric variables including wind speed, pressure-level features, and humidity. ERA5 served as the environmental context for predicting potential fire spread.

All datasets were clipped to the continental United States and temporally aligned to ensure consistency between meteorological conditions and fire observations.

2.3 Baseline Exploration in ParaView

Before model development, we used ParaView to explore both datasets and to replicate elements of previous visualization research in wildfire modeling. ParaView helped us:

- Inspect the spatial distribution of fire detections over time
- Overlay atmospheric variables onto geospatial grids
- Identify which environmental parameters appeared most visually correlated with fire movement
- Understand the visual interpretability and limitations of standard scientific visualization tools

This exploratory phase directly informed our model design: it helped us choose which ERA5 variables to include and clarified that a 2D predictive visualization would be the most interpretable outcome for our collaborator.

2.4 Model Architecture and Training

We trained a deep learning model designed to take environmental fields (e.g., wind speed, temperature, oxygen-related variables) and produce a 2D prediction of likely fire spread. While the model was intentionally simple, our goal was to evaluate whether environmental features alone could yield meaningful spatial predictions and whether an automated output could help our collaborator see patterns not obvious from raw data For the Training setup, we trained the model on OSCAR, Brown's high-performance computing cluster and experimented with multiple epoch settings to evaluate training stability and determine whether the model benefited from longer training cycles. Because we lacked the time and data to pursue a rigorous model selection process, our approach remained exploratory. The focus was on generating usable visualizations rather than optimizing predictive accuracy. The model generates a 2D grid representing the estimated probability of fire occurrence in the next time step. These outputs were later rendered into visual heatmap overlays.

3 Evaluation

Our evaluation had two components: Visualization Quality and Feedback From Collaborator (Formative Evaluation). We produced several forms of exploratory visualizations, including:

- Heatmap overlays aligning predicted fire intensity with the geographical map
- Side-by-side comparisons between model predictions and actual FIRMS-detected fire points
- Timeline slices displaying how predictions shifted in response to changes in wind speed

These visualizations were used qualitatively rather than quantitatively; the primary goal was to understand whether the output was interpretable and helpful.

We conducted an informal evaluation with our collaborator, asking:

- Whether the visualizations were understandable
- Whether the predictive outputs aligned with her intuition
- Whether such a tool would be useful in future research

She responded positively, noting that the pipeline could be useful for quickly scanning large datasets and that the visual overlays made it easier to identify potential regions of interest. She also indicated clear directions for improvement, including longer prediction horizons, more atmospheric variables, and higher spatial resolution.

4 Conclusion

Our project demonstrates that even a simple deep learning model, when paired with meaningful visualizations, can provide insights into wildfire spread. By integrating FIRMS and ERA5 data, using ParaView for preliminary exploration, and training on OSCAR, we created a workflow that produced informative heatmap-based predictions. The project ultimately served as a prototype for what a more advanced predictive visualization system could look like.

With more time, we would expand the dataset, improve spatial resolution, incorporate more physically meaningful variables, and conduct a more rigorous evaluation. Still, our collaborator found the current tool helpful, reinforcing our main takeaway: even imperfect predictive models can provide valuable insights when paired with well-designed visualizations.

Acknowledgements

We thank our collaborator Divya Banesh, Los Alamos National Laboratory, for their patience and insights during our final project, in addition to supporting us in our project plans. We thank Professor Laidlaw and the course TA for their guidance, and valuable teachings this semester.

5 Contributions

5.1 Olivia Heng

5.1.1 Intellectual Contributions

I contributed intellectually by leading our team's engagement with prior scientific and visualization work. I analyzed existing fire-spread modeling approaches and explored them using ParaView to understand how atmospheric variables and fire detections had been visualized in previous research. This analysis shaped our project direction by informing what datasets to include, what visual encodings would be useful, and what types of predictions our collaborator needed.

I also led communication and scheduling with our national-lab collaborator, helping translate her research needs into concrete project goals. Throughout the semester, I coordinated timelines, guided the framing of our research questions, and helped structure our written reports and presentations. Additionally, I contributed to the interpretation of our model's evaluation results, especially how they aligned with collaborator expectations and practical wildfire-analysis use cases.

5.1.2 Practical Contributions

I managed all exploratory visualization work in ParaView, including importing and visualizing the FIRMS and ERA5 datasets, reconstructing prior workflows, and producing baseline visualizations that informed our model design. I prepared and organized progress reports, drafted major portions of our written deliverables, and led our in-class presentations.

I obtained access to the OSCAR compute cluster and handled logistics related to running our model on the system. I coordinated meeting notes, organized documentation of collaborator requirements, and synthesized these into explicit evaluation criteria for the model's usefulness. I also contributed to refining the final evaluation section by summarizing feedback from our collaborator and integrating it into our final report.

5.2 Aditi Kannan

5.2.1 Intellectual Contributions

I contributed intellectually by leading the development of our predictive model. I researched candidate architectures and datasets, evaluated their suitability for wildfire-spread prediction, and selected the modeling direction we pursued. I also developed approaches for evaluating the model's predictive behavior based on collaborator feedback, including selecting metrics and determining how predictions should be compared to atmospheric variables.

I contributed to discussions on how to expand the dataset, identifying additional ERA5 variables and sources that would strengthen the model. I also helped refine our problem framing by clarifying what aspects of fire spread the model should learn from the available data.

5.2.2 Practical Contributions

I implemented and trained the deep learning model on OSCAR, including preparing data inputs, running repeated training sessions with different epoch counts, and debugging the training pipeline. I handled integration of additional datasets once identified through research, including preprocessing and incorporating them into the training workflow.

I contributed to our evaluation methods by generating predicted outputs, producing intermediate results for inspection, and helping analyze how model behavior changed with different training configurations. I also helped prepare written documentation describing our modeling approach and contributed to finalizing the results section of the report.

Towards 3D Visualization of Power Transmission Grid Systems

Yonas Amha* Aidan LaBella† Korey Sam‡

Brown University

ABSTRACT

Power transmission grids are complex systems that require extensive monitoring. Visualizing the data from these complex systems has become a challenge in recent years. Previous works ultimately fail to completely capture all the dynamics of these multivariate systems. We show that users are both able to more accurately depict transmission grid time-series data in a 3D environment compared to the 2D environment. We also show that users spend less time depicting such features in a 3D environment a 2D environment

Keywords: Virtual reality, human-computer interaction, evaluation.

1 INTRODUCTION AND AIMS

The primary aim of this project to is implement a 3-dimensional power transmission grid visualization utility. Specifically, we aim to provide an immersive virtual reality experience for all types of users, whether they have a background in electrical engineering or not. The **scientific contribution** is that subject matter experts (SMEs) will be able to more quickly gather insights into the dynamics of power grid transmission systems, specifically when there are anomalies or other grid disruptions. Our **visual contribution** will be to provide an innovative and intuitive visualization utility that is effective at displaying the complexity of the multivariate data that power transmission grids produce.

2 DATA

The data used for this project was obtained from the Western Electricity Coordinating Council (WECC). The WECC oversees The Western Interconnection and is responsible for creating, monitoring, and enforcing reliability standards throughout that region. The WECC provided 200MB worth of electrical grid data in CSV format. The data contains information regarding generator location, frequency, voltage, as well as other multivariate variables.

3 SIGNIFICANCE AND RELATED WORK

The intricate workings of power transmission grid systems are difficult to explain, even for SMEs. The current state of the art is limited as the data cannot be explained in an intuitive manner to those that are not SMEs. Although those who work with this type of data may have developed an innate intuition as to what features are important during analysis, those who are not experts (such as shareholders, executives, project managers, etc) will likely not have such an intuition. Additionally, the current 2D representations contain lots of visual clutter, which can make analysis tricky and features left indiscernible. This is where 3D data representations, including immersive virtual reality, have the potential to provide visual information in ways the previous state of the art has not been able to.

The most recent work concerning the effectiveness of current power transmission data visualization methods has explored the usefulness and correctness of existing power grid data, such as with 2D contour maps [1], which were found to largely misrepresent the data. The main weakness of these 2D contour map representations is their inability to provide a streamlined representation of multivariate data, resulting in indiscernible data streams, particularly for voltage violations (i.e. the voltage exceeds or goes below its standard operating range). The human subjects tasked with finding these violations in the data were ultimately only able to correctly identify (on average) five out of fifteen violations. We expect 3D representations to be a plausible solution to this problem as 3D representations, in general, are able to display data at a higher fidelity.

In 2011, Zhu et. al [6] took a data-driven approach to visualization of power grid data. This includes the paradigm of "query-driven" visualization, where the authors focus their visualization efforts on the data that users frequently query. This only addresses the issue of how large amounts of data are filtered, not necessarily the mechanics how it is presented. Our approach will differ in that we propose a volume-agnostic approach in the sense that users should be able to visualize all of the data, not just the most frequently queried data.

In the year 2000, Overbye et al [4] provided a novel 3D visualization method of transmission grid data. This approach was extremely limited in the sense that the authors only provided 3D projects on to existing 2D maps. Our approach will differ in that we will implement a 3D visualization method with existing 3D representations of transmission grid locations.

Other notable work includes physics-based visualizations [5], the use of broadband cable TV data [2] and web-based applications [3].

4 METHODOLOGY

4.1 3D Implementation

The 3D visualization tool is implemented as a Python PyVista standalone desktop application that runs with a given set of CSV files. There are two types of CSV files the implementation expects: metadata and time-series readings. Metadata files contain information about the grid's infrastructure, such as generator locations and connections between generators, buses and loads. The time-series CSV files are expected to contain per-timestep readings for the various entities that are seen in the metadata. Using these two types of data, users can "walk" through the data timestep-by-timestep. This included a "slider" bar so that the user could control the animation speed. Users were encouraged to use this function in conjunction with the pan function, as a way to visualize the grid flow before starting any tests.

*e-mail: yamha@cs.brown.edu

†e-mail: alabella@cs.brown.edu

‡e-mail: korey_sam@brown.edu

4.2 2D Implementation

The 2D visualization was implemented as an interactive oscillation dashboard, designed specifically for analyzing generator frequency oscillations in the Western Electricity Coordinating Council (WECC) power grid. The dashboard presents multiple coordinate views that enable users to observe and analyze oscillator behavior across generators and transmission lines in bot temporal and phase-space representations.

The oscillation dashboard visualizes simulation data, displaying three primary variables generator frequency (f) measured in Hz, power output (P) in MW derived from frequency deviation, and transmission line current (I) in Amperes. The dashboard enables users to observe how oscillations propagate through the system and identify generators exhibiting anomalous frequency behavior.

The dashboard provides interactive controls for exploration: animation playback with adjustable speed, time navigation via slider or keyboard shortcuts, multi-select lists for generators and transmission branches, and variable toggles. These controls enable users to isolate specific generators of interest, adjust the temporal window to observe oscillation patterns at different timescales, and customize visual encodings to highlight features relevant to their analysis.

5 Evaluation

The participants in the evaluation phase of this project were engineers from the New York Power Authority (NYPA). The evaluation phase is split into two phases; testing and user feedback. In the testing phase, users are tasked with completing two tests. In the first test, we asked the participants to take WECC 3D model accuracy test. The test tasks the participants with identifying several specific generators within the visualization along with navigating to specific timesteps. Once the participants navigate to a specific timestep, they are tasked with the estimate the frequency of the generator. The frequency guesses will be compared to a ground truth value (which indicates the actual frequency of the selected bus), so as to determine the accuracy of the user participant estimates. The time it took for each participant took identify each generator is also recorded. The idea behind recording the time taken to identify each generator within the test is to determines whether participants can quickly and accurately identify the features within the visualizations. The second test featured participants examining the 2D Visualization, which represents the current state-of-the-art. User then identified the same features within the 2D Visualization as they did the 3D Visualization and data from both tests will be compared so as to determine which visualization performed better empirically.

In second phase of the testing, all participants completed a NASA Task Load Index (TLX) survey for both the 2D and 3D Visualizations. A NASA TLX is a questionnaire that assesses the subjunctive mental workload of performing a specific task. With regards this project, NASA TLX will help assess whether our 3D visualization is more intuitive then the current state of the art.

6 Conclusion

In this work, we demonstrate that 3D visualization methods for power transmission grid systems outperform the 2D methods. While our experiments were limited to 3D desktop environments, we believe there could be added value in porting the software to an immersive virtual reality (VR) package. Additionally, our software does not provide for geospatial context (such as with map overlays), which could help users find the grid entities sooner.

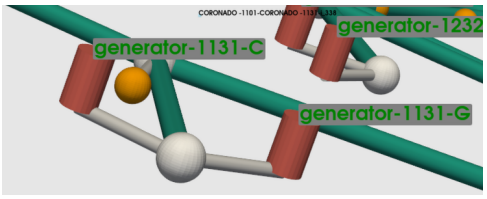

Figure 1: Two generators in the 3D implementation.

