Evaluating the Reliability and Spatial Accuracy of Crowd-Sourced Flash Floods Reports

HIRES Interns: Keili Vargas
Mentors: JaeUng Yu, Paniz Peiravani
Faculty Advisors/Researchers: Naresh Devineni
Address/Affiliation - CUNY CREST High School Initiative in Remote Sensing of Earth System Engineering and Sciences (HIRES)
The City College of New York, NY 10031

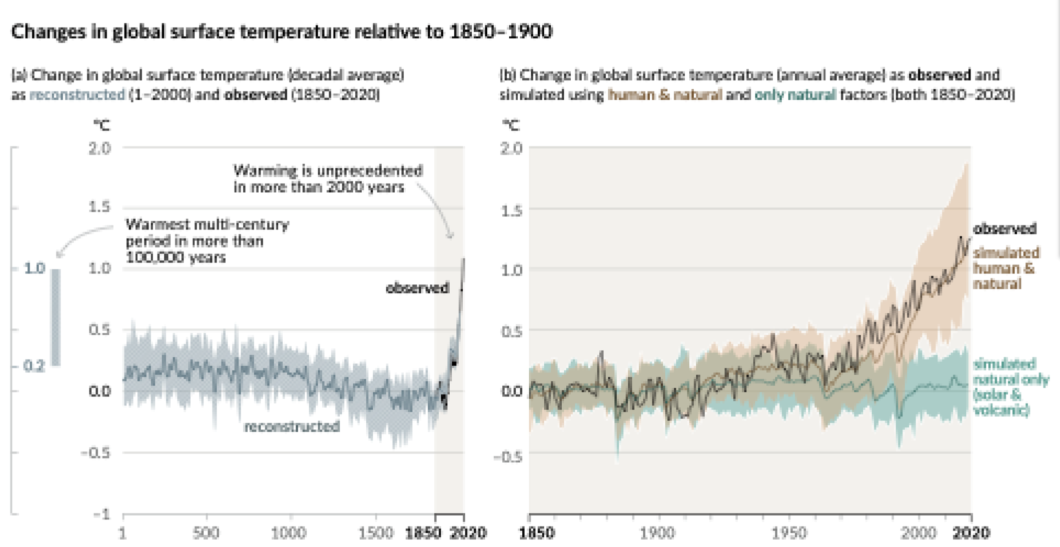
Research Objectives:

- Compare flash flood reports from each platform for the same time period and verify the actual affected locations to determine which platform has the best reliability.

"Our main objective is to test the reliability of different sources"

Background:

- The impact of climate change has been further clarified in the IPCC's Sixth Assessment Report, which attributes it to human activities.
- Numerous studies have indicated that both the intensity and frequency of rainfall are continuously increasing due to climate change.
- In urban areas, increased rainfall, combined with extensive impervious surfaces, raises the likelihood of flooding, which has led to tangible damage in New York City.



"From where do you get FLOOD WARNING?"

CROWD-SOURCED DATA

- Crowdsourced data is collected from various organizations for specific purposes and offers the advantage of obtaining large amounts of information even under limited conditions.
- However, since the reliability of the collected data may be low, only statistically significant data should be used when applying it.

Study Area:

The focus of this study is the **Greater New York area**, which includes **New York City** and its surrounding counties in **New York, New Jersey, and Connecticut**, covering approximately **13,318 square miles**. To ensure accurate spatial analysis and visualization*.

※**QGIS** was used to project and map the coordinates of both crowd-sourced and observation data.

Key locations analyzed include:

- **Bronx, Manhattan, Queens, Brooklyn, Staten Island, Long Island, Nassau County, Orange County, Putnam County, Richmond County, Suffolk County, Rockland County**, and **Westchester County**.

The primary objective is to **evaluate spatial accuracy** by calculating the **Root Mean Square Error (RMSE)** between crowd-sourced data and official observation data.

Methods:

The flowchart illustrates the process used to assess the reliability of crowd-sourced data by comparing it to verified observation data.

