Compound Flood Alert System using Multimodal Hypercube-RAG

IDC 6940 Capstone in Data Science
Student: Mohankrishna Kolla
Course Instructor: Dr. Ananda M. Mondal
Mentor: Dr. Leonardo Bobadilla
Project Repository: https://github.com/FIU-MoRA-Lab/Hypercube-RAG-FIU/tree/main

***Abstract*—Compound flooding events arise when multiple flood drivers—such as storm surge, tides, and extreme rainfall—interact in space and time. These events are difficult to monitor in real time because the relevant information is scattered across numerical sensors, remote-sensing grids, hydrologic models, and unstructured text such as reports and alerts. This paper presents progress toward a *Compound Flood Alert System* that leverages Hypercube-based Retrieval-Augmented Generation (Hypercube-RAG) to fuse text and physical data into a unified retrieval and reasoning framework. I first summarize the construction and failure of a text-only hypercube built from a general ScienceDaily corpus, then describe the integration of multi-source hydrometeorological datasets (IMERG, MRMS, NOAA CO-OPS) into a multimodal hypercube. Preliminary evaluation on a Compound Flooding QA dataset using GPT-4o indicates that multimodal Hypercube-RAG dramatically improves semantic similarity, correctness, and completeness of answers compared to both no-retrieval and unimodal text-only RAG baselines. I conclude with a roadmap for turning this research prototype into a real-time, explainable alerting system.**

***Index Terms*—compound flooding, retrieval-augmented generation, Hypercube-RAG, hydrometeorology, explainable AI**

I. Introduction

Compound flooding refers to flood events driven by the interaction of multiple processes, such as coastal storm surge, tides, river discharge, and intense rainfall occurring simultaneously or in close succession. When these drivers overlap, they can produce higher water levels, longer-lasting inundation, and larger spatial footprints than any driver would cause alone, leading to substantial damages to infrastructure, ecosystems, and communities.

Despite advances in hydrologic and coastal modeling, stakeholders such as emergency managers, utilities, and residents still struggle to obtain *timely and explainable* information about compound flood risk. Operational products exist for individual drivers (e.g., tide gauges, rainfall estimates, storm surge forecasts), but integrating and interpreting conflicting signals across sources remains largely a manual and cognitively heavy task.

Retrieval-Augmented Generation (RAG) provides a promising path to combine large language models (LLMs) with external knowledge bases. However, standard RAG pipelines are optimized for text-only retrieval and are poorly suited for the multi-modal, spatio-temporal nature of compound flooding data.

A. Problem Statement and Goal

Relevant information for compound flooding is distributed across:

- unstructured text (scientific literature, agency reports, news, alerts),
- structured time series (water levels, discharge, radar-derived rainfall),
- gridded remote-sensing products (satellite precipitation, re-analysis),
- hydrologic or hydrodynamic model outputs.

Traditional text-only RAG cannot natively handle these heterogeneous modalities. Conversely, purely numerical systems lack the narrative and contextual grounding that text provides.

Goal: Develop a multimodal Hypercube-RAG system that unifies domain-relevant text and sensor/model data into a single retrieval framework for compound flooding analysis and alerts.

B. Research Questions

This work is guided by the following research questions:

1) **RQ1 (Corpus Quality):** How does the choice of text corpus affect RAG performance for compound flooding Q&A?
2) **RQ2 (Multimodal Fusion):** Can Hypercube-RAG effectively align text and physical data along shared dimensions (e.g., location, time, event type) for retrieval?
3) **RQ3 (Answer Quality):** Does multimodal Hypercube-RAG improve correctness, completeness, and semantic quality of answers compared to text-only RAG or no-retrieval baselines?
4) **RQ4 (Toward Alerts):** What design patterns are needed to turn Hypercube-RAG into an explainable alerting pipeline for compound flooding?