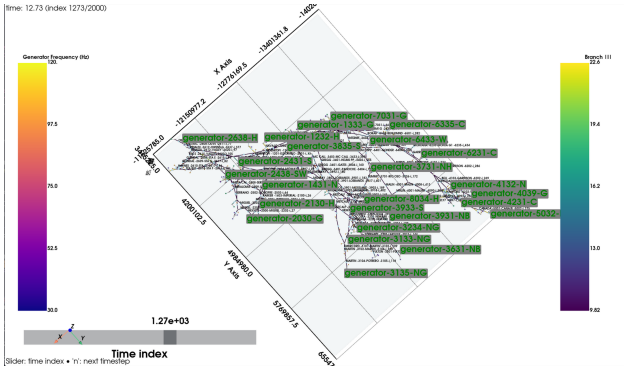

Figure 2: A top-down view of the 3D implementation.

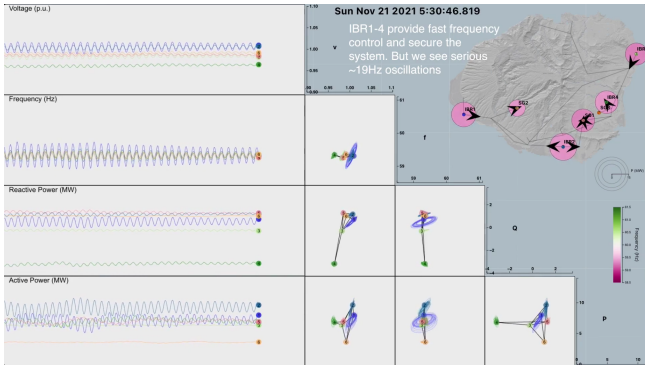

Figure 3: A screenshot of the 2D implementation.

	User 1	User 2	User 3	User 4	User 5	User 6	User 7	User 8
Generator Lolation Time (s)								
Time to seek to target index (s)								
Identified Frequency	100	120						
Actual Frequency (Hz)	120	120	120	120	120	120	120	120
Frequency Percent Error	**0.16**	**0.0**						
Identified Branch Current (a)	110	125						
Actual Branch Current (a)	123	123	123	123	123	123	123	123
Current Percent Error	**0.11**	**0.02**						

Figure 4: 3D Software Results

	User 1	User 2	User 3	User 4	User 5	User 6	User 7	User 8
Generator Lolation Time (s)								
Time to seek to target index (s)								
Identified Frequency								
Actual Frequency (Hz)	120	120	120	120	120	120	120	120
Frequency Percent Error								
Identified Branch Current (a)								
Actual Branch Current (a)	123	123	123	123	123	123	123	123
Current Percent Error								

Figure 5: 2D Software Results

ACKNOWLEDGEMENTS

The authors wish to thank Dr. Kenny Gruchalla from the National Renewable Energy Lab (NREL), as well as Rahul Kadavil and Dr. Tung Lam Nguyen from the New York Power Authority (NYPA).

REFERENCES

[1] K. Gruchalla, S. Molnar, and G. Johnson. Reevaluating contour visualizations for power systems data. *IEEE Transactions on Smart Grid*, 14(6):4877–4887, 2023.

[2] G. Johnson, K. Gruchalla, M. Ingram, N. Guba, R. Cruickshank, and S. Caruso. Vis-saga: Visual analytics for situational awareness of grid anomalies. In *2023 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pages 17–21. IEEE, 2023.

[3] G. Johnson, S. Molnar, N. Brunhart-Lupo, and K. Gruchalla. Architecture for web-based visualization of large-scale energy domains. In *2024 IEEE Workshop on Energy Data Visualization (EnergyVis)*, pages 46–51. IEEE, 2024.

[4] T. J. Overbye and J. D. Weber. Visualization of power system data. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, pages 7–pp. IEEE, 2000.

[5] P. C. Wong, K. Schneider, P. Mackey, H. Foote, G. Chin Jr, R. Guttromson, and J. Thomas. A novel visualization technique for electric power grid analytics. *IEEE Transactions on Visualization and Computer Graphics*, 15(3):410–423, 2009.

[6] J. Zhu, E. Zhuang, C. Ivanov, and Z. Yao. A data-driven approach to interactive visualization of power systems. *IEEE Transactions on Power Systems*, 26(4):2539–2546, 2011.

7 INDIVIDUAL CONTRIBUTIONS

7.1 Yonas

As aa Co-PI on this project, I handled the evaluation phase of the project as well as the management of the logistical information. I helped with coordination and communication between our group and Kenny and Rahul. I helped prune the WECC CSV metadata so that it was compatible with both our 3D PyVista Visualization as well as the 2D interactive oscillation dashboard. I also created the final presentation PowerPoint that our group will eventually present as well as providing weekly updates regarding the project's progress.

7.2 Aidan

As PI of this project, I handled the novel contribution (i.e. the 3D software). This involved developing the codebase (with feedback from team-mates), testing the software and communicating with collaborators about their needs/wants. I also designed experiments and methods for the analysis of our new software. In addition to developing code, I also developed a google form survey that was key for gathering results. Finally, I have written user-friendly documentation for our users and ensured our 3D software can be deployed on any machine that can run python and conda.

7.3 Korey

As a Co-PI of this project, my responsibilities fell within the realm of the 2d baseline visualization for comparison and evaluation of our proposed 3d visualization. This looked like coordinating with NYPA and NREL to gain access to the codebase that they demonstrated in an early meeting, and configuring the method, as well as connecting it to the WECC data that we are using. Then it looked like deploying and testing the configuration and establishing the evaluation for the 2D implementation, and packaging things for shipping to test users.

VisArch: An Interactive Dashboard for Comparing LiDAR Preprocessing Methods at Chachapoya Settlements

Charly Castillo* PI
Sir Dylan Allotey† Co-PI
Parker VanValkenburgh‡ Collaborator

ABSTRACT

We present *VisArch*, a three-part visualization for exploring how LiDAR preprocessing affects archaeological visibility. The approach improves upon standard visibility modeling, which typically relies on a single cleaned surface, making it difficult to see how alternative preprocessing methods change results. *VisArch* addresses this issue with side-by-side views of raw, Random Forest, and XGBoost-classified surfaces and controls for toggling preprocessing modes and inspecting visibility, making the interpretive consequences of preprocessing choices more explicit. The results suggest potential analytical value for archaeologists and may improve clarity for non-expert users.

Keywords: Archaeological visualization, Machine learning, Spatial analysis, Interpretability, LiDAR.

1 INTRODUCTION

The Chachapoya were a highland Andean culture of the Late Intermediate Period (circa 1100–1450 C.E.), known for their hilltop settlements characterized by dense clusters of circular stone houses [3]. Light Detection and Ranging (LiDAR) uses automated laser pulses to produce high-density 3D models and is now widely applied to these sites, including those hidden beneath heavy vegetation [3]. LiDAR datasets can be processed through manual, semi-automated, or machine learning (ML) methods. However, manual labeling is slow, difficult to reproduce, and dependent on individual expertise, limiting how quickly large datasets can be analyzed [4]. Thus, ML is increasingly used in archaeological data analysis.

By comparing raw, Random Forest, and XGBoost-based surfaces and their associated visibility networks, this research clarifies how preprocessing decisions alter interpretations of Chachapoya architecture. At the same time, highlighting previously obscured structures contributes to broader heritage and tourism discussions. By drawing attention to lesser-known Peruvian sites beyond the overrun Machu Picchu, the project reinforces the value of investing in LiDAR analysis at understudied sites.

2 RELATED WORK

While ML has entered archaeological LiDAR workflows, we know little about how specific classifiers and preprocessing pipelines reshape the surfaces archaeologists interpret. Most work reports accuracy for a single chosen model, but says less about how different modeling choices might change what appears as architecture, vegetation, or noise. This gap makes it difficult to judge the impact of preprocessing on visibility analysis and site interpretation.

Previous research has highlighted the potential of these methods. Abate et al. use a Random Forest classifier in CloudCompare to segment LiDAR data from an unmanned aerial system at the Kastrí-Pandosia site in Epirus, Greece [1]. Cerrillo-Cuenca and Bueno-Ramírez apply XGBoost to digital elevation models (DEMs) from the Valdecañas Reservoir in Cáceres, Spain, to predict zones of erosion and sedimentation [2]. Both studies show that ML can improve detection, but they treat the resulting surfaces as fixed products. This study builds on this work by directly comparing Random Forest and XGBoost preprocessing at Ollape and beginning to examine how those choices affect intra-site visibility. Unlike previous work, we focus on making these differences explicitly visible to support archaeological interpretation.

3 METHODOLOGY

This study utilized airborne LiDAR data from the Chachapoya site of Ollape in present-day Peru, which contains approximately 240 architecture features. Our workflow included data labeling, classification, surface generation, visibility modeling, and dashboard design. First, the full point cloud was manually labeled into two classes: Vegetation (1) and Other (0), with the latter including structures, ground, water, and shadows. In total, the dataset included approximately 4.83 million vegetation points and 0.48 million non-vegetation points.

Then, we trained an XGBoost classification model using the following point-level features: Z coordinates, intensity, return number, number of returns, and RGB values. For comparison, we implemented a Random Forest model from Abate et al. We chose this approach because their method is open-source, implemented in CloudCompare's 3DMASC plugin, and has been used to detect archaeological features in areas covered by tree vegetation [1]. However, because it was designed and tuned for the Kastrí-Pandosia site, we treat the model as a conceptual benchmark and retrain it on the Ollape labels.

Each classifier was used to generate a surface model by removing vegetation points, alongside an unfiltered surface from the raw point cloud. First, we removed all points with predicted vegetation probability ≥ 0.5 and rasterized those remaining into 0.5 m digital terrain models (DTMs) using the `knnidw` spatial interpolation algorithm. We also produced 0.5 m digital surface models (DSMs) with the `p2r` points-to-raster algorithm in `rasterize_canopy` and derived canopy height models (CHMs) by subtracting each DTM from its corresponding DSM. Then, structure-to-structure visibility was computed for each surface. Lastly, these outputs were visualized using side-by-side comparisons and interactive layer controls.

4 RESULTS

Our classification results show clear differences between the RF baseline and the XGBoost models for the Ollape dataset. For vegetation, RF attains a precision of 0.68, a recall of 0.79, and an F1 of 0.73, whereas XGBoost with random sampling reaches a precision of 0.9892, a recall of 0.9955, and an F1 of 0.9924. The "Other" class shows a similar pattern, with RF scoring 0.70 precision, 0.57 recall, and 0.63 F1, compared with 0.9513, 0.8904, and 0.9198 for XGBoost with random sampling. In both classes, the boosted model yields higher values on all three metrics reported in Table 1 relative to the RF baseline.

*e-mail: charly_castillo@brown.edu
†e-mail: sir_dylan_allotey@brown.edu
‡e-mail: parker_vanvalkenburgh@brown.edu

To examine spatial generalization, we trained a second XGBoost model using a spatially blocked split based on 5 m by 5 m tiles. In this setting, vegetation precision is 0.9891, recall is 0.9957, and F1 is 0.9924, which are effectively unchanged relative to the random split. For the "Other" class, precision is 0.9531, recall is 0.8896, and F1 is 0.9203. Compared with the random split, the tile split produces a small trade-off for the "Other" class: recall decreases slightly, while precision and F1 increase slightly.

Table 1: Classification performance at Ollape

Class	Metric	RF	XGB (random)	XGB (tile)
Vegetation (1)	Precision	0.68	0.9892	0.9891
	Recall	0.79	0.9955	0.9957
	F1	0.73	0.9924	0.9924
Other (0)	Precision	0.70	0.9513	0.9531
	Recall	0.57	0.8904	0.8896
	F1	0.63	0.9198	0.9203

Note. *XGB (random) and XGB (tile) denote models trained using a standard random point-based split and a spatially blocked split based on 5 m × 5 m tiles, respectively.*

The *VisArch* dashboard organizes the classification and surface outputs into a single interactive view. A control panel on the left contains a dropdown menu that selects the preprocessing method, including the raw surface, the Random Forest baseline, and the random XGBoost variant. Users can zoom in on the map to focus on specific areas of Ollape. Radio buttons allow the user to switch between DTM, DSM, and CHM, and text under each label summarizes what that surface represents. Beneath the controls, a metrics table reports precision, recall, and F1 for vegetation and other classes for each method, matching the values in Table 1. The main panel displays the chosen raster layer using a continuous color scale, so changing either the technique or surface type immediately updates the map with the corresponding model for Ollape.

