Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

GeoAI Platform for Natural Language Exploration of Multidimensional Flood Risk Scenarios

Aditya Barve

asbarve@mit.edu

Massachusetts Institute of Technology

Rutvik Deshpande

rutvik_d@mit.edu

Massachusetts Institute of Technology

Miho Mazereeuw

Massachusetts Institute of Technology

Research Article

Keywords: Flood Risk, Geospatial AI, Natural Language Interface

Posted Date: October 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11540644/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

GeoAI Platform for Natural Language Exploration of Multidimensional Flood Risk Scenarios

Aditya Barve
Massachusetts Institute of Technology
Cambridge, USA
asbarve@mit.edu

Rutvik Deshpande
Massachusetts Institute of Technology
Cambridge, USA
rutvik_d@mit.edu

Miho Mazereeuw
Massachusetts Institute of Technology
Cambridge, USA
mmaz@mit.edu

Abstract

While many coastal jurisdictions increasingly use flood models that incorporate future climate change projections, floodplain managers, planners and community organizers frequently need to answer questions that draw on multiple data sources and model outputs. A single compound planning question requires stitching together results from inundation outputs, infrastructure records, and socioeconomic datasets that currently live in disconnected silos. The need for complex workflows and specialized tooling constrains the questions practitioners can ask and the speed with which they can get the results. This paper presents a geospatial AI platform that addresses this gap through three architectural contributions: (i) A multidimensional scenario store that organizes 141 flood scenarios within a single queryable structure along with other physical and socioeconomic datasets; (ii) A deterministic tool-routed LLM agent that maps compound natural language queries into geospatial tool calls; and (iii) A bidirectional map-chat interface that preserves spatial grounding throughout the analytical workflow.

CCS Concepts

• **Information systems → Geographic information systems**; • **Human-centered computing → Natural language interfaces**.

Keywords

Flood Risk, Geospatial AI, Natural Language Interface

ACM Reference Format:
Aditya Barve, Rutvik Deshpande, and Miho Mazereeuw. 2026. GeoAI Platform for Natural Language Exploration of Multidimensional Flood Risk Scenarios. In *The 34th ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL '26), November 03–06, 2026, Riverside, CA, USA*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3841645.3843394

1 Introduction

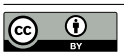
Coastal cities face compound flood hazards from projected sea-level rise, heavy precipitation, and storm surge. The resulting risk calculus is inherently multidimensional, combining hazard variables such as rainfall intensity, sea-level rise, and storm surge with socioeconomic vulnerability factors and expected effects of adaptation interventions. This multidimensionality creates a steep practical barrier for decision-makers who must translate these interacting probabilities into planning actions. Answering a question like *"Which areas in my city will need more evacuation assistance than others in worst-case flooding by 2050?"* requires understanding of modeled scenarios and comparative flooding across different areas, shelter locations, and socio-economic barriers to evacuation. A non-expert community organizer, an emergency manager during an active event, or a planner at a community meeting cannot currently perform this work on the fly. Institutional fragmentation and local-level resource constraints further compound this bottleneck. While flood models are often produced at the county or state scale, the decisions they inform are made at the municipal level, where resources are often limited. As a result, common planning workflows often require manual coordination across a fragmented data landscape.

1.1 Related Work

Coastal adaptation planning requires evaluating intersecting variables such as storm return periods, sea-level rise trajectories, surge conditions, and the mitigating effects of specific adaptation tiers, along with socio-economic datasets [8]. However, currently, data platforms where this exploration can take place are either advanced geospatial analytical platforms that require specialist operators, sequential layer loading, and manual dataset integration. Or, they are public-facing lookup tools like the *FEMA Flood Map Service Center, First Street Foundation*, and *NOAA Sea Level Rise Viewer*, which lower the access barrier at the cost of limiting interaction to pre-computed single-scenario visualization [5, 7, 11].