II. Background and Related Work

A. Compound Flooding

Recent hydrology and coastal engineering studies highlight the importance of jointly modeling tide, surge, river discharge, and precipitation when estimating flood risk in low-lying coastal regions. Compound events often produce non-linear impacts and are expected to become more frequent under climate change scenarios due to sea-level rise and more intense rainfall. This motivates decision-support tools that can reason about multiple drivers simultaneously and communicate that reasoning clearly.

B. Retrieval-Augmented Generation and Geo-RAG

RAG constrains LLMs by grounding their responses in external knowledge bases. Conventional RAG focuses on textual documents, retrieving relevant passages via dense or sparse retrieval before generation. Geo-RAG frameworks extend this idea to geoscience settings, retrieving rasters, digital elevation models, and equations while enforcing physics and domain rules during reasoning.

C. Hypercube-RAG

Hypercube-RAG organizes data into multi-dimensional “hypercubes” keyed by axes such as location, time, variable type, and data source. It excels at aligning structured geoscience data, while a RAG front-end uses that structure to retrieve coherent slices for a given query. This combination naturally supports multimodal storage (text, sensors, rasters) and dynamic AI reasoning, making it a promising substrate for compound flooding applications.

III. System Overview

Fig. 1 illustrates the conceptual Compound Flood Alert System built on Hypercube-RAG.

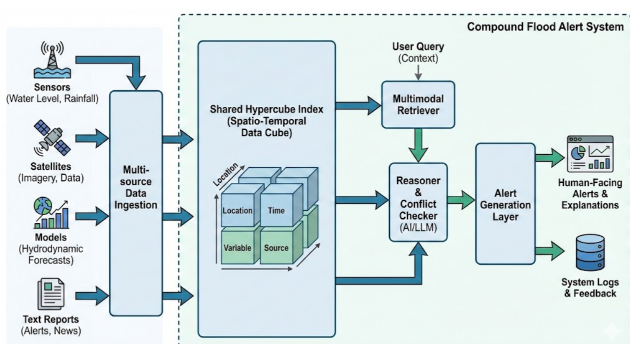

Fig. 1. Conceptual overview of the Compound Flood Alert System. Multi-source data are ingested into a shared hypercube, enabling multimodal retrieval, reasoning, conflict checking, and alert generation.

The system comprises four major components:

1) **Ingestion and Hypercube Construction:** Pulls data from multi-source feeds (sensor APIs, satellite products, document streams) and encodes them into a shared hypercube index keyed by dimensions such as location, time, data type, and source.
2) **Retriever:** Given a query with spatial and temporal context, retrieves relevant slices of both text and physical data.
3) **Reasoner and Conflict Checker:** Uses LLM-based reasoning plus domain heuristics to synthesize answers, check consistency across sources (e.g., rainfall vs. water level), and flag disagreements.
4) **Alert Layer:** Translates retrieval outputs into human-facing alerts that explain *why* risk is elevated, while logging decisions for backtesting.

IV. Data Sources

A. Textual Data and QA Construction

1) ScienceDaily Corpus: The initial text corpus was built by scraping 19,456 ScienceDaily articles. Only about 900 articles (roughly 1%) were related to storms, tides, or flooding; the majority focused on biology, chemistry, or medicine. This domain mismatch became critical when constructing a text-only hypercube: even with a large document count, the hydrology signal was weak and noisy.

2) Compound Flooding QA Dataset: To evaluate the system, a Compound Flooding QA dataset was manually constructed from domain-relevant sources such as NOAA, USGS, and NASA educational materials and key papers. Example questions include:

- “What is fluvial flooding and how can it combine with surge?”
- “Why do compound floods often extend farther inland than expected?”
- “What hydrological conditions precede major compound flood events?”

Each QA pair was stored in JSON, forming the text-based benchmark for comparing retrieval configurations.

B. Physical Data: Sensors and Remote Sensing

To move beyond text-only RAG, three main physical datasets were integrated, focusing on Florida coastal regions:

- **IMERG (GPM):** Half-hourly satellite precipitation (mm/hr) at grid points around NOAA tide gauges.
- **MRMS:** Radar-derived precipitation rate sampled at or near coastal stations.
- **NOAA CO-OPS:** Half-hourly water level observations (meters) from coastal tide gauges.