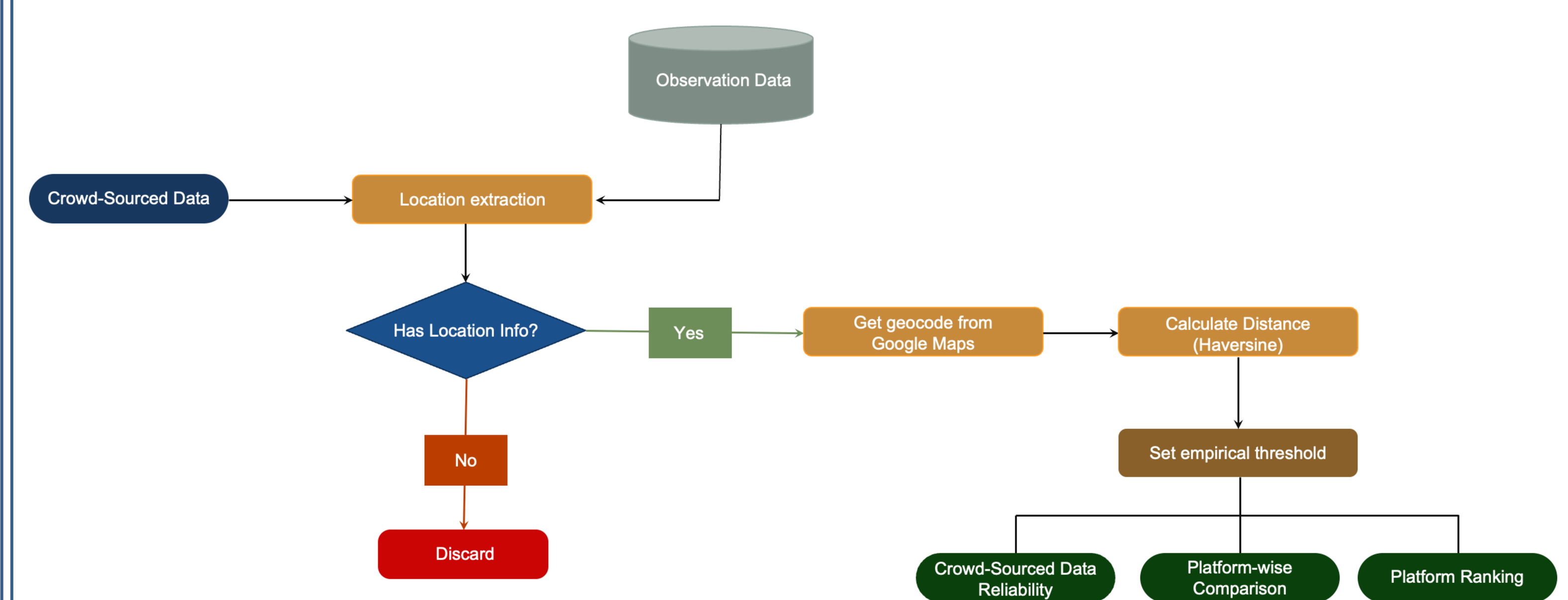
Crowd-sourced reports are first collected and passed through a **location extraction** step to identify spatial information associated with each entry. At the same time, **observation data** is used as a reference for validation.

Each report is checked for valid location information:

- If no location is found, the entry is **discarded**.
- If location data is present, the location is **geocoded** using the **Google Maps search engine** to obtain coordinates.

$$d = 2r\arcsin\left(\sqrt{\sin^2\left(\frac{\phi_2-\phi_1}{2}\right)+\cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\lambda_2-\lambda_1}{2}\right)}\right)$$

➡ **Empirical Threshold**



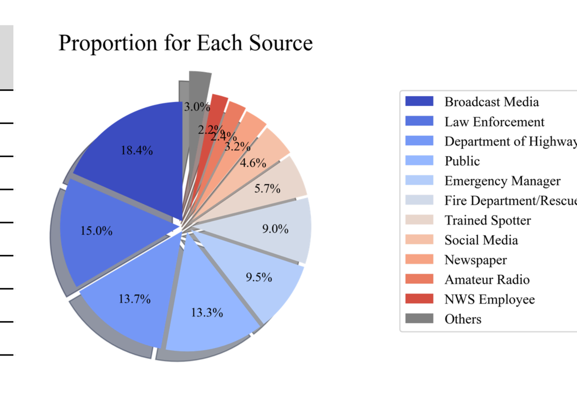
The flowchart illustrates the process used This process supports:

i. **Quantifying the reliability** of crowd-sourced data,
ii. **Comparing platforms** based on data accuracy,
iii. **Ranking platforms** according to their spatial reporting reliability.

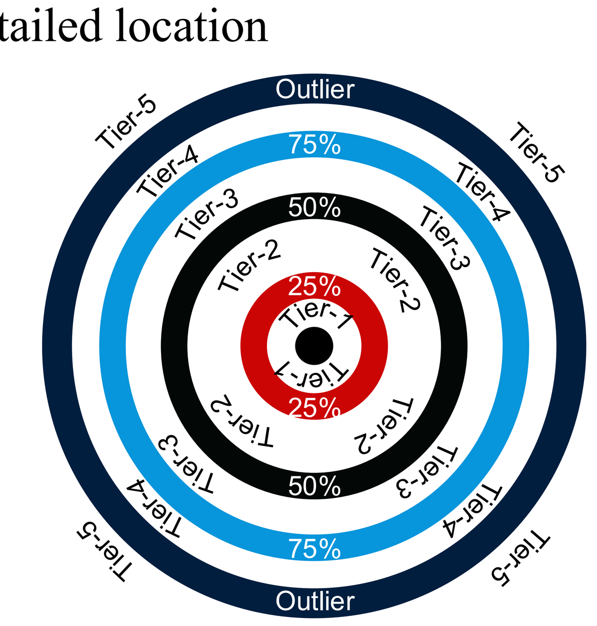
Results and Observations:

- This study utilizes data obtained from the NOAA Storm Data publication, specifically compiling flash flood events and corresponding reports related to the Greater New York area.