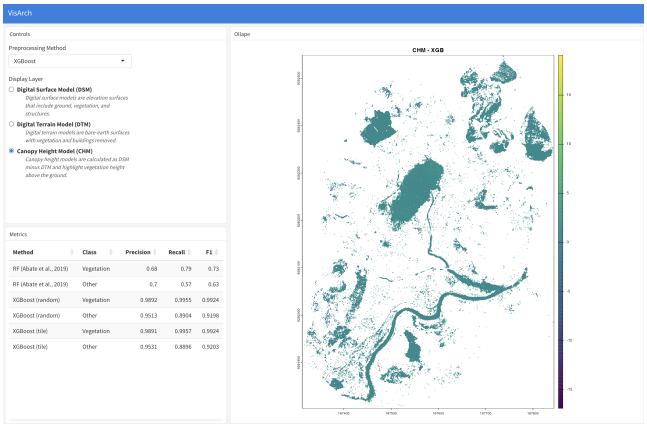

Figure 1. *VisArch dashboard for exploring LiDAR preprocessing outcomes at Ollape, showing the control panel, summary metrics, and an interactive map view of a selected surface.*

We are also adding toggles that encode model confidence and structure-to-structure visibility, including a confidence overlay that uses green for more confident classifications and yellow and red for less confident classifications, once appropriate thresholds have been defined.

5 Evaluation

Then, we conducted a non-expert study with six participants. All participants were unfamiliar with LiDAR and had no background in archaeology. Part 1 of the survey presented images of Ollape generated by the different preprocessing methods and asked participants to compare them. XX of 6 participants chose the RF-derived image as the easiest to understand, while XX of 6 preferred the XGBoost version for identifying possible structures. The raw surface was most often described as cluttered, with participants noting issues such as "XX" or "XX."

Part 2 focused on the dashboard layout using Likert-style questions on a 1-5 scale. Ratings indicated that the dashboard layout was generally understandable, with a mean clarity score of M = XX (SD = XX). Participants also felt that comparing the three preprocessing methods within the same view was reasonably feasible, with M = XX (SD = XX), although several comments pointed to specific uncertainties, such as "XX" or "XX".

Placeholder for survey result charts.

Figure 2. *Interpretation of the survey result charts will go here.*

Lastly, Part 3 assessed how people directly recognize and delineate non-vegetation features and how accurately they do so. The drawing task yielded XX returned sketches. Of these, XX outlines closely matched the reference non-vegetation regions, XX overshot the target areas by including extra surroundings, and XX undershot them by missing portions of the structures or ground that we labeled as "Other."

6 Conclusion

This study helps clarify how LiDAR preprocessing choices influence interpretations of Ollape. XGBoost consistently outperforms the Random Forest baseline for both Vegetation and Other classes, and this advantage remains stable under random and spatially blocked sampling. These results suggest that boosting algorithms can provide more reliable vegetation removal. However, the small changes in the "Other" class under tile sampling highlight the need to continue checking spatial generalization in complex terrain.

Beyond raw performance, our surface models and dashboard show how different classifiers reshape what users see. By deriving and placing DSMs, DTMs, and CHMs in a single interactive view, we can observe when structures change during preprocessing choices. The non-expert study indicates that these visual differences are noticeable to viewers without archaeological training, and that the dashboard helps them compare versions of the site and reason about where structures might be.

Feedback from Dr. VanValkenburgh suggests that *VisArch* could be helpful for XX. Together, these results point toward a workflow in which archaeologists routinely inspect multiple preprocessing pipelines rather than treating a single output as fixed. Future work should apply this approach to additional Chachapoya sites to test whether the patterns observed at Ollape generalize across different topographies.

ACKNOWLEDGEMENTS

The authors would like to thank Professor David Laidlaw and their collaborator, Parker VanValkenburgh, for their guidance and support throughout this project. The authors also thank their classmates for their feedback throughout the research process.

REFERENCES

[1] N. Abate et al. An open-source machine learning–based methodological approach for processing high-resolution uas lidar data in archaeological contexts: A case study from epirus, greece. *Journal of Archaeological Method and Theory*, 32:1–34, 2025.

[2] E. Cerrillo-Cuenca and P. Bueno-Ramírez. Predictive archaeological risk assessment at reservoirs with multitemporal lidar and machine learning (xgboost): The case of valdecañas reservoir (spain). *Remote Sensing*, 17:1–32, 2025.

[3] P. VanValkenburgh et al. Lasers without lost cities: Using drone lidar to capture architectural complexityat kuelap, amazonas, peru. *Journal of Field Archaeology*, 45:575–588, 2020.

[4] B. Štular et al. Airborne lidar point cloud processing for archaeology. pipeline and qgis toolbox. *Remote Sensing*, 13:1–37, 2021.

7 CHARLY CASTILLO

7.1 Intellectual Contributions

As the Primary Investigator, I established the overall research plan by defining our main question about how well XGBoost classification performs in identifying vegetation compared with the pre-existing RF method. I reviewed relevant work and used it to decide which variables to use as features and which performance metrics and error analyses would be most informative. I also planned how to best incorporate the survey results into the project to assess our dashboard's practical usefulness. Lastly, throughout the semester, I led the interpretation of our results in progress report presentations.

7.2 Practical Contributions

I wrote all of the project's R scripts. I implemented the full classification pipeline, including reading in the labeled data, building training and test sets, fitting the XGBoost model, tuning its parameters, computing summary statistics, and applying the final model to the full dataset to produce cleaned `.las` files. Additionally, I designed the user survey, distributed it to study participants, and analyzed the responses for inclusion in the results. I substantially revised the project README to clearly document how to run the code and reproduce our outputs. In addition, I created and updated most of our presentation materials, including progress-report slides, and wrote all of the final paper.

8 SIR DYLAN ALLOTEY

8.1 Intellectual Contributions

I contributed intellectually by helping coordinate meetings and facilitating communication with our collaborator and segmenting our data for training. I also suggested points to emphasize on our slides to present our main findings clearly.

8.2 Practical Contributions

I handled cleaning and organizing the raw LiDAR data and labeling the segments as Vegetation or Other. Additionally, I ran 3DMASC using the agreed-upon settings and saved its outputs for comparison with those from our XGBoost method. I drafted an initial version of the project README and prepared preliminary, which was later refined for the final presentation.

OLaVR: Oceanographic Layers in Virtual Reality

Juliana Escuti*

Brown University

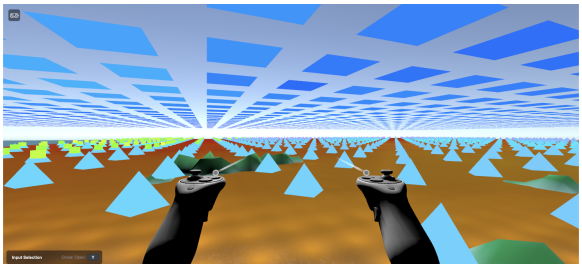

Figure 1: Four variables visualized in Fiji: Temperature, Sea Level Anomaly, Salinity, and Windspeed

ABSTRACT

Traditional methods of visualizing multivariate oceanographic data involve layering data points over a 2D map or comparing two 2D maps. OLaVR (Oceanographic Layers in Virtual Reality) is a novel method of visualizing multivariate oceanographic data in 3D. It uses decals to represent different data sources. This will improve accuracy in evaluation and decrease perceptual effort in comparison to CoastWatch.

Keywords: Virtual reality, multifield visualization, oceanography.

1 INTRODUCTION

Understanding complex ocean feedback loops requires analysis of multiple data fields. One method of visualizing these datasets, called a coordinated multi-view approach, described by Kehrer et al [3], involves comparing multiple graphs and maps, displaying coordinated data from multiple different sources. This method has been proven to aid in hypothesis generation, but has a disadvantage in that the viewer of these linked views must move their eyes back and forth between visualizations to compare data at a certain point. Lipsa et al classify multifield visualization as a current challenge for visualization in the physical sciences [2]. They note how the coordinated multi-view approach does not have a focus on mapping visualizations of natural phenomenon in physical space. Including multiple dimensions of data on a map can increase visual clutter and decrease the effectiveness of a visualization. Lipsa et al establish that visualizing multiple fields in an accurate way simultaneously to examine relationships between fields is the goal of such a visualization.

Current state of the art visualizations include Ocean Data Viewer (ODV) [6] and NOAA CoastWatch, available through the NOAA website. Both of these visualizations are desktop based and primarily 2D. This makes the multifield visualization problem described by Lipsa et al inherent to these two visualizations. One solution is reducing the dimensionality of the data by taking it from a 2 dimensional map to a 1 dimensional plot along a single line of longitude or latitude. This can obscure important spatial patterns in the data, like currents, which do not normally lie cleanly on a single line of longitude or latitude [5]. To remedy this, Nobre developed a web based tool that can visualize multidimensional oceanographic datasets in 2D, representing points through color and tracing contours in the data [4]. While this strategy was deemed effective and accurate, it is still two dimensional, and can only visualize on field at a time, like temperature. Another strategy involves utilizing three dimensions, and a type of visual glyph called a "decal" because it sticks to the surface of three dimensional meshes. In a paper by Rocha, this method was shown to be effective for three data fields, density, salinity, and ocean currents. OLaVR builds off the most successful aspects of these visualizations, decals and color maps.

*e-mail: juliana_escuti@brown.edu

2 OLAVR DESIGN

OLaVR uses decals to layer oceanographic data fields on a 3D surface representing the topography and bathymetry at a certain location. These decals include: points, continuous shading, and various kinds of meshes. Typical mesh shapes, such as spheres and cubes were chosen, alongside unconventional meshes such as capsules, eggs, and prisms. These unconventional meshes were chosen because due to their tapered shape, they could allow for a better view of distant data in a perspective camera view. While most of the decals lie on the 3D surface, they can also be positioned along the Y axis as appropriate or for maximum visibility. Users are able to navigate in this 3D space in virtual reality.

2.1 Implementation

OLaVR was developed in Unity 6.3 using C#. Unity was chosen for its ease of integration with virtual reality development pipelines [7]. To aid in performance, rather than rendering thousands of individual meshes, OLaVR uses point clouds. Out of the box, Unity does not support point cloud visualization. However, Unity's particle system can be used to generate point clouds using a simple C# script that reads data and spawns a particle for each data point. The shape, location, and color of each particle can be customized, allowing the particles to fit the data. The location of each particle matches the geospatial coordinates of each data point, and the color corresponds to the data value.

The data values are normalized before visualization to ensure the color map shows a visible change in value. A perceptually uniform color space had to be used to avoid subtle changes in perceived hue, saturation, and brightness that do not match patterns in the data [1]. Several color maps are available to users: sequential and diverging.

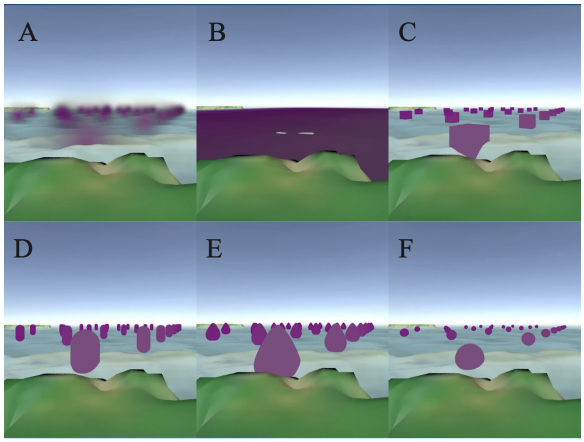

Figure 2: Individual Decals. A- Points, B- Continuous, C- Cube Mesh, D- Capsule Mesh, E- Egg Mesh, F- Sphere Mesh

These color maps are implemented using a custom shader script in C#.

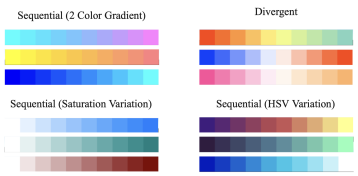

Figure 3: Sample of Color Map Options.

3 Evaluation

3.1 User Study

Two user studies were conducted to determine the effectiveness of OLaVR. The first user study aimed to test the understandability of each decal. A group of users were selected who were familiar with the Outer Banks island chain off the coast of North Carolina, with varying levels of familiarity with oceanography. For each decal, the coarse grained accuracy (over wide patterns in the data), fine grained accuracy (value at each point), and perceptual effort required to understand each decal were tested. Results are shown in Table 1. The accuracy scores were determined by averaging each user's score. All 6 decals scored equally well for coarse grained accuracy. Points scored the best for fine grained accuracy, and spheres scored the worst. Users were asked to score how much perceptual effort each visualization required, and those scores were averaged to produce the effort score. Users said the continuous visualization required the least effort to understand, while cubes scored the most.

The second user study compared OLaVR to CoastWatch. OLaVR scored better than CoastWatch for all sections, except for the section with three variables. Confidence generally decreased as more variables were added, and effort increased. Surprisingly, users scored higher on the section with four values than the section with three variables. Strangely, users had a relatively high confidence level for relatively lower accuracy for Coastwatch, and relatively lower confidence for higher accuracy with OLaVR. This could be due to a lack of familiarity with this novel visualization.