LLMs have been used to support flood situational awareness, but they primarily focus on information access rather than structured reasoning over future climate scenarios and socio-economic datasets [10]. Within GIS software, LLM-based geospatial workflows have successfully demonstrated the use of natural language interfaces to retrieve and generate GIS queries, climate-related information, and automate spatial workflows [1, 9, 14]. Existing platforms, however, do not address the specific challenge that adaptation planning requires: resolving compound queries simultaneously across different storm return periods, socioeconomic datasets, and infrastructure impacts while keeping analytical outputs coupled to an interactive map state for ease of visual communication. This paper presents a new geospatial AI platform that addresses this gap through three contributions:

(1) Zarr-backed scenario store encoding 141 flood scenarios, infrastructure layers, socioeconomic indicators, and adaptation impact tables under a common spatial index;
(2) Tool-routed LLM agent that processes compound natural language queries into typed analytical function calls executed against this index; and



This work is licensed under a Creative Commons Attribution 4.0 International License.
SIGSPATIAL '26, Riverside, CA, USA
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2950-8/2026/11
https://doi.org/10.1145/3841645.3843394

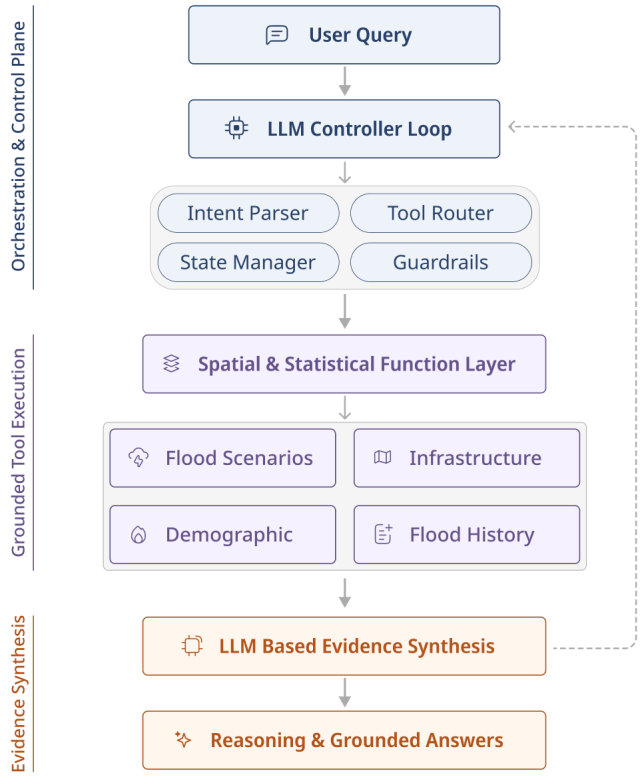


Figure 1: System architecture showing three functional layers: orchestration (query parsing and tool routing), tool execution (spatial and statistical functions across data), and evidence synthesis (LLM-based summarization with map-synchronized output).

(3) Bidirectional map-chat interface that keeps answers spatially grounded through synchronized map outputs.

2 Datasets

2.1 Hydrodynamic Flood Scenarios

The primary dataset in this framework is a multidimensional flood scenario dataset produced by the Broward County Resilience Unit at Broward County, Florida [3]. This dataset is derived from countywide hydrodynamic inundation modeling under a structured combination of future climate and storm scenarios, in addition to proposed adaptation projects under the county resilience plan [2]. The resulting data store comprises 141 precomputed flood scenarios, covering a range of modeled conditions from plausible near-term risks to longer-horizon tail-risk scenarios. Organized across multiple rainfall return periods, these scenarios range from 1-year to 100-year events; three sea-level-rise conditions: Current, 2 ft (2050), and 3.3 ft (2070); and three surge conditions: no surge, 20-year, and 100-year. Each scenario combination is evaluated under three states: (1) baseline, representing existing infrastructure without additional intervention; (2) Tier 1, a near-term adaptation suite of infrastructure investments such as drainage improvement and seawall construction (2025–2040); and (3) Tier 2, a longer-horizon strategy that builds on Tier 1 with raised seawalls, advanced pumping systems, and expanded green infrastructure (2040–2070).

2.2 Resilience Impact Data

The county resilience plan provides additional datasets relevant to the county's flood risk assessment and preparedness. To measure flood risk exposure, the county has curated infrastructure assets, categorized into emergency shelters and critical facilities, consisting of fire stations, hospitals, schools, and law enforcement facilities that the county administration is responsible for maintaining operational access to during disasters. Additionally, the resilience plan includes census tract-level financial impact data assessing benefits, economic damages, and service disruptions resulting from adaptation proposals.