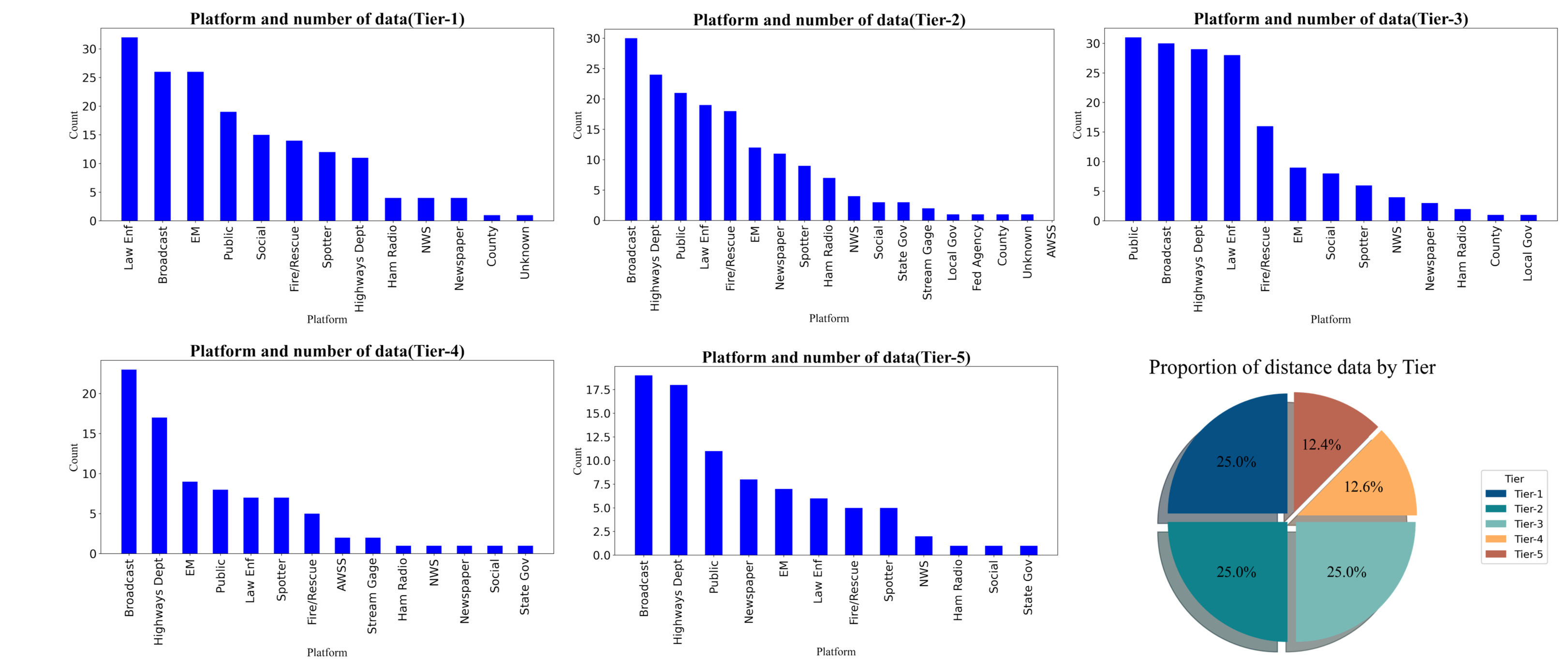
No.	Platform	Num of Data	No.	Platform	Num of Data
1	Other Federal Agency	1	10	Newspaper	19
2	AWSS	2	11	Social Media	27
3	Local Official	2	12	Trained Spotter	34
4	Unknown	2	13	Fire Department/Rescue	53
5	County Official	3	14	Emergency Manager	56
6	River/Stream Gage	4	15	Public	79
7	State Official	4	16	Department of Highways	81
8	NWS Employee	13	17	Law Enforcement	89
9	Amateur Radio	14	18	Broadcast Media	109



- Out of a total of 949 events, the analysis was conducted using only the 676 events for which contained detailed location information.
- The distances from the observation locations were divided into five tiers based **on the empirical threshold** - 25th percentile(0.185km), 50th percentile(0.778km), 75th percentile(2.998km), and upper outliers(7.218km) - of the entire dataset.