A deliberately different month was used for one of the datasets to stress-test temporal alignment, forcing the hypercube pipeline to handle non-overlapping windows.

TABLE I
Overview of physical datasets used in the initial multimodal Hypercube-RAG prototype.

Dataset	Description	Study Periods
IMERG	Half-hourly satellite precipitation (mm/hr) within 25 km of NOAA CO-OPS tide stations.	Jun 1–30, 2025
MRMS	30-minute radar-derived precipitation rate at ground-based stations.	Jun 1–30, 2025
CO-OPS	Half-hourly water levels (m) at coastal tide gauges.	Sep 21–Oct 22, 2025

Visualizations of CO-OPS water levels for Pensacola and Cape Canaveral reveal strong tidal oscillations largely independent of rainfall, providing a useful contrast when reasoning about compound events.

V. Exploratory Data Analysis

This section summarizes basic exploratory data analysis (EDA) for the IMERG, MRMS, and CO-OPS datasets used in the multimodal Hypercube-RAG prototype.

A. Dataset Overview

Table II provides a high-level description of the three physical datasets and their role in the compound flooding use case.

TABLE II
OVERVIEW OF PHYSICAL DATASETS USED IN COMPOUND FLOODING EDA.

Dataset	Description
IMERG	Half-hourly satellite precipitation (mm/hr) across grid points within ≈25 miles of NOAA CO-OPS tide stations.
MRMS	30-minute radar-derived precipitation rate sampled directly at CO-OPS tide gauge locations.
CO-OPS	Half-hourly water level observations (meters) from NOAA coastal tide gauge stations.
Purpose	Provide real-world physical context (rainfall + water levels) to support multimodal Hypercube-RAG retrieval.
Study Periods	IMERG + MRMS: Jun 1–30, 2025; CO-OPS: Sept 21–Oct 22, 2025.

B. IMERG Precipitation EDA

TABLE III
IMERG HALF-HOURLY PRECIPITATION EDA SUMMARY.

Characteristic	Value
Rows	1,051,200
Grid Points	730
Time Range	2025-06-01 to 2025-06-30
Columns	`time, point, lon, lat, precip_mm_per_hr`
Mean Precipitation (mm/hr)	0.226
Std Dev	1.116
Median	0.000
75th Percentile	0.000
Max Value	48.960
Missing Values	0
Saved Plots	`imerg_precip_timeseries.png, imerg_precip_hist.png`

Daily mean precipitation (Fig. 2) shows mostly dry conditions with intermittent bursts of higher rainfall, which is consistent with the skewed distribution summarized in Table III.

C. MRMS Precipitation EDA

Fig. 3 shows daily mean MRMS precipitation rate. Compared to IMERG, MRMS captures more localized high-intensity spikes, which are important for compound flooding scenarios.

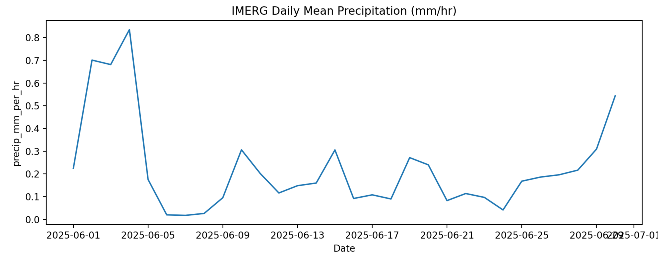

Fig. 2. IMERG daily mean precipitation (mm/hr) for June 2025.

TABLE IV
MRMS 30-MINUTE PRECIPITATION RATE EDA SUMMARY.