2.3 Socioeconomic & Historical Claims Data

Beyond the flood modeling data from the county, the system integrates federally curated socioeconomic datasets. The Centers for Disease Control and Prevention's (CDC) Social Vulnerability Index (SVI) provides tract-level information related to socioeconomic status, demographic trends, and household characteristics [4]. American Community Survey (ACS) provides granular demographic information, such as poverty rate, age distribution, median income, disability, etc., which can help in assessing the demographic profile of a region [13]. Another important data source for flood risk assessment is the National Flood Insurance Program (NFIP) claims data curated by the Federal Emergency Management Agency (FEMA), which is used as a historical reference for comparing modeled risk patterns with past reported losses [6].

3 Methodology

The system architecture is organized into three main functional layers: data synthesis, query routing and grounded tool execution, and finally the evidence synthesis layer, as shown in Figure 1.

3.1 Data Preparation and Integration

To translate the county's existing resilience datasets into queryable data infrastructure, the framework restructures official scenario outputs, facility inventories, and impact tables into formats that support efficient spatial retrieval. The compound flood model produces a spatially explicit inundation depth grid for each combination of rainfall intensity, storm surge return period, sea-level rise, and adaptation tier across the county. The depth grids have 125 ft resolution in the *NAD83(HARN) / Florida East projection*. The array is chunked along the spatial and scenario dimensions, allowing individual subsets to be retrieved independently, without loading the entire dataset using Zarr [15]. The administrative boundary polygons and critical facilities data are cleaned and stored as *GeoParquet* files in the same bucket, enabling compressed columnar access and direct spatial joins. The financial-impact-related datasets are stored as flat, comma-separated values (CSV) keyed by census tracts and sliced to boundary geometry when the tool needs to rank or compare tracts by economic benefits. The rest of the socioeconomic datasets are not stored locally; instead, they are fetched from federal APIs during the query and cached per session.

Table 1: Tool taxonomy and system functions

Domain	Tool	Description
Scenario Selection	`filter_scenarios`, `get_scenario_info`	Resolves natural language queries into scenario indices and returns available return periods for rainfall, sea-level rise, surge, and adaptation tier values.
Flood Risk	`get_flood_depth_at_point`, `get_flood_depth_in_area`, `find_min_flood_conditions_at_point`, `find_min_flood_conditions_in_area`, `get_all_scenarios_area_stats`	Computes cross-scenario statistics, including percent area flooded, minimum flood conditions, and inundation depth for the selected area of interest under queried scenarios.
Ranking & Comparison	`rank_areas_by_flood_risk`, `compare_areas_by_flood_risk`, `compare_adaptation_tiers`, `query_financial_impacts`	Compares and ranks multiple census tracts or cities by selected metrics across scenarios, including coverage, maximum depth, flood probability, and avoided residential damages.
Facility Assessment	`get_facility_flood_risk`, `get_nearest_facilities`, `get_facility_types`, `get_facilities_in_aoi`	Samples inundation depth at curated infrastructure locations within an area of interest and ranks facilities by exposure.
Parcels, Streets and Plans Search	`get_parcel_flood_risk`, `get_street_flood_risk`, `search_resilience_plan`	Samples inundation depth and returns street- and parcel-level flooding. Plan search goes through text embeddings of the resilience plan to augment answers.

3.2 Query Orchestration

A core design requirement of the system was to make semantically equivalent queries produce consistent results. To achieve this, the system uses a model-swappable architecture built on an LLM function-calling loop and uses OpenAI's *GPT-4o mini* to facilitate spatial querying and statistical summarization [12]. This framework encompasses a set of flood risk query functions with analytical capabilities supporting scenario exploration, spatial exposure analysis and impact assessment as defined in Table 1. Each of these functions is an LLM callable tool with a fixed signature and a specified JSON schema describing its parameters, defaults and valid ranges of values. The system matches the user query with the tool descriptions and emits a structured function call with validated arguments. The server executes these functions using the data stored on the cloud and returns a deterministic result.

The tool descriptions and schemas serve as the guardrails to avoid hallucinations and errors. Each tool schema encodes domain constraints like valid rainfall values, valid scenario indices, area names from boundary catalogs, and categories available in facility datasets. The system prompt creates an additional layer, mapping natural language patterns to specific tools. As the LLM output is a function call rather than generated text, validation is deterministic and can be checked against predefined schemas and valid arguments. For compound questions, like, *'Rank hospitals in my city by flood depth under worst case conditions'*, the system resolves the request into a set of sequential tool calls: first `filter_scenarios()` resolves the 'worst case' semantic into concrete scenario indices, then the `get_facility_flood_risk()` tool selects queried facilities and runs a flood risk analysis on those within the selected area of interest.