- In the case where ranking is based solely on the number of Tier-1 events—those closest to the observation points—there is a tendency to place more confidence in datasets with a larger total number of events. To address this, an approach that accounts for the total number of events was applied.



Conclusions and Discussions:

- In this study, the reliability of various platform reports regarding flash flood locations was evaluated by comparing them to the actual event locations.
- As shown in the table, the comparison using general statistical measures revealed that it is difficult to make fair comparisons due to the varying amounts of data available across platforms.

Source	Broadcast Media	Law Enforcement	Department of Highways	Public	Emergency Manager	Fire Department/ Rescue	Trained Spotter	Social Media	Newspaper	Amateur Radio	NWS Employee
Num of Data	109	89	81	79	56	53	34	27	19	14	13
Standard Deviation	1.75	1.41	1.69	1.36	1.93	1.36	1.89	0.94	0.91	1.20	1.48
MAE	2.36	1.77	2.42	1.82	2.32	1.70	2.32	1.20	1.14	1.40	1.82
RMSE	1.58	1.07	1.74	1.21	1.28	1.03	1.36	0.74	0.69	0.72	1.07

- Therefore, the RMSE for each tier was calculated, and the inverse of each RMSE was used to assign weights to the corresponding tiers. These weights were then applied to the number of data points in each tier to determine a weighted ranking based on the relative importance of each dataset.

Step 1. Weighted Score(W_score)

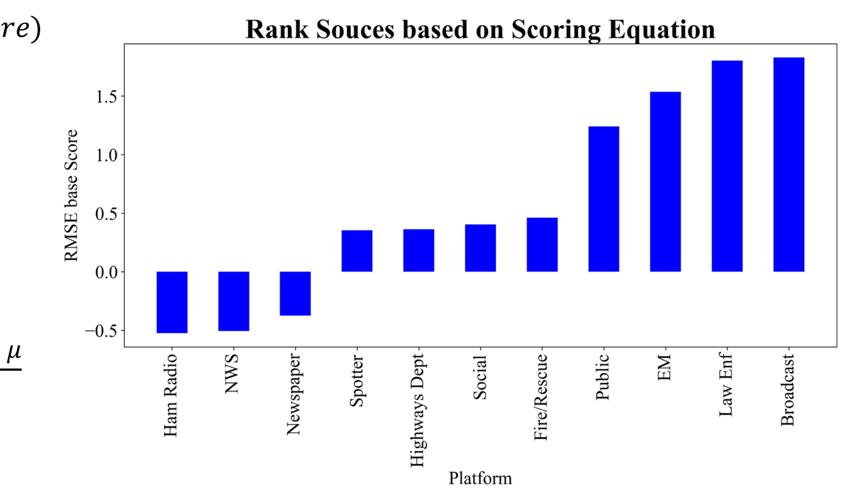
$$W - Score_{i,j} = \sum_{i=1}^{N_p}\sum_{j=1}^{4}\frac{N_{i,j}}{RMSE_{i,j}}$$

$i = Num\ of\ Souces$
$j = Num\ of\ Tier$
$N = Number\ of\ Data$

Step 2. Z − score

$$Normailzed\ Score_i = \frac{W\_Score_i - \mu}{\sigma}$$

$\mu = Mean\ of\ W\_score$
$\sigma = SD\ of\ W\_score$



Rank	Source	Score
1	Broadcast Media	1.83
2	Law Enforcement	1.80
3	Emergency Manager	1.54
4	Public	1.24
5	Fire Department/Rescue	0.46
6	Social Media	0.40
7	Department of Highways	0.36
8	Trained Spotter	0.35
9	Newspaper	-0.37
10	NWS Employee	-0.51
11	Amateur Radio	-0.52

Future Studies:

- This study aims to enhance the reliability of report data from each platform.
- A re-evaluation will be conducted by expanding the spatial scope to include flash flood data across the entire United States.
- In future research, the results derived from this method will be applied to a language model to evaluate model performance by platform and assess its applicability.

References:

Masson-Delmotte, Valérie, et al. "Climate change 2021: the physical science basis." *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* 2.1 (2021): 2391.
NYC Hazard Mitigation Plan, https://nychazardmitigation.com/documentation/hazard-profiles/flooding/
National Oceanic and Atmospheric Administration, Strom Events Database ,https://www.ncdc.noaa.gov/stormevents/

This summer research project was supported by Pinkerton Foundation (HIRES), The authors would like to express their gratitude to their faculty advisors Naresh Devineni, mentors JaeUng Yu and Paniz Peiravani. The statements contained within the poster are not the opinions of the funding agency or the U.S. government but reflect the author's opinions.