Table 1: User Study 1

Decal	Coarse Accuracy(%)	Fine Accuracy(%)	Effort (%)
Point	100%	100%	42%
Continuous	100%	71%	32%
Cube Mesh	100%	85%	54%
Capsule Mesh	100%	71%	48%
Egg Mesh	100%	85%	51%
Sphere Mesh	100%	57%	48%

Table 2: User Study 2

	Accuracy(%)	Confidence(%)	Effort(%)
CoastWatch	75%	56%	70%
OLaVR (One Variable)	83.75%	52%	66%
OLaVR (Two Variables)	81.6%	71%	66%
OLaVR (Three Variables)	72.5%	52%	76%
OLaVR (Four Variables)	79.4%	42%	57%

3.2 Case Study

Collaborators Simon Su and Guangming Zheng, oceanographers at NIST, participated in a case study to determine the effectiveness of the visualization with experts. They found the proximity of the rendered meshes to continuously rendered data points to be advantageous, since multiple data fields could be visualized in the same image. They noted that the diverging color map made it easy to visualize fronts, although contouring major fronts with an outline would have helped even further. However, they pointed out that for areas with a sharp increase in temperature in a single spot, the normalization process made it hard to determine changes in temperatures elsewhere. They also noted that because the implementation used a point cloud, it would lend itself well to visualizing time based motion of particles.

4 Future Work

The case study revealed two areas of interest for future work using the OLaVR framework. The first is contouring. Fronts are areas where significant change in the data occurs rapidly, forming a distinct boundary. A gradient could be calculated to determine where the strongest front boundary in the data is, then that boundary could be rendered as another layer in the three dimensional space. Another area of exploration is the utilization of the particle system to show data that varies in velocity. For example, wind velocity can be visualized as an arrow, but using the particle system, it could also be visualized as a point moving through space.

5 Conclusion

Overall, this project was successful. Users, both experts and novices at oceanography, found the visualization to be immersive and visually compelling. Although accuracy increased overall for OLaVR, confidence in correct answers remains low. While there are still open threads, such as increasing user confidence and implementing the features suggested in the case study, OLaVR achieved its stated goals.

Acknowledgements

The author wishes to thank Simon Su and Guangming Zheng, and David Laidlaw.

REFERENCES

[1] A. P. Bernice E. Rogowitz, Alan D. Kalvin and A. Cohen. Image which trajectories through which perceptually uniform color spaces produce appropriate colors scales for interval data? *Imaging Science and Technology*, 1999.

[2] S. J. C. J. C. R. R. W. Dan R. Lipsa, Robert S. Laramee and H. Michelle A. Borkin; Pfister. Hypothesis generation in climate research with interactive visual data exploration. *IEEE*, 2008.

[3] P. M. H. D. A. S. Johannes Kehrer, Florian Ladstädter and H. Hauser. Visualization for the physical sciences. *Computer Graphics Forum*, 2012.

[4] C. Nobre. Oceanpaths: Visualizing multivariate oceanography data. *Harvard University*, 2016.

[5] A. M. Paul Spence, John C. Fyfe and A. J. Weaver. Southern ocean response to strengthening winds in an eddy-permitting global climate model. *Journal of Climate*, 2010.

[6] R. Schlitzer. Data analysis and visualization with ocean data view. *CMOS Bulletin*, 2015.

[7] Unity Technologies. *Unity 6.3 User Manual*, December 2025.

OHCA-Vis: Interactive Visualization of Determinants in Pediatric Out-of-Hospital Cardiac Arrest

Ryan Huang* Haonan Zhang† Latha Ganti‡

Brown University

Abstract

We present a novel method called *OHCA-Vis* for visualizing pediatric out-of-hospital cardiac arrest (OHCA) data using interactive alluvial plots. State-of-the-art methods—like registry dashboards and static alluvial plots—can become overly convoluted and are not conducive to interpreting key insights, such as which variables (e.g. drug administered) are most important to survival. *OHCA-Vis* addresses these limitations with two features: 1) alluvial plot interactivity and 2) color-encoded feature importance. Users are able to interactively select and reorder variables of interest and visualize their importance from color-encoded nodes based on Gini scores of a Random Forest classifier on the selected variables. User study results suggest that this approach improves interpretability and insight generation for OHCA data.

Keywords: Alluvial diagrams, interactive visualization, pediatric out-of-hospital cardiac arrest, emergency medical services, feature importance

1 Introduction

Pediatric out-of-hospital cardiac arrest (OHCA) is rare but highly lethal, with survival rates generally below 10% [7, 6]. EMS registries capture rich information about demographics, interventions, and clinical outcomes, yet current reporting tools rely largely on static tables and simple charts that make it difficult to understand how multiple factors jointly shape patient trajectories [5].

Static alluvial diagrams provide an intuitive way to visualize categorical flows but can quickly become cluttered when many variables or categories are shown [2, 8]. At the same time, machine-learning models can estimate feature importance for predicting survival, but their outputs typically appear in separate views, forcing users to manually connect them back to patient pathways.

We present *OHCA-Vis*, an interactive alluvial visualization system designed to support exploratory analysis of pediatric OHCA data. The system enables users to reorder variables, reduce visual clutter through lightweight layout optimization, and view random-forest feature-importance scores directly encoded in node colors. Our goal is to offer a more interpretable and cohesive representation of how clinical and contextual factors influence outcomes.

2 Related Work

Previous work has focused on static dashboards, tables, and survival summaries that aggregate patient outcomes by individual factors such as age, sex, or intervention type [3, 4]. These methods are limited because they obscure interactions between variables and don't have features for exploration on how factors affect patient outcomes.

*e-mail: ryan_y_huang@brown.edu

†e-mail: haonan_zhang2@brown.edu

‡e-mail: latha_ganti@brown.edu

Alluvial and Sankey-style visualizations have been used to summarize categorical flows and transitions in various domains, including bibliometrics and demographic studies [10]. Moreover, alluvial plots have been used to evaluate factors in COVID-19 survival rates [9]. However, there has been no work in making these plots interactive or exploratory for users to evaluate and choose which variables to visualize. Prior studies show that increasing the number of flows reduces diagram interpretability [1]. With the amount of data that exists in EMS data, the number of flows would inevitably increase. The introduction of an editable and interpretable plot is hypothesized here to improve user understanding, as the plot would be easily simplified and user-modifiable.

3 Methodology

Our dataset contains 133 pediatric OHCA cases with 29 variables spanning demographics, event characteristics, and prehospital interventions.

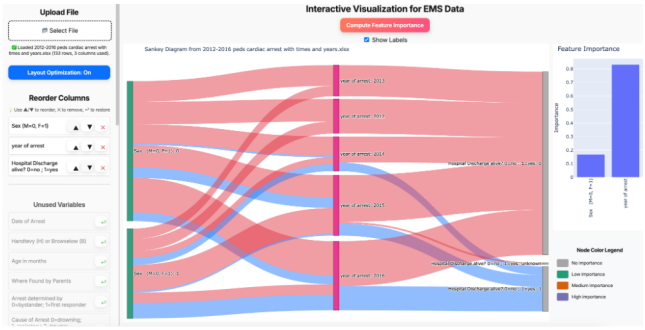

Figure 1: The OHCA-Vis interface, showing the variable-selection panel, interactive alluvial plot, and integrated Random Forest feature-importance view.

3.1 Data Preprocessing

To prepare the data for visualization, raw registry variables contained heterogeneous labels, missing values, and inconsistent category definitions. The resulting cleaned dataset was used directly by the interactive Dash application.

3.2 Alluvial Diagram Construction

Each user-selected variable forms a column in the alluvial diagram, with nodes representing category values and edges showing patient flows across variables, ending in the survival outcome. The diagram uses a simple ordering strategy to maintain a readable left-to-right progression of pathways, while node heights scale with category frequencies so that common trajectories occupy more visual space.

3.3 Interactive System Features

The interface allows users to upload datasets, select variables, reorder columns, and hide infrequent categories to simplify the diagram. Users can also toggle a layout optimization mode that automatically reduces visual clutter by adjusting column ordering and

minimizing edge crossings. Together, these interactions support exploratory analysis by enabling users to examine alternative pathway structures and investigate relationships between variables. The visualization updates immediately in response to user actions, allowing rapid comparison of different configurations.

3.4 Feature-Importance Integration

To augment interpretability beyond visual inspection, OHCA-Vis provides a model-driven feature-importance view. A Random Forest classifier (100 estimators) is trained using the currently selected variables, and normalized Gini importance scores are extracted. These scores are mapped to a perceptually uniform color scale applied to node backgrounds, allowing highly influential variables to stand out. Because the model retrains dynamically whenever variable selections change, users can observe how feature relevance shifts across different analytic scenarios.

4 Results

To assess *OHCA-Vis*, 15 participants, consisting of medical school students and faculty from the Warren Alpert Medical School, completed a series of task-based assessments designed to measure how effectively they can extract insights from the visualization compared to standard alluvial plots.

4.1 *OHCA-Vis* Improves Interpretation Accuracy

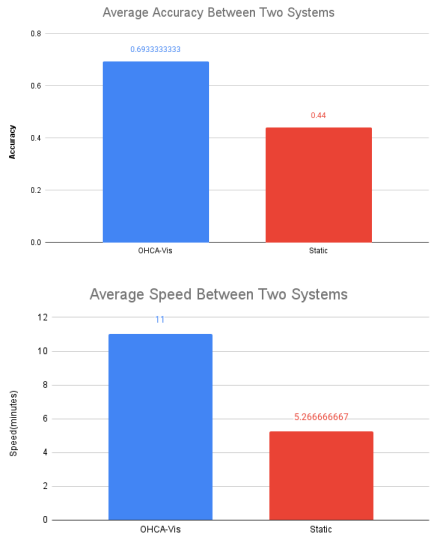

Figure 2: Task performance comparison. (Top) Average accuracy between the two systems. (Bottom) Average response time.

The assessment included multiple choice questions regarding variables and their correlation with survival rate. The speed and accuracy of answering these questions were then measured. Participants were determined to be able to answer the questions with 57% higher accuracy when using *OHCA-Vis* compared to the static plot.

However, we find that users require more time to answer the questions by 48% on average. We attribute this to getting familiar with the tool. While static alluvial plots were provided as a static image, participants had to navigate to our deployed application and learn to use the tool as they answered the questions.

4.2 *OHCA-Vis* Allows Users to Gather More Insights

The assessment also included a text box for users to list insights gained from the visualization. It was found that participants were

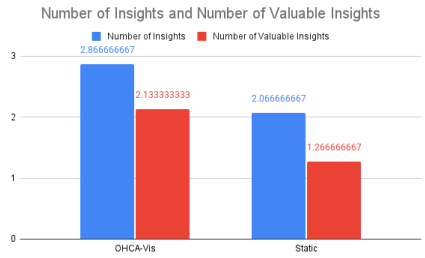

Figure 3: Average number of insights and valuable insights extracted using each system.

able to extract 38% more insights about the OHCA dataset using *OHCA-Vis* compared to static alluvial plots. Moreover, these insights were then assessed by an expert (our collaborator), who classified whether each insight was "valuable" or "not valuable". It was determined that participants using *OHCA-Vis* extracted 169% more valuable insights. Moreover, the proportion of valuable insights relative to total insights was 12.6% higher when using *OHCA-Vis*.

4.3 Users Preferred *OHCA-Vis* Over Static Alluvial Plots

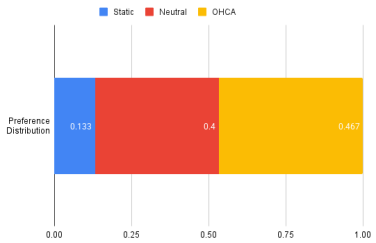

Figure 4: User preference distribution between static alluvial plots and *OHCA-Vis*.

Participant preferences in using the plots were also assessed. It was found that 46.7% of participants preferred using *OHCA-Vis*, 40% were indifferent, and 13.3% preferred using static alluvial plots—illustrating a slight preference for *OHCA-Vis*.

5 Conclusion

We presented *OHCA-Vis*, an interactive alluvial visualization system designed to support interpretation of pediatric out-of-hospital cardiac arrest data. By enabling users to reorder variables, explore patient pathways, and view integrated feature-importance cues, the system provides a more interpretable and flexible alternative to conventional static alluvial plots.

Our evaluation demonstrates that *OHCA-Vis* meaningfully improves users' analytical performance: participants achieved higher accuracy when answering survival-related questions, generated more total insights, and produced substantially more valuable insights, with a higher proportion of valuable-to-total insights. Participants also expressed a clear preference for the interactive system.

While promising, the current prototype has limitations, including the use of global feature-importance scores and the absence of subgroup-specific explanations. Future work will focus on expanding filtering capabilities, incorporating additional model-based explanation methods, and conducting formal task-based studies with EMS providers and pediatric emergency physicians. We anticipate that the design principles underlying *OHCA-Vis* can generalize to other clinical and registry-based contexts where outcomes emerge from complex, multi-step care pathways.

ACKNOWLEDGEMENTS

The authors wish to thank Professor David Laidlaw and their collaborator, Latha Ganti, for their guidance throughout the project.