Characteristic	Value
Rows	33,097
CO-OPS Stations	23
Time Range	2025-06-01 to 2025-06-30
Columns	`station_id, time, lat, lon, precip_rate, source_file`
Mean Precipitation Rate (mm/hr)	0.262
Std Dev	3.190
Median	0.000
Min / Max	−3.000 / 184.700
Missing Values	0
Notes	Negative values represent radar artifacts and should be treated as noise.
Saved Plots	`mrms_precip_timeseries.png, mrms_precip_hist.png`

D. CO-OPS Water Level EDA

The empirical distribution of water levels and daily means are shown in Fig. 5. Tidal variability and gradual changes over the month are evident and largely independent of the June rainfall patterns.

E. Cross-Dataset Insights

The combined EDA across IMERG, MRMS, and CO-OPS yields several cross-dataset insights, summarized in Table VI.

VI. Hypercube Construction and RAG Pipeline

A. Text-Only Hypercube

The first experiment built a text-only hypercube over the ScienceDaily corpus. Named entity recognition (NER) was used to extract six entity types: location, date, event, organization, person, and theme. Each article was mapped into this six-dimensional space, and candidate slices were retrieved based on query entities.

Due to the off-topic corpus, the resulting hypercube was both *skewed* and *domain poor* for compound flooding. Answers generated by GPT-4o using unimodal RAG frequently ignored

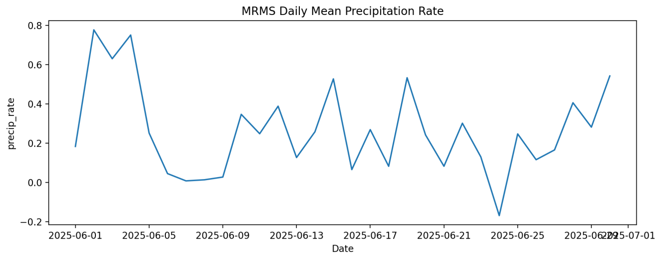

Fig. 3. MRMS daily mean precipitation rate (mm/hr) for June 2025.

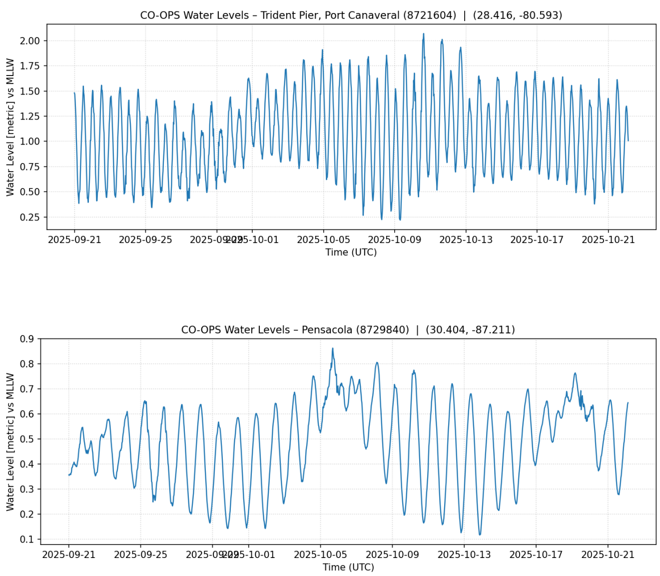

Fig. 4. CO-OPS water level time series for two Florida coastal stations: (top) Trident Pier, Port Canaveral (8721604), and (bottom) Pensacola (8729840). Both stations exhibit strong tidal oscillations with varying superimposed meteorological effects.

hydrologic mechanisms or hallucinated irrelevant biology content.

B. Multimodal Hypercube

In the second phase, the hypercube was extended to include physical data dimensions such as precipitation (IMERG, MRMS) and tidal height (CO-OPS). Measurements were keyed by station ID, geolocation, and timestamp, then discretized into time windows aligned with QA events. The multimodal hypercube thus allowed retrieval of co-located text *and* numerical slices, such as "recent heavy rainfall near a given tide gauge" or "tide peaks coinciding with strong storms."