3.3 User Interaction

The user interaction is built around two coupled surfaces - an interactive map where users select an area of interest (AOI) as an address, census tract, or city, and a chat interface where natural language queries are resolved against that AOI. The query related to this AOI is asked within the chat interface, which triggers the LLM agent in the backend. For example, a user who selects Fort Lauderdale as the area of interest can ask a series of questions about the same geography without having to select it again. The threading of selection state across chained queries makes multi-step workflows feel conversational. To visualize the spatial grounding of answers, every analytical output is anchored to its geographical context through bidirectional map coupling. This map visualization is in one of the following formats: flood depth grids, facility pins, or ranked choropleths. The textual answers are summarized versions of analytical outputs that explain the results succinctly and in plain English. Table 2 traces a representative workflow, showing how user queries are resolved into scenario filtering, spatial sampling, external data retrieval, and map-grounded outputs, along with reported execution time and peak memory use.

3.4 Validation

The platform was tested through 189 distinct queries authored by the research team to span all five tool categories in Table 1 over a different range of geographies (address, tract, municipal boundaries). As the system emits typed function calls rather than freely generated text, we separate two failure modes: execution reliability and routing correctness.

- `Execution`: 172 (91%) queries completed successfully with a median latency of 11.8 s (p90 35.3 s). All 17 failures were infrastructure-related rather than incorrect reasoning: 15 exhausted memory on large county-scale queries, 2 failed on external API availability.

Table 2: End-to-end execution log for a representative workflow.

Facility triage under compound flooding	
Query	"Show me operational critical facilities under a 1-in-50 yr rainfall and 20-yr storm surge." (*Area of Interest:* Pompano Beach)
Execution	Resolves scenario conditions via exact-match filtering (9 of 141 scenarios matched across 3 tiers and 3 SLR levels), samples flood depths at 39 critical facilities within the city boundary, and ranks by maximum exposure.
Tool Calls	$\texttt{LLM}_1$ (3.6 s) → `filter_scenarios` (1 ms) → $\texttt{LLM}_2$ (3.8 s) → `get_facility_flood_risk` (6.8 s) → $\texttt{LLM}_3$ (7.2 s) → `synthesize` (1.9 s)
Performance	25.0 s total interaction time | 289 MB peak memory
Output	Ranked table of 10 most-exposed facilities, color-coded map markers, and scenario pills for all 9 matched conditions.

- `Routing`: For each query, we checked that the emitted tool usage and the final answer matched the intended analysis. In 170 of 172 cases, the system produced accurate results; in 2 cases, it invoked the correct tools but resolved them against an unintended scenario set. Both of these failures were traced to ambiguous language about probabilistic return period, as terms such as *"Worst case"* and *"1-in-50 flood"* can admit multiple resolvable mappings onto the scenario index. While this evaluation focused on deployment reliability and routing fidelity, determining whether the returned analysis is what the planner would have chosen and how much time-to-answer reduction is achieved compared to the manual GIS workflow requires a detailed participant study, which we leave to future work.

4 Discussion

The framework presented above demonstrates that structured multidimensional scenario storage, a deterministic tool-routed LLM agent, and a bidirectional map-chat interface can make multidimensional flood data widely accessible. In Broward County, this has institutional significance. Broward County's planning environment, with 31 municipalities under a countywide resilience framework and two decades of adaptation planning, provides the coordinated foundation this approach requires. Countywide resilience plans often bring multiple jurisdictions into a shared adaptation framework, yet the evidence behind those plans remains distributed across separate departments, datasets, and tools. The framework collapses those data silos into a single interface and a planning question into a single natural-language query resolved within seconds.

While the framework itself is scalable, it remains limited by the quality and coverage of the underlying inundation models. The system can not answer questions beyond modeled scenarios, and modeling errors, missing scenarios, or outdated facility records propagate directly into system outputs. Ambiguous language can also produce misrouted analysis. Deterministic function calling reduces these issues, but does not completely eliminate them. Finally, the conversational ease of the platform can convey greater certainty than is warranted when the underlying model carries significant uncertainty. Future work thus includes enhancing uncertainty communication as a central design requirement.

Acknowledgments

This work was conducted in collaboration with Broward County, Florida, in support of the county's Resilient Broward initiative. The authors thank Dr. Jennifer Jurado, Chief Resilience Officer of Broward County, for her guidance and support. This work is part of the MIT Climate Grand Challenge on Weather and Climate Extremes, supported by MIT Climate Grand Challenges.