REFERENCES

[1] A. Arunkumar. Bayesian modelling of alluvial diagram complexity. manuscript, Arizona, United States of America, 2021.

[2] D. Catalogue. Chart snapshot: Alluvial diagrams. https://datavizcatalogue.com/blog/chart-snapshot-alluvial-diagrams-examples/, 2024. Accessed 2025-12-09.

[3] J. Chishtie. Interactive visualization applications in population health and health services research: Systematic scoping review. manuscript, Toronto, Canada, 2022.

[4] O. M. Christen. Dashboard visualization of information for emergency medical services. manuscript, Biel, Switzerland, 2020.

[5] O. M. Christen et al. Emergency medical services information visualization. *Studies in Health Technology and Informatics*, 272:81–84, 2020.

[6] E. L. Fink et al. Unchanged pediatric out-of-hospital cardiac arrest incidence and survival rates with regional variation in north america. *Resuscitation*, 107:121–128, 2016.

[7] A. Kämäräinen et al. Out-of-hospital cardiac arrests in children. *Journal of Emergencies, Trauma and Shock*, 3(3):273–276, 2010.

[8] B. Koneswarakantha. Data exploration with alluvial plots: easyalluvial. https://erblast.github.io/easyalluvial/articles/data_exploration.html, 2025. Accessed 2025-12-09.

[9] W. C. K.-T. T. Po-Tsung Yen, Tsair-Wei Chien. Using the alluvial diagram to display variable characteristics for covid-19 patients and research achievements on the topic of covid-19, epidemiology, pathogenesis, and vaccine (cepv): Bibliometric analysis. manuscript, Tainan, Taiwan, 2023.

[10] A. W. K. Yeung. Data visualization by alluvial diagrams for bibliometric reports, systematic reviews and meta-analyses. manuscript, Hong Kong, Hong Kong, 2018.

6 Ryan Huang Individual Contributions

6.1 Intellectual Contributions

As the Primary Investigator, I contributed to the design and implementation of the *OHCA-Vis* system, including interactive interface development, feature-importance integration, and preparation of the manuscript. I coordinated all project activities, maintained continuous communication with our clinical collaborator, Latha Ganti, and ensured scientific and clinical relevance throughout the system's development. Additionally, I designed and administered the user study, including task construction, participant coordination, and quantitative and qualitative analysis of user-study outcomes.

6.2 Practical Contributions

I implemented the core technical components of *OHCA-Vis*, including the Plotly application. I also designed the entire layout and core features, such as variable selection. I also implemented feature-importance visualization using a Random Forest classifier. I integrated this into the application layout and researched different color schemes for color-encoding (ultimately deciding on color-blind friendly colors) as shown in Figure 1.

7 Haonan Zhang Individual Contributions

7.1 Intellectual Contributions

As a collaborator on the project, I contributed to the conceptual development of the *OHCA-Vis* system, particularly in areas related to data processing, layout organization, and the design of the user study. I developed the initial user-study questionnaire, including the design and wording of the evaluation tasks and prompts. I also led the initial drafting of the manuscript, working with Ryan to refine the writing and align the narrative with the system's analytical goals. Throughout the project, I collaborated closely with Ryan in refining system behavior and incorporating feedback into both the manuscript and the project presentations.

7.2 Practical Contributions

My practical contributions focused on implementing key technical components aligned with our project milestones, including the data-processing workflow and the layout-optimization feature. The optimization module reduces edge crossings in the alluvial diagram using barycentric ordering and pairwise node-swapping techniques, improving the readability of flow patterns. I also deployed the *OHCA-Vis* system to a web server to enable remote participation during the user study. In addition, I produced visual summaries of user-study data to support the interpretation of participant performance and feedback.

Uncovering patterns of Institutional Trust and Perception

Shravya Munugala* Gordan Milovac†

Brown University

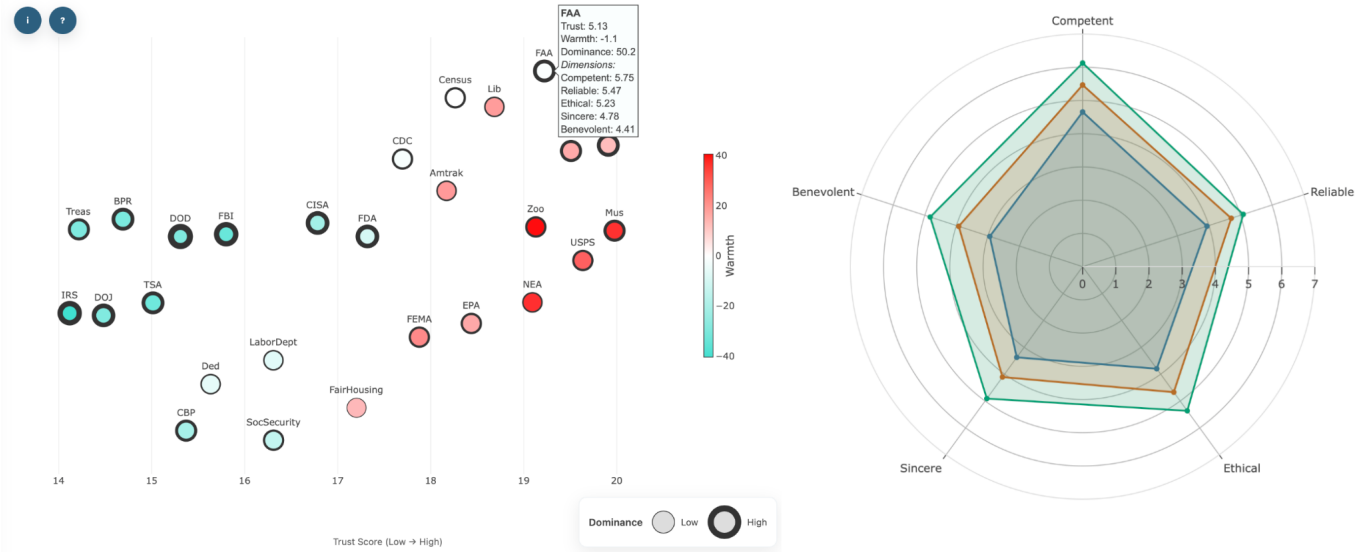

Figure 1: Our UMAP and Overlying Radial Graph approach

Abstract

We present an interactive visualization system for revealing patterns of public trust in government agencies through rich, multi-dimensional survey data. Traditional 2D charts often oversimplify these relationships and limit meaningful exploration. To address this, our system extends prior radial graph approaches with support for overlaid comparisons and introduces a dual-visualization strategy: a UMAP-based overview for clustering agencies by similarity and overlaid radial graphs for detailed, multi-attribute inspection of selected agencies. The system is deployed as a web platform for broad accessibility. We evaluated its effectiveness against baseline visualizations in a controlled study with 25 participants.

The empirical results demonstrated that our system improved overall efficiency, reducing task completion time by approximately 18.5% compared to the baseline visualizations, while maintaining high confidence and strong accuracy in similarity-based tasks. The system was perceived as easier to navigate, provided quicker access to relevant information, and supported a more streamlined workflow for observing complex patterns and relationships across agencies.

Keywords: Data Visualization, Uniform Manifold Approximation and Projection (UMAP), Dimensionality Reduction, Radial Graphs, Multi-Dimensional Data, Interactive Visualization, Usability Study, Human-Computer Interaction (HCI), Institutional Trust

*e-mail: shravya_munugala@brown.edu

†e-mail: gordan_milovac@brown.edu

1 Introduction

Public trust in governmental and institutional agencies is a critical component of civic engagement, policy compliance, and social cohesion. Accurately tracking and interpreting public sentiment across multiple dimensions is essential for researchers, policymakers, and the media. However, complex multi-attribute survey data is often visualized using simple bar charts or static 2D displays that fail to capture its multidimensional nature. These traditional approaches oversimplify patterns, limit exploration, and obscure relationships between agencies and trust dimensions.

To address this gap, this paper introduces an interactive, interpretable visualization system for multi-attribute trust data. We combine techniques from machine learning and human–computer interaction to provide richer insight into how government agencies are perceived across multiple factors. Our system adapts established visualization principles to create a tool that is both scientifically informative and practically accessible. The core contributions of this work are:

Dual-Visualization Interactive System: A deployed web platform integrating two complementary techniques. **UMAP** provides a dimensionality-reduced overview that organizes agencies by similarity across their full trust profiles, enabling high-level pattern discovery. **Interactive Overlying Radial Graphs** support direct, multi-attribute comparison of selected agencies.

Empirical Validation: A controlled user study with 25 participants quantitatively and qualitatively comparing our interactive system to conventional 2D baseline methods, evaluating task accuracy, speed, and perceived workload (NASA-TLX).

2 Related work

Effective visualization of multi-dimensional data is a persistent challenge in Information Visualization (InfoVis) and Data Science [3]. Datasets with five or more dimensions, such as agency trust profiles, exceed the limits of human perceptual capacity [6]. Traditional visualizations (e.g., bar charts, scatterplots) are typically limited to two or three variables, forcing analysts to rely on manual cycling of 2D views or data aggregation that obscures latent patterns [2].

To address this, glyph-based visualizations encode multi-attribute entities into single graphical objects. The radial graph (radar chart) is a popular glyph for up to 10 dimensions [1]. While strong for individual profile representation, radial graphs are weak for direct inter-entity comparison due to high cognitive load from evaluating overlapping or many separate plots [1]. Our work directly addresses this by introducing an interactive, overlying radial graph system for immediate, efficient comparison.

Dimensionality Reduction (DR) is another approach, projecting high-dimensional data into a lower-dimensional space while preserving structure. Modern nonlinear methods like t-SNE and UMAP have replaced older linear methods (MDS, PCA) [4]. UMAP is preferred for its better preservation of global structure and scalability, making it ideal for the high-level pattern discovery stage. We employ UMAP to create a similarity map of agencies, fulfilling the "overview first" step of the information seeking mantra [5].

Our system synthesizes these two distinct paradigms: DR for global pattern discovery and adapted glyphs for detailed comparison, creating a single interactive web platform that allows seamless movement from a global agency similarity overview (UMAP) to multi-attribute comparison (Overlying Radial Graphs).

3 Methodology

This paper details our novel visualization system for exploring multi-dimensional institutional trust data, emphasizing the iterative design process.

3.1 Data and Preprocessing

The underlying data contains public perception scores for approximately 30 U.S. government agencies across five key trust dimensions and two key style dimensions (Warmth and Dominance). Raw survey scores were normalized prior to visualization to ensure equal contribution from all dimensions to similarity calculations and visual encodings.

3.2 Dual-Visualization System Implementation

The system is a cohesive, interactive web platform integrating two distinct visualization components:

UMAP Similarity Map: We apply the UMAP algorithm to reduce the data (comprising five trust dimensions) to a two-dimensional scatterplot. The proximity between nodes represents the similarity of the agencies' full trust profiles. An alphabetical legend allows quick agency location and identification. This view serves as the system's Overview function [5].

Overlying Radial Graph Comparison: This component serves as the Details-on-Demand function. It adapts the classic radial graph, mapping each of the seven dimensions (five trust + Warmth/Dominance) to an axis. Our innovation allows users to dynamically select multiple agencies via a search-and-select list and render their graphs overlying one another, using color and transparency to facilitate direct, simultaneous comparison across all dimensions. This design allows for rapid assessment of where agency profiles align and diverge.

3.3 Pilot Study and Iterative Refinement

We conducted a pilot study with 10+ participants to validate and refine the initial design, comparing tasks performed on the interactive system (Visualization B) against the 2D baseline (Visualization A). Feedback led to substantial revisions:

- **UMAP Simplification:** Initial dual encodings (Warmth: color, Dominance: size) were confusing and created information overload. **Revision**: We removed these encodings, simplifying the UMAP to represent only the five trust dimensions, using an alphabetical legend for identification.
- **Usability and Terminology:** Participants found the original "blue = warm" color scale counterintuitive and struggled with abstract terminology. **Revision**: The Warmth scale was redesigned using a more intuitive light-blue-to-red gradient, and clearer explanations of the psychological constructs and UMAP rationale were added.
- **Interface Improvements:** Issues with the initial dropdown selector and cluttered overlaid radial graphs were noted. **Revision**: The dropdown was replaced with a robust search-and-select interface, and layout elements were adjusted for clarity.
- **Fairer Evaluation:** Difficulties accessing the baseline 2D plots raised fairness concerns. **Revision**: Baseline access was streamlined, task instructions were clarified, and the prior workload questions were replaced with the standardized NASA-TLX questionnaire.

These critical changes shaped the final system and study protocol for the final 25-participant evaluation.