C. RAG Configurations

Three retrieval configurations were evaluated:

1) **No Retrieval:** GPT-4o answers questions without external context.
2) **Unimodal Hypercube-RAG:** GPT-4o receives only text retrieved from the ScienceDaily hypercube.

TABLE V
CO-OPS HALF-HOURLY WATER LEVEL EDA SUMMARY.

Characteristic	Value
Rows	34,382
CO-OPS Stations	23
Time Range	2025-09-21 to 2025-10-22
Columns	`variable, units, datum, time, water_level, quality, sigma, station_id, station_name, lat, lon`
Mean Water Level (m)	0.740
Std Dev	0.384
Median	0.658
Min / Max	−0.210 / 2.862
Missing Values	0

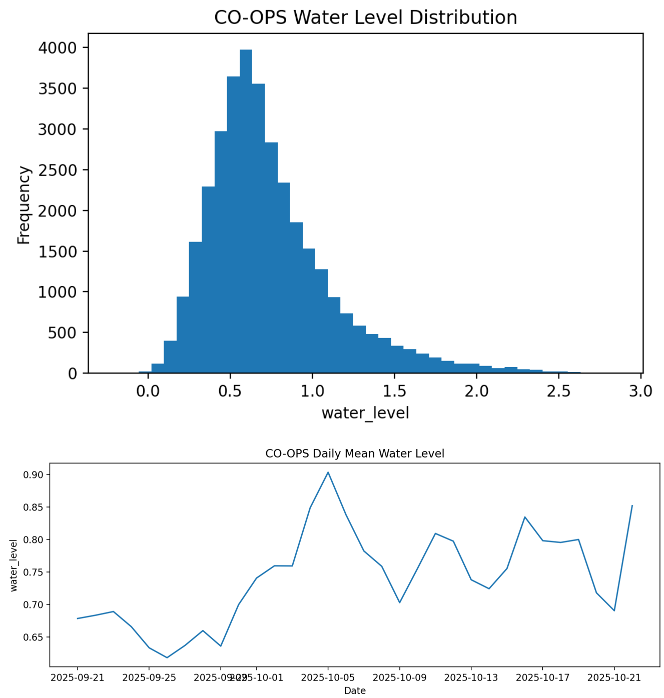

Fig. 5. CO-OPS water level EDA: (top) distribution across all stations and timestamps, and (bottom) daily mean water level (m) for Sept–Oct 2025.

3) **Multimodal Hypercube-RAG:** GPT-4o receives both text snippets and structured summaries of physical data (e.g., rainfall intensity and water level trends).

The multimodal prompt explicitly instructed the model to synthesize information from rainfall, water levels, and textual descriptions, and to explain how multiple drivers interact.

VII. Evaluation Results

Answer quality was assessed with an LLM-based judge along four dimensions on a 0–100 scale: F1 token overlap, semantic

TABLE VI
CROSS-DATASET INSIGHTS FOR COMPOUND FLOODING.

Finding	Interpretation
Rainfall Distribution	IMERG + MRMS reveal generally dry conditions with occasional high-intensity precipitation bursts.
Tidal Variability	CO-OPS water levels show clear tidal cycles independent of rainfall.
Temporal Alignment	Datasets allow time-synchronized analyses across rainfall and water level signals.
Compound Flooding Use Case	Supports reasoning about stormwater impacts, surge interactions, and rainfall–water level coupling.
Multimodal RAG Impact	Enables retrieval based on sensor-derived physical features instead of relying solely on text.

similarity, correctness, and completeness.

TABLE VII
AUTOMATIC EVALUATION METRICS ON THE COMPOUND FLOODING QA DATASET.

Metric	No RAG	Unimodal	Multimodal
F1 (%)	10.0	10.0	8.9
Semantic (%)	26.9	26.1	**72.6**
Correctness (%)	5.6	5.3	**96.0**
Completeness (%)	69.2	67.3	**94.7**

Table VII suggests several findings:

- Neither no-retrieval nor unimodal RAG produced meaningful improvements; both struggled to generate hydrologically grounded answers due to missing or irrelevant context.
- The multimodal configuration dramatically improved semantic similarity, correctness, and completeness, indicating that physical data slices provided a strong grounding signal.
- F1 scores remained low overall, reflecting diverse legitimate phrasings, but qualitative review confirmed that multimodal answers were substantially more informative and domain-appropriate.

