Research Letter

# Addressing Climate Hazards and Resilience in the Greater Boston Inner Core: Semiautomated Content Analysis of Municipal Reports

Nina Cesare[1], PhD; Chad W Milando[2], PhD; Allison M James[2], MS; Muskaan Khemani[2], MPA; Madeleine K Scammell[2], DSC; Beverly Ge[2], BA; Mariangelí Echevarría-Ramos[3], MPH; Amruta Nori-Sarma[4], PhD; Jonathan I Levy[2], SCD

[1]Center for Health Data Science, School of Public Health, Boston University, Boston, MA, United States
[2]Department of Environmental Health, School of Public Health, Boston University, Boston, MA, United States
[3]Mystic River Watershed Association, Arlington, MA, United States
[4]T.H. Chan School of Public Health, Harvard University, Boston, MA, United States

**Corresponding Author:**

Nina Cesare, PhD
Center for Health Data Science
School of Public Health, Boston University
715 Albany Street
Boston, MA 02118
United States
Phone: 1 617-358-2281
Email: ncesare@bu.edu

## Abstract

Semiautomated web scraping and content analysis tools can be used to identify and summarize municipal climate reports, lending high-level insight into policy priorities across broad geographic and political units. A semiautomated content analysis of reports within the Greater Boston Inner Core uncovers homogeneity in climate priority and variation in community outreach efforts. Across the study area, reports highlighted floods and storms as focal hazards. Surveys and community meetings were common outreach tools, but municipalities varied in their approach. The analytic framework and tools used in this analysis may be of use to others seeking to improve existing environmental policy or coordinate policy decisions across communities.

***JMIR Public Health Surveill 2026;12:e100771;*** doi: 10.2196/100771

**Keywords:** climate resilience; community planning; hazard identification; outreach strategies; content analysis

## Introduction

The frequency and severity of extreme weather events is increasing, and municipalities respond with strategies to protect residents [1]. Although communities are unique, there may be overlaps in how neighboring communities perceive hazards and plan resilience strategies. Therefore, analysis of municipal reports can provide valuable insights for coordination.