4 User Study

We conducted an online user study with approximately 25 anonymized participants to evaluate how well users could interpret and compare multi-dimensional trust information using a baseline system (Visualization A) and our proposed system (Visualization B). The goal of the study was to assess how effectively each visualization supported tasks involving similarity judgments, trust-profile comparisons, and recognition of patterns across 27 agencies. Each visualization component was evaluated separately to isolate its individual contribution to user performance. In Visualization A, participants used a Warmth–Dominance scatter plot and individual radial graphs to answer targeted comparison questions. In Visualization B, they completed analogous tasks using overlaid radial graphs and a UMAP similarity map, which required selecting nearest neighbors, identifying balanced or unbalanced profiles, and interpreting overall trust structure. For both systems, participants also rated ease of use, confidence, and mental workload, and provided qualitative feedback to help identify the strengths and limitations of each visualization.

5 Results

Participants demonstrated noticeably better performance with the proposed system compared to the baseline. Tasks were completed about 20% faster on average (baseline: 9.08 minutes; proposed: 7.40 minutes), indicating that users were able to interpret and compare agencies more efficiently. Accuracy also improved: while the baseline produced several mid-range accuracy scores (some tasks around 55–75%), the proposed system consistently achieved high accuracy, with many tasks falling in the 95–100% range.

The NASA-TLX style ratings (shown in the radial graph/Figure 2) support these trends. Participants reported notably lower effort and higher ease of comparison with the proposed system. This led to significantly higher confidence in their answers. Frustration levels were comparable, but factors like mental demand and physical demand were both negligibly higher with the interactive tool,

suggesting the trade-off primarily improved usability and accuracy, rather than simply reducing cognitive load.

When asked to choose which approach they preferred overall, the majority of participants selected the proposed system, citing clearer comparisons and easier interpretation of multi-dimensional information. Taken together, these results show that the proposed visualization improved speed, accuracy, and user experience for understanding trust patterns across agencies.

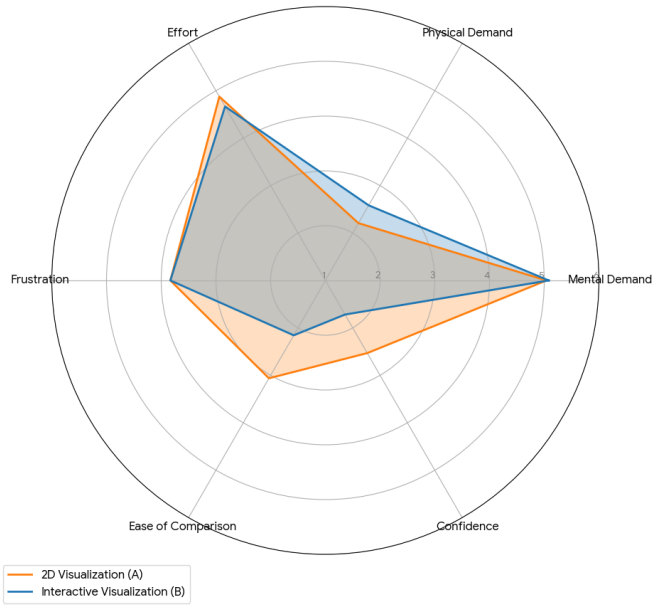

Figure 2: Perceived Workload (smaller = better).

6 Conclusion

This project introduced a new visualization system designed to help users interpret complex, multi-dimensional trust data more effectively. Through our user study, we found that the proposed system consistently supported faster, more accurate, and less cognitively demanding analysis compared to the baseline. Participants were able to identify similarities, differences, and overall trust patterns across agencies more efficiently, and most reported preferring the proposed system for its clarity and ease of comparison.

Together, these results demonstrate that integrating multi-agency overlays with a similarity-based embedding provides a more intuitive and powerful framework for understanding institutional trust data. The project highlights how thoughtful visualization design can meaningfully improve both analytical performance and user experience when working with high-dimensional information.

Acknowledgements

We would like to thank Dr. Bertram Malle for his invaluable collaboration and continuous support throughout this project. We are also grateful to Dr. David Laidlaw for his guidance and feedback over the course of the semester, which greatly strengthened the direction and execution of our work.

References

[1] F. Beck, M. Behrisch, and T. Schreck. Radial charts survey: Visualization, aggregation, and interaction. *IEEE Transactions on Visualization and Computer Graphics*, 23(10):2201–2218, Oct. 2017.

[2] J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey. *Graphical Methods for Data Analysis*. Wadsworth International Group, 1983.

[3] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI Magazine*, 17(3):37–54, 1996.

[4] L. McInnes, J. Healy, and J. Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

[5] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages (VL '96)*, pages 336–343, July 1996.

[6] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers, second edition, 2004.

CONTRIBUTIONS

Shravya Munugala

Intellectual Contributions

I initiated the project and defined its overall research direction, proposing the core idea of modeling multi-dimensional trust relationships through a structured, low-dimensional visualization framework. I developed the conceptual rationale for adopting a UMAP-based trust landscape by evaluating trade-offs in topology preservation, separability, and interpretability across dimensionality-reduction options. I designed the evaluation methodology, specifying the controlled comparison between the 2D baseline and our system, selecting quantitative metrics, and ensuring alignment in data exposure and task design. I also crafted the logic for non-obvious evaluation questions to probe genuine relational understanding, and structured the conceptual user workflow so that transitions across the embedding view, agency-level metrics, and comparison tasks remained cognitively coherent.

Technical Contributions

I contributed across the full technical stack of the project, spanning data preparation, visualization development, system configuration, and empirical analysis. I built the preprocessing pipeline for the trust-dimension dataset, handling normalization, aggregation, missing values, and preparing the data for the visualization and analysis steps. I developed the complete 2D baseline visualization website as a performance-optimized, low-latency interface with controlled visual encoding to serve as a methodologically rigorous comparison condition. I configured the UMAP-based trust-space representation, introduced an index-based alphabetical legend for scalable agency selection, and refined the dimensional structure by removing high-variance interpersonal attributes identified during pilot diagnostics. I managed the data and participant logs and carried out the quantitative analysis. I also contributed to writing sections of the final report and ensured that our work stayed aligned with collaborator feedback.

Gordan Milovac

Intellectual Contributions

My primary Intellectual Contributions were central to the successful design, refinement, and validation of the project's core output. I played a key role in the initial brainstorming and design, co-creating the innovative dual-visualization system that integrates the UMAP similarity map for pattern discovery with a new version of the radial graph. Specifically, I designed the concept of the Interactive Overlying Radial Graphs, making it possible to dynamically compare multiple entities, which fixed the established problem of using standard radial charts for profile comparisons. A key part of this was solving a major usability issue identified in the pilot study: I proposed and implemented the intuitive Ice-to-Fire color scale for the Warmth dimension, which fixed both the confusing coloring and the visibility issues with the Dominance factor. I also developed and ran the pilot study, using its findings to drive major revisions. This work helped ensure the final 25-participant study was strong, including adding the standard NASA Task Load Index (NASA-TLX) for measuring workload.

Technical Contributions

My Technical Contributions ensured the visualization system was fully built, deployed, and confirmed to be effective. I was responsible for developing and publishing the complete interactive web platform, including setting up the domain, which made our solution widely accessible. I built both the UMAP component, which required precise mapping of seven dimensions using color and size, and the dynamic rendering logic for the Overlying Radial Graphs. I made key technical improvements based on user feedback, such as replacing the initial difficult selection menu with a usable selection list. Beyond coding, I provided strong project support by creating and giving the weekly update presentations, managing collaborator communication smoothly, and contributing to the project's final materials by helping analyze the user study data and writing sections of the final report.

Socially-Responsible Visualization: Creating Context-Driven, Multi-View Clustering Methods for Population Genetics

Skylar Sargent Walters* Kyle Wisialowski† Alex Diaz-Papkovich Sohini Ramachandran

Brown University

Abstract

We present an interactive, multi-view PCA dashboard for human population genetics that centers interpretability and social responsibility in genomic visualization. The system links each individual's PCA position to concise gene-level annotations, enables simultaneous comparison across multiple principal component views, and supports group-level exploration through lasso-based selection and interpretation panels. In a within-subject study using chromosome 1 data, the interpretable dashboard improved quiz performance, enhanced users' understanding of PCA variability and genetic drivers, and reduced support for racialized, misinterpreting narratives relative to a static PCA baseline. These findings suggest that embedding interpretive context directly into clustering visuals can both aid exploratory analysis and strengthen robustness against deliberate misuse of population structure plots. The dashboard can be accessed here.

Keywords: PCA, population genetics, interpretability, socially-responsible computing, human-computer interaction.

1 Introduction

Visualizing the genetic relationships between individuals and populations is a key element of population genetics research [6, 4]. However, human genome data is very high-dimensional. As a result, scientists often employ dimensionality reduction techniques such as Principal Component Analysis (PCA) to create accessible visualizations. Here, each individual represents a point, and their position is determined by the pattern of single-nucleotide polymorphisms (SNPs), or mutations, over each individual's genome.

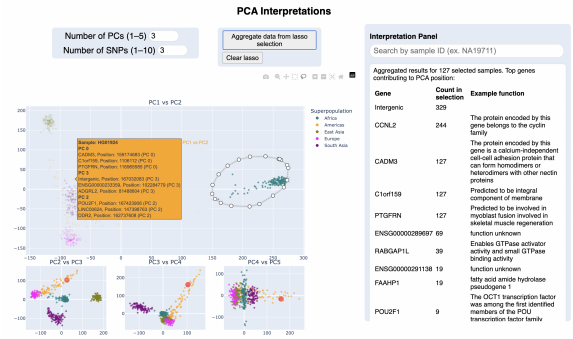

Figure 1: PCA Analysis dashoard includes hoverable gene interpretations, lasso capabilities to analyze driving differentiating factors in groups, multiview plots for tracking points across PCs, and an interpretation panel for extended responses.

*e-mail: skylar@brown.edu

†e-mail: kyle_wisialowski@brown.edu

However, PCA's emphasis on cluster differentiation creates misleading or even dangerous inferences. Given the abstract nature of the optimization algorithm, users may struggle to intuit meaning behind the distances between points; further, these distances change depending on the visualized principal components (PCs) [5]. As a result,malicious actors have manipulated PCA plots in support of claims about white supremacy and eugenics [2].

To address these challenges, we developed a dashboard to emphasizes plot sensitivity and expose the underlying genetic features that contribute to individual positioning (Fig. 1). This includes a number of novel elements. First, we have **multi-view visualization**, enabling users to view the same data point across multiple PCA plots; second, we have an **integrated interpretation panel** for exploration of biological function through a hover or a click; third, we have **subsample interpretations**, in which users can select a series of points to conduct a group interpretation.

2 Related Work

Past work in improving population-structure visualizations have highlighted the need for interpretability in clustering. Critically, PCA provides no explicit interpretation mechanism for understanding why individuals appear in particular positions within the embedding, and its views can change significantly depending on the PCs used.

Recent methods such as GAUDI have been applied to other clustering methods, such as UMAP, and incorporated integrated interpretation using SHAP values to quantify feature contributions [3].

To address limitations in single-view PCA, prior work has explored integrating multiple datasets or representations into a shared low-dimensional space. Approaches such as joint PCA aim to extract components that summarize variation across several views, reducing view-specific distortions and emphasizing shared structure [9]. Although this improves stability, it still lacks mechanisms for interpreting feature contributions within the embedded space.

3 Methodology

This study used chromosome 1 data from Phase 3 of the 1000 Genomes Project, which contains 6 million SNPs across 3,500 individuals [1]. To visualize this data, we filtered the genomes based on read quality, then constructed a genotype matrix with dimensions `number of SNPs x number of individuals`. Thus, we could characterize each individual as a vector based on the presence or absence of each SNP.

To find the most contributing genes for each position, we extracted the absolute value of the PC loading score for each SNP. These SNPs were then assigned gene identity and function with the `pyranges` and `mygene` packages [7, 8]. We then paraphrased these captions with GPT-5 to shorten the text (prompts here).

We surveyed Brown University's Population Genetics course (BIOL 1430), to assess the performance of our visualization. Where relevant, we used within-group comparisons between our interactive plot and the baseline static plots; higher, positive values indicate a preference for the interactive system, while lower, negative values indicate a preference for the static system.

4 RESULTS

4.1 User Performance Improved in Interpretable Plot

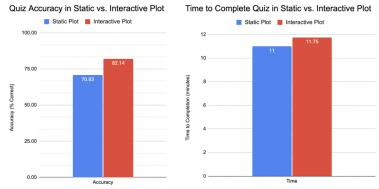

Figure 2: While the quiz for the interpretable visualization took an average of 0.75 minutes longer, users had higher performance than on the baseline visualization.

To understand how well users can navigate the visualization interface for exploration and discovery, we had users complete a short quiz. These questions covered basic interpretation of gene functions, gene relationships to PC clustering, relative point locations across PCs, and comparisons of relatedness.

When using the interpretable software, users scored almost 12% higher than when using the standard visual. However, it took 45 seconds longer to complete the quiz, which we attribute to taking time to get accustomed to the dashboard functionality (Fig. 2).

4.2 Users preferred the interactive visualization in use cases around pattern recognition, accessibility, exploration, and gene contribution.