References

[1] Amulya Bhattaram, Justin Chung, Stanley Chung, Ranit Gupta, Janani Ramamoorthy, Kartikeya Gullapalli, Diana Marculescu, and Dimitrios Stamoulis. 2025. GeoFlow: Agentic Workflow Automation for Geospatial Tasks. In *Proceedings of the 33rd ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL '25)*. Association for Computing Machinery, New York, NY, USA, 1150–1153. doi:10.1145/3748636.3763217

[2] Broward County, Florida. 2025. Countywide Resilience Plan 2025. https://www.broward.org/resilience/Pages/default.aspx

[3] Broward County, Florida. 2026. Broward County GeoHub. https://geohub-bcgis.opendata.arcgis.com/ Geospatial open data portal.

[4] Centers for Disease Control and Prevention and Agency for Toxic Substances and Disease Registry, Geospatial Research, Analysis, and Services Program. 2022. CDC/ATSDR Social Vulnerability Index 2022 Database: Florida. https://www.atsdr.cdc.gov/place-health/php/svi/svi-data-documentation-download.html State-level Social Vulnerability Index database.

[5] Federal Emergency Management Agency. 2026. *FEMA Flood Map Service Center*. Federal Emergency Management Agency. https://msc.fema.gov/ Accessed: 2026-06-16.

[6] Federal Emergency Management Agency (FEMA). 2026. FIMA NFIP Redacted Claims - v2 | FEMA.gov. https://www.fema.gov/openfema-data-page/fima-nfip-redacted-claims-v2

[7] First Street Foundation. 2025. *Flood Factor® Flood Risk Model Methodology*. First Street Foundation. https://firststreet.org/methodology/flood Accessed: 2026-06-16.

[8] Shih-Chun Hsiao, Wen-Son Chiang, Jiun-Huei Jang, Han-Lun Wu, Wei-Shiun Lu, Wei-Bo Chen, and Yun-Ta Wu. 2021. Flood risk influenced by the compound effect of storm surge and rainfall under climate change for low-lying coastal areas. *Science of The Total Environment* 764 (April 2021), 144439. doi:10.1016/j.scitotenv.2020.144439

[9] Jinghong Hu, Ligang Sun, and Xiliang Liu. 2025. GeoAgent: An Agentic AI Framework for Spatial Query Understanding and Interactive Geospatial Intelligence. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Generative and Agentic AI for Multi-Modality Space-Time Intelligence (GeoGenAgent '25)*. Association for Computing Machinery, New York, NY, USA, 54–60. doi:10.1145/3764915.3770719

[10] Rafaela Martelo, Kimia Ahmadiyehyazdi, and Ruo-Qian Wang. 2026. Towards democratized flood risk management: An advanced AI assistant enabled by GPT-4 for enhanced interpretability and public engagement. *Environmental Modelling & Software* 197 (Feb. 2026), 106821. doi:10.1016/j.envsoft.2025.106821

[11] NOAA Office for Coastal Management. 2026. *Sea Level Rise Viewer*. National Oceanic and Atmospheric Administration. https://coast.noaa.gov/digitalcoast/tools/slr.html Accessed: 2026-06-16.

[12] OpenAI. 2024. GPT-4o mini: advancing cost-efficient intelligence. https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/. Accessed: 2026-06-16.

[13] U.S. Census Bureau. 2026. American Community Survey 2024 5-Year Estimates. https://api.census.gov/data/2024/acs/acs5 Data accessed through the Census Bureau Application Programming Interface.

[14] Saeid Ashraf Vaghefi, Dominik Stammbach, Veruska Muccione, Julia Bingler, Jingwei Ni, Mathias Kraus, Simon Allen, Chiara Colesanti-Senni, Tobias Wekhof, Tobias Schimanski, Glen Gostlow, Tingyu Yu, Qian Wang, Nicolas Webersinke, Christian Huggel, and Markus Leippold. 2023. ChatClimate: Grounding conversational AI in climate science. *Communications Earth & Environment* 4, 1 (Dec. 2023), 480. doi:10.1038/s43247-023-01084-x

[15] Zarr Developers. 2026. *Zarr Storage Specification*. Zarr Developers. https://zarr-specs.readthedocs.io/ Accessed: 2026-06-16.