VIII. Discussion

A. Impact of Corpus Quality

The failure of the ScienceDaily-based unimodal hypercube underscores that RAG performance is limited by the relevance and density of the underlying corpus. For compound flooding, curated sources from NOAA, USGS, FEMA, and coastal hazard literature are more appropriate than general science news.

B. Value of Multimodal Grounding

The success of the multimodal configuration suggests that physical data can strongly anchor LLM reasoning even when textual context is imperfect. When the model is told explicitly that "3-hour rainfall totals exceed a threshold while tide levels are simultaneously elevated," it can more reliably connect those conditions to mechanisms such as backwater effects and reduced drainage capacity.

C. Toward Explainable Alerts

The architecture naturally supports human-understandable alerts. A future alert message might state:

"Compound flood risk is elevated for the next 6 hours near Station 8729840 because (1) storm surge is predicted to coincide with peak tide, (2) IMERG and MRMS both indicate heavy rainfall upstream, and (3) historical events with similar patterns produced moderate flooding in low-lying areas."

Such explanations can also surface disagreements, e.g., when MRMS shows localized extreme rainfall that IMERG underestimates, or when neighboring tide gauges diverge.

IX. Future Work and Roadmap

Short-term priorities include:

- Replacing or augmenting the ScienceDaily corpus with hydrology-focused sources.
- Automating data ingestion and hypercube updates for near-real-time operation.
- Building a simple dashboard for querying locations, viewing sensor conditions, and inspecting Hypercube-RAG explanations.

Longer-term extensions include:

- Incorporating satellite imagery and flood maps via CLIP-style vision–language models.
- Maintaining a longer-term hypercube spanning multiple years of events for case-based reasoning.
- Involving domain experts to score explanations for scientific correctness and operational usefulness.
- Coupling Hypercube-RAG with physics-based models or physics-informed neural surrogates to support rapid scenario analysis.

X. Conclusion

This paper described progress toward a Compound Flood Alert System built on a multimodal Hypercube-RAG framework. A text-only experiment showed that a large but off-topic corpus fails to support meaningful hydrologic reasoning. By integrating IMERG, MRMS, and CO-OPS data into the hypercube, the system gained access to physically meaningful signals that substantially improved answer correctness, completeness, and semantic similarity on a Compound Flooding QA dataset.

Future work will focus on corpus curation, real-time pipelines, user-facing interfaces, and physics-aware extensions, with the broader goal of delivering an explainable, data-driven alerting framework for communities facing increasing compound flood risks.

References

[1] J. Shi, S. Zhou, B. Jin, W. Hu, R. Tian, S. Wang, G. Narasimhan, and J. Han, "Hypercube-based retrieval-augmented generation for scientific question-answering," *arXiv preprint* arXiv:2505.19288, 2025.

[2] R. Yu, S. Luo, R. Ghosh, L. Li, Y. Xie, and X. Jia, "RAG for geoscience: What we expect, gaps and opportunities," *arXiv preprint* arXiv:2508.11246, 2025.

[3] A. Radford *et al.*, "Learning transferable visual models from natural language supervision," in *Proc. ICML*, 2021, pp. 8748–8763.
[4] F. Liu *et al.*, "RemoteCLIP: A vision–language foundation model for remote sensing," *arXiv preprint* arXiv:2306.11029, 2023.
[5] NASA, "Global Precipitation Measurement Mission (GPM)," https://gpm.nasa.gov/.
[6] NOAA, "Center for Operational Oceanographic Products and Services (CO-OPS)," https://tidesandcurrents.noaa.gov/.
[7] USGS, "Water Data for the Nation," https://waterdata.usgs.gov/.