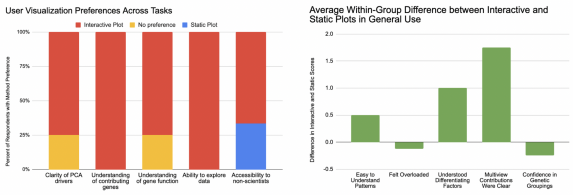

Figure 3: **Left** Users had the highest preference for the interpretable plot in all metrics. **Right** In a within-group study, users found the interpretable plot favorable across key metrics.

We explored preferences of users across tasks to better understand key abilities and limitations of the interpretable method as compared to the baseline. Across all tasks, users had the highest preference for the interpretable method, indicating that it has stronger abilities in explaining contributing genes and their functions, depicting PCA drivers, providing exploration, and being accessible to non-scientists. Critically, in the interactive plot, they were less confident in the genetic groupings; this is a key result, as we want users to understand the variation intrinsic to PCA.

4.3 Interactive plot reinforces the flexibility and variable view of PCA visualization

A key element of ensuring individuals de-center the division of peoples in PCA is understanding the arbitrariness inherent to PCA; looking at two different PCs will provide entirely different views of clusters. To assess this, we had users conduct a Likert survey for each visual. We found that the interpretable method was 1) more transparent about PCA variability, 2) enforced an understanding of what genes drove each view, and 3) prevented the over-interpretation of distances (Fig. 4). These results show promise in conveying that our visual provides a less rigid approach to PCA interpretation, thus promoting the togetherness of groups as opposed to separation.

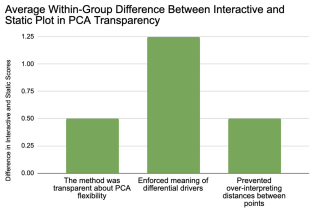

Figure 4: Users found the interactive plot scored better for PCA transparency, explanation of drivers, and preventing over-interpretation of distances.

4.4 The interpretation-based method enforces caution in user decision-making and strengthens robustness to misinterpretation

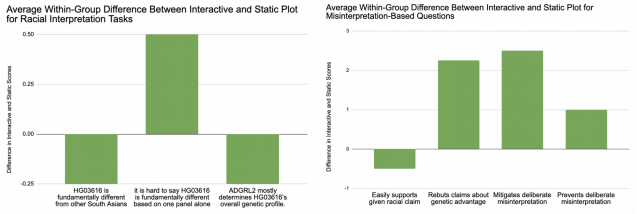

Figure 5: User survey results in misinterpretation settings

To study how users may **unintentionally** draw biased conclusions from each visualization, we asked users to support or refute three statements about individual HG03616. For the first statement, the interactive plot more strongly refuted HG03616 being fundamentally different from other South Asians, supporting a holistic view of relationships. For the second, users again favored the interpretable plot to show that PCA-based distinctions involve significant gray areas. Finally, users indicated that the interpretable visualization more convincingly disproved the notion that racial differentiation could be attributed to a single gene (Fig. 5, left).

To test robustness to deliberate misinterpretation, users rated the effectiveness of the baseline and interpretable method in 1) supporting a given racialized claim and, 2) rebutting a given racialized claim, and 3) mitigating and preventing deliberate misinterpretation (Fig. 5, right). Users found that, compared to the static baseline, the interpretable method supported racialized claims less easily, and instead strongly rebuked said claim. Additionally, when asked about ability to mitigate and entirely prevent weaponization, users favored the interactive visualization, providing evidence for the effectiveness of our visualization system in creating socially-robust visuals.

5 CONCLUSION

This new approach improved user understanding of contributing DNA regions to PCA formation, embeds notions of plot sensitivity, and increases robustness to deliberate manipulation.

More broadly, we hope this dashboard can be applied to larger sections of the human genome. This study was conducted exclusively with chromosome 1, thus we hope to show that similar findings will occur across chromosomes. Additionally, this dashboard and interpretation-in-visualization method will be applied to other dimensionality reduction techniques, such as UMAP and t-SNE.

While no single solution will be able to solve deliberate misinterpretation of data, it is critical to build safeguards and education against this use. In establishing an accessible platform for both scientists and non-scientists alike, this dashboard pushes closer to having a socially-responsible platform for human genomic visualization.

ACKNOWLEDGEMENTS

The authors wish to thank Prof. David Laidlaw for his project guidance and for teaching CS2370. Additionally, we want to thank Alex Diaz-Papkovich and Sohini Ramachandran for their guidance in this project. Finally, we thank the BIOL1430 students and Prof. Dan Weinreich for acting in the user survey

REFERENCES

[1] A. Auton, G. R. Abecasis, D. M. Altshuler, R. M. Durbin, G. R. Abecasis, D. R. Bentley, A. Chakravarti, A. G. Clark, P. Donnelly, E. E. Eichler, P. Flicek, S. B. Gabriel, R. A. Gibbs, E. D. Green, M. E. Hurles, B. M. Knoppers, J. O. Korbel, E. S. Lander, C. Lee, H. Lehrach, E. R. Mardis, G. T. Marth, G. A. McVean, D. A. Nickerson, J. P. Schmidt, S. T. Sherry, J. Wang, R. K. Wilson, R. A. Gibbs, E. Boerwinkle, H. Doddapaneni, Y. Han, V. Korchina, C. Kovar, S. Lee, D. Muzny, J. G. Reid, Y. Zhu, J. Wang, Y. Chang, Q. Feng, X. Fang, X. Guo, M. Jian, H. Jiang, X. Jin, T. Lan, G. Li, J. Li, Y. Li, S. Liu, X. Liu, Y. Lu, X. Ma, M. Tang, B. Wang, G. Wang, H. Wu, R. Wu, X. Xu, Y. Yin, D. Zhang, W. Zhang, J. Zhao, M. Zhao, X. Zheng, E. S. Lander, D. M. Altshuler, S. B. Gabriel, N. Gupta, N. Gharani, L. H. Toji, N. P. Gerry, A. M. Resch, P. Flicek, J. Barker, L. Clarke, L. Gil, S. E. Hunt, G. Kelman, E. Kulesha, R. Leinonen, W. M. McLaren, R. Radhakrishnan, A. Roa, D. Smirnov, R. E. Smith, I. Streeter, A. Thormann, I. Toneva, B. Vaughan, X. Zheng-Bradley, D. R. Bentley, R. Grocock, S. Humphray, T. James, Z. Kingsbury, H. Lehrach, R. Sudbrak, M. W. Albrecht, V. S. Amstislavskiy, T. A. Borodina, M. Lienhard, F. Mertes, M. Sultan, B. Timmermann, M.-L. Yaspo, E. R. Mardis, R. K. Wilson, L. Fulton, R. Fulton, S. T. Sherry, V. Ananiev, Z. Belaia, D. Beloslyudtsev, N. Bouk, C. Chen, D. Church, R. Cohen, C. Cook, J. Garner, T. Hefferon, M. Kimelman, C. Liu, J. Lopez, P. Meric, C. O'Sullivan, Y. Ostapchuk, L. Phan, S. Ponomarov, V. Schneider, E. Shekhtman, K. Sirotkin, D. Slotta, H. Zhang, G. A. McVean, R. M. Durbin, S. Balasubramaniam, J. Burton, P. Danecek, T. M. Keane, A. Kolb-Kokocinski, S. McCarthy, J. Stalker, M. Quail, J. P. Schmidt, C. J. Davies, J. Gollub, T. Webster, B. Wong, Y. Zhan, A. Auton, C. L. Campbell, Y. Kong, A. Marcketta, R. A. Gibbs, F. Yu, L. Antunes, M. Bainbridge, D. Muzny, A. Sabo, Z. Huang, J. Wang, L. J. M. Coin, L. Fang, X. Guo, X. Jin, G. Li, Q. Li, Y. Li, Z. Li, H. Lin, B. Liu, R. Luo, H. Shao, Y. Xie, C. Ye, C. Yu, F. Zhang, H. Zheng, H. Zhu, C. Alkan, E. Dal, F. Kahveci, G. T. Marth, E. P. Garrison, D. Kural, W.-P. Lee, W. Fung Leong, M. Stromberg, A. N. Ward, J. Wu, M. Zhang, M. J. Daly, M. A. DePristo, R. E. Handsaker, D. M. Altshuler, E. Banks, G. Bhatia, G. del Angel, S. B. Gabriel, G. Genovese, N. Gupta, H. Li, S. Kashin, E. S. Lander, S. A. McCarroll, J. C. Nemesh, R. E. Poplin, S. C. Yoon, J. Lihm, V. Makarov, A. G. Clark, S. Gottipati, A. Keinan, J. L. Rodriguez-Flores, J. O. Korbel, T. Rausch, M. H. Fritz, A. M. Stütz, P. Flicek, K. Beal, L. Clarke, A. Datta, J. Herrero, W. M. McLaren, G. R. S. Ritchie, R. E. Smith, D. Zerbino, X. Zheng-Bradley, P. C. Sabeti, I. Shlyakhter, S. F. Schaffner, J. Vitti, D. N. Cooper, E. V. Ball, P. D. Stenson, D. R. Bentley, B. Barnes, M. Bauer, R. Keira Cheetham, A. Cox, M. Eberle, S. Humphray, S. Kahn, L. Murray, J. Peden, R. Shaw, E. E. Kenny, M. A. Batzer, M. K. Konkel, J. A. Walker, D. G. MacArthur, M. Lek, R. Sudbrak, V. S. Amstislavskiy, R. Herwig, E. R. Mardis, L. Ding, D. C. Koboldt, D. Larson, K. Ye, S. Gravel, The 1000 Genomes Project Consortium, Corresponding authors, Steering committee, Production group, Baylor College of Medicine, BGI-Shenzhen, Broad Institute of MIT and Harvard, Coriell Institute for Medical Research, E. B. I. European Molecular Biology Laboratory, Illumina, Max Planck Institute for Molecular Genetics, McDonnell Genome Institute at Washington University, US National Institutes of Health, University of Oxford, Wellcome Trust Sanger Institute, Analysis group, Affymetrix, Albert Einstein College of Medicine, Bilkent University, Boston College, Cold Spring Harbor Laboratory, Cornell University, European Molecular Biology Laboratory, Harvard University, Human Gene Mutation Database, Icahn School of Medicine at Mount Sinai, Louisiana State University, Massachusetts General Hospital, McGill University, and N. National Eye Institute. A global reference for human genetic variation. *Nature*, 526(7571):68–74, Oct. 2015. Publisher: Nature Publishing Group.

[2] A.-H. R. Carlson, Henn. Counter the weaponization of genetics research by extremists, Oct. 2022.

[3] P. Castellano-Escuder, D. K. Zachman, K. Han, and M. D. Hirschey. GAUDI: interpretable multi-omics integration with UMAP embeddings and density-based clustering. *Nature Communications*, 16(1):5771, July 2025. Publisher: Nature Publishing Group.

[4] N. Duforet-Frebourg, K. Luu, G. Laval, E. Bazin, and M. G. Blum. Detecting Genomic Signatures of Natural Selection with Principal Component Analysis: Application to the 1000 Genomes Data. *Molecular Biology and Evolution*, 33(4):1082–1093, Apr. 2016.

[5] E. Elhaik. Principal Component Analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. *Scientific Reports*, 12(1):14683, Aug. 2022. Publisher: Nature Publishing Group.

[6] P. Paschou, E. Ziv, E. G. Burchard, S. Choudhry, W. Rodriguez-Cintron, M. W. Mahoney, and P. Drineas. PCA-Correlated SNPs for Structure Identification in Worldwide Human Populations. *PLOS Genetics*, 3(9):e160, Sept. 2007. Publisher: Public Library of Science.

[7] E. B. Stovner and P. Sætrom. PyRanges: efficient comparison of genomic intervals in Python. *Bioinformatics*, 36(3):918–919, Feb. 2020.

[8] C. Wu, A. Mark, and A. I. Su. MyGene.info: gene annotation query as a service, Sept. 2014. Pages: 009332 Section: New Results.

[9] S. Yi, Z. Lai, Z. He, Y.-m. Cheung, and Y. Liu. Joint sparse principal component analysis. *Pattern Recognition*, 61:524–536, Jan. 2017.

6 CONTRIBUTIONS: SKYLAR WALTERS

6.1 Intellectual Contributions

As PI for the project, I was responsible for the conception of the visualization system. I co-developed this framework by reviewing past research in PCA for human genomics and work in social visualization, with particular emphasis on sources [2] and [5]. This project remained largely similar to our finished creation, though we ultimately decided to focus on PCA instead of working with both PCA and UMAP.

After creating the group, I proposed adding a feature that addressed subsample interpretations (ie., our lasso tool), which Kyle and I discussed in-depth before deciding to implement as a supplementary feature.

6.2 Practical Contributions

Much of my technical contributions came in the form of writing Python Scripts and creating the Dash app. Having a background in population genetics, I was already familiar with processing human genome data. I pulled all of the ChrY and Chr1 data from 1000 Genomes, then created the genotype matrices with the GenotypeArray package. Further, I wrote scripts to filter to only the high-quality reads to ensure the robustness of our PCA. Following this, I coded the PCA that we used in our experimental ChrY data, then the real PCA we employed in the final Chr1 data.

With this, I was able to create an interpretability software that returned the relevant locus location, but not the actual function. With Kyle's integration of MyGene and Pyranges, we were able to unite these elements.

Using the format and structure of Kyle's HTML draft, I was able to write the Dash app for use in the survey. This entailed restructuring much of our preliminary code to work in Dash as opposed to Jupyter notebooks, and broadly covered 1) having alternative views correspond to the main plot, 2) having hovers show up, and 3) adapting Kyle's gene interpretation panel.

Being in BIOL1430, I wrote the survey/quiz, and also administered it in the course.

7 CONTRIBUTIONS: KYLE WISIALOWSKI

7.1 Intellectual Contributions

When I began working on the project, Skylar had already developed a first-draft of the interactive PCA plot with hoverable tooltips displaying the top 3 contributing SNP locations for each datapoint. My first task was to figure out how map the SNP location to a gene label. After a bit of research, I was able to find a Python library called MyGene that made it very easy to query annotated gene data.

Once we had decided to use MyGene, the next challenge was figuring out how to effectively present gene descriptions to users. I helped brainstorm a redesign of our interactive display to include an exploratory side panel that would allow the user to investigate detailed gene information without overloading the main interface.

7.2 Practical Contributions

After experimenting with the MyGene library as a method of acquiring gene descriptions, I ended up sticking with this approach and integrated it into our final pipeline. I was able to include the description query functionality into the existing architecture that populated the hoverable displays on the plot.

While we were still figuring out how we wanted to summarize the gene descriptions, I decided to write a HTML-generating script that would extend the PyPlot interface to include a large side panel. Since PyPlot was limiting us to tooltips as our only method of information display, it was hard to customize the existing script. The HTML program allowed me to include a completely separate side panel that Skylar could further refine by converting the HTML design into a PyDash program.

LiDAR Classification and Intervisibility Visualization for Chachapoya Archaeology

Yiyang Zhang* Yuqiao Guan† Parker VanValkenburgh

Brown University

ABSTRACT

LiDAR point clouds are very common data formats in archaeology. They effectively record the spatial and visual information captured from a region in a top-down view, providing accurate distance measurements and structural information about archaeological sites. However, such systems suffer from dense vegetation and complex geography. In the context of this project, dense vegetation hinders users from effectively identifying the structures of human-constructed sites and determining mutual visibility through the coverage of dense vegetation. Our work aims to address these issues by aggressively removing vegetation and reconstructing the remaining sites into explicit geometries that accurately reflect their topological structures.

Keywords: Reconstruction, LiDAR-based point clouds, site extraction

1 INTRODUCTION

LiDAR (Light Detection and Ranging) is a technique commonly used in archaeology to collect point cloud data that captures the geographical information of a site. They are proven to be very effective and accurate. However, in vegetation-dense areas like mid and southern America, vegetation may conceal human-constructed sites that researchers are interested in and prevent them from understanding the internal structures of these sites. Besides that, the researchers are also interested in the spatial distribution of sites and their intervisibility - it is a demonstration social hierarchy between the residents of the sites. Effective vegetation removal and site extraction will contribute significantly in determining both issues.

Our approach involves a 3-step pipeline: patch-level vegetation filtering, surface reconstruction and Laplacian smoothing of reconstruction. The first step mobilizes PDAL (Point Data Abstraction Library) to classify ground and non-ground points and remove the non-ground ones. The second step mobilizes Poisson Reconstruction to reconstruct the surfaces from point clouds, with a final step of applying Laplacian smoothing to improve the geometry integrity.

2 EXPOSITION

2.1 Overeview of the Pipeline

Our progressing pipeline transforms an unprocessed LiDAR point-cloud dataset that contains coverage of a forested Chachapoya hilltop, producing a smoothed mesh reconstruction of the architectural regions with the tall vegetation removed, emphasizing the structures. This workflow comprises four main steps:

- **Patch-level vegetation filtering** using joint height and color clues
- **Poisson surface reconstruction** of a watertight mesh from the segmented point cloud
- **Laplacian smoothing** to reduce small-scale noise and stabilize the reconstructed mesh

The philosophy of our approach is to aggressively remove vegetation and optimize the results with maximized exploitation of computation resources, then reconstruct the sites through Poisson Surface reconstruction or Point2Mesh models. This approach provides researchers with explicit geometry that reflects accurate internal structures through providing an explicit geometry instead of sparse points that is not reflective of structural information.

2.2 Ground-Based Vegetation Removal with PDAL

The raw LiDAR point cloud represents terrain, vegetation, and human constructions. Vegetation removal is performed indirectly through ground classification, following standard airborne LiDAR processing workflows as implemented in the Point Data Abstraction Library (PDAL). Rather than explicitly identifying vegetation objects, this approach separates ground from non-ground points based on local geometric properties of the surface.

The method assumes that the ground surface is locally smooth and continuous and above-ground objects (including vegetation) may cause abrupt height variations over flat areas. A ground classification filter will be applied to the point cloud, operating on spatial proximate objects to evaluate elevation differences, slope changes, and local relief. Points consistent with the expected terrain height variation are labeled as ground, while points that exceed these constraints are classified as non-ground.

Ground elevation is estimated by analyzing the lowest returns within local neighborhoods and progressively refining a terrain surface model. This process is conceptually similar to progressive morphological filtering and other widely used bare-earth extraction techniques in LiDAR analysis.

Vegetation removal is then achieved by keeping only points classified as ground and removing all non-ground points. This results in a bare-earth representation of the landscape in which vegetation canopy, understory, and other elevated features are removed. Unlike point-wise semantic classification, this method relies purely on geometric relationships and does not incorporate spectral or intensity information.

While effective for producing digital terrain models and revealing large-scale topographic features, this approach may misclassify low vegetation or erode small above-ground structures. As a countermeasure, we manually select the human-constructed structures from the remaining point clouds and mobilize surface reconstruction algorithms/models to build an explicit geometry.

2.3 Poisson Surface Reconstruction

After vegetation removal, the remaining points represent the union of ground and unchanged structural surfaces, but we can still only observe them as a sparse sample. To obtain a more visually coherent and interpretable representation, we reconstruct a mesh via Poisson surface reconstruction [2].

*e-mail: yiyang_zhang3@cs.brown.edu

†e-mail: yuqiao$_{guan}$@cs.brown.edu

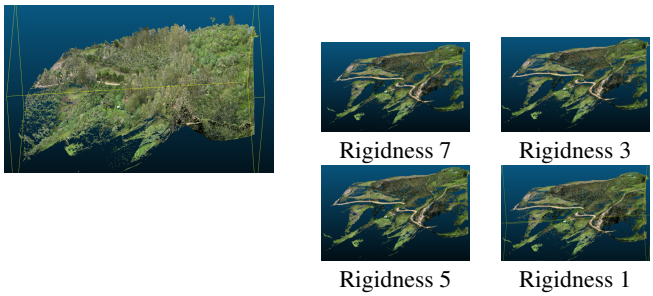

Figure 1: Comparison of reconstruction results.

Poisson surface reconstruction treats the input as a point set where each point has an estimated normal vector, which we compute through principal component analysis (PCA) on its k-nearest neighbors. The algorithm treats these normals as samples of the gradient of an unknown indicator function χ, then solves a Poisson equation

$$\nabla^2 \chi = \nabla \cdot \mathbf{V},$$

where $\mathbf{V}$ is a vector field derived from the oriented points. The resulting scalar field is globally optimized to fit all input points at the same time.

In our case, Poisson reconstruction is well suited to the archaeological data, since it avoids local inconsistencies that can caused in purely using local triangulation on noisy data. Also, where vegetation removal has created small gaps (between tree trunks and adjacent walls), the implicit formulation naturally interpolates plausible surfaces, often closing wall faces and terrace tops.

Another attempt we have made is using Point2Mesh. It is a "technique that reconstruct a surface mesh from an input point cloud. This approach "learns" from a single object, by optimizing the weights of a CNN to deform some initial mesh to shrink-wrap the input point cloud [1]." One of the particular strength of this model is computational simplicity: since there exists a model for it, we can use the model for inference instead of spending enormous computation power on reconstruction. Another major strength is that this model provides a water-tight reconstruction, which ensures the geometric integrity of the reconstruction quality. The users were asked in the user study to determine the model created from different approaches.

2.4 Laplacian Smoothing

While Poisson reconstruction provides a relatively smooth surface, the mesh can still show small stair-step-like artifacts inherited from the irregularity of the sampled points. We apply a Laplacian-based mesh smoothing operator to further regularize the mesh structure [3].

This method acts as a low-pass filter on vertex positions:

$$\mathbf{x}_i^{(t+1)} = \mathbf{x}_i^{(t)} + \lambda \sum_{j \in \mathcal{N}(i)} w_{ij} \left(\mathbf{x}_j^{(t)} - \mathbf{x}_i^{(t)} \right)$$

where $\mathbf{x}_i$ is the position of vertex i, $\mathcal{N}(i)$ is its 1-ring neighborhood, w_{ij} are normalized weights, and λ is a small step size. It cancels high-frequency noisy vertices while preserving lower-frequency geometric features.

We apply a small number of smoothing iterations with small step sizes (¡5) to preserve sharp wall tops and terrains while avoiding over-smoothing.

3 User Studies

All users were presented with the following results and asked which of the following models do the consider as the best in terms of geometry integrity: 1. Pure point clouds from LiDAR files. 2. Results from Poisson Surface reconstruction algorithm without Laplacian smoothing. 3. Results from Point2Mesh model without Laplacian smoothing. 4. Results from Poisson Surface reconstruction algorithm with Laplacian smoothing. 5. Results from Point2Mesh model without Laplacian smoothing. Our results show that most users consider 5 as the best-performing processing pipeline.

4 Conclusion

We present a pipeline for improving the interpretability of LiDAR point clouds in vegetation-dense archaeological landscapes. By aggressively removing vegetation and reconstructing the human-constructed features into explicit surface geometries, we aim to address two key challenges in archaeological LiDAR point cloud dataset analysis: the visual obstruction caused by dense canopy coverage and the difficulty of reasoning about site structure and intervisibility from sparse point data.

Our pipeline integrates ground-based vegetation filtering using PDAL with surface reconstruction techniques, including Poisson surface reconstruction and learning-based Point2Mesh models, followed by Laplacian smoothing to enhance geometric coherence. This combination enables the transformation of point clouds damaged by aggressive filtering into watertight and visually stable meshes that better capture the topology and spatial organization of human-constructed. Compared to raw point clouds, the reconstructed geometries provide a clearer representation of human constructions.

User study results indicate a preference for reconstructed meshes over raw LiDAR point clouds, with learning-based reconstruction methods receiving the highest ratings in terms of perceived geometric integrity. These findings suggest that explicit surface reconstruction, particularly when paired with aggressive vegetation suppression, can significantly enhance the usability of LiDAR data for archaeological analysis.

Despite these improvements, our approach remains sensitive to the quality of ground classification and may erode subtle above-ground features during vegetation removal. Future work will focus on incorporating semantic-aware filtering, adaptive parameter tuning, and visibility analysis directly on reconstructed meshes to further support archaeological interpretation. Overall, this work demonstrates the value of combining traditional LiDAR processing with modern surface reconstruction techniques to better illustrate the structural information embedded in complex, vegetation-covered landscapes.

Acknowledgements

The authors wish to thank the dataset and guidance offered by Professor Parker VanValkenburgh, who made his best effort to support the development of this project.

REFERENCES

[1] R. Hanocka, G. Metzer, R. Giryes, and D. Cohen-Or. Point2mesh: A self-prior for deformable meshes. *ACM Trans. Graph.*, 39(4), 2020.

[2] M. Kazhdan, M. Bolitho, and H. Hoppe. Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing*, volume 7, 2006.

[3] A. Nealen, T. Igarashi, O. Sorkine, and M. Alexa. Laplacian mesh optimization. In *Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*, pages 381–389, 2006.

5 CONTRIBUTION

5.1 Yiyang Zhang

My contributions to this project is the development and testing of the patch-level filtering and reconstruction pipeline for the LiDAR data. I implemented the local patch-based vegetation filter that combines height-above-ground and RGB green-channel dominance, including the design of per-patch statistics, threshold selection, and rules for preserving structurally ambiguous regions. I also integrated Poisson surface reconstruction into the pipeline to ensure that architectural features such as walls and terraces remained coherent in the resulting meshes.

5.2 Yuqiao Guan

I contributed to the project by introducing Point2Mesh model into the pipeline and tested the PDAL pipeline for vegetation removal. Also, I implemented the pipeline to incorporate extracted sites and post-process with Point2Mesh models to observe its advantage and disadvantage over Poisson surface reconstruction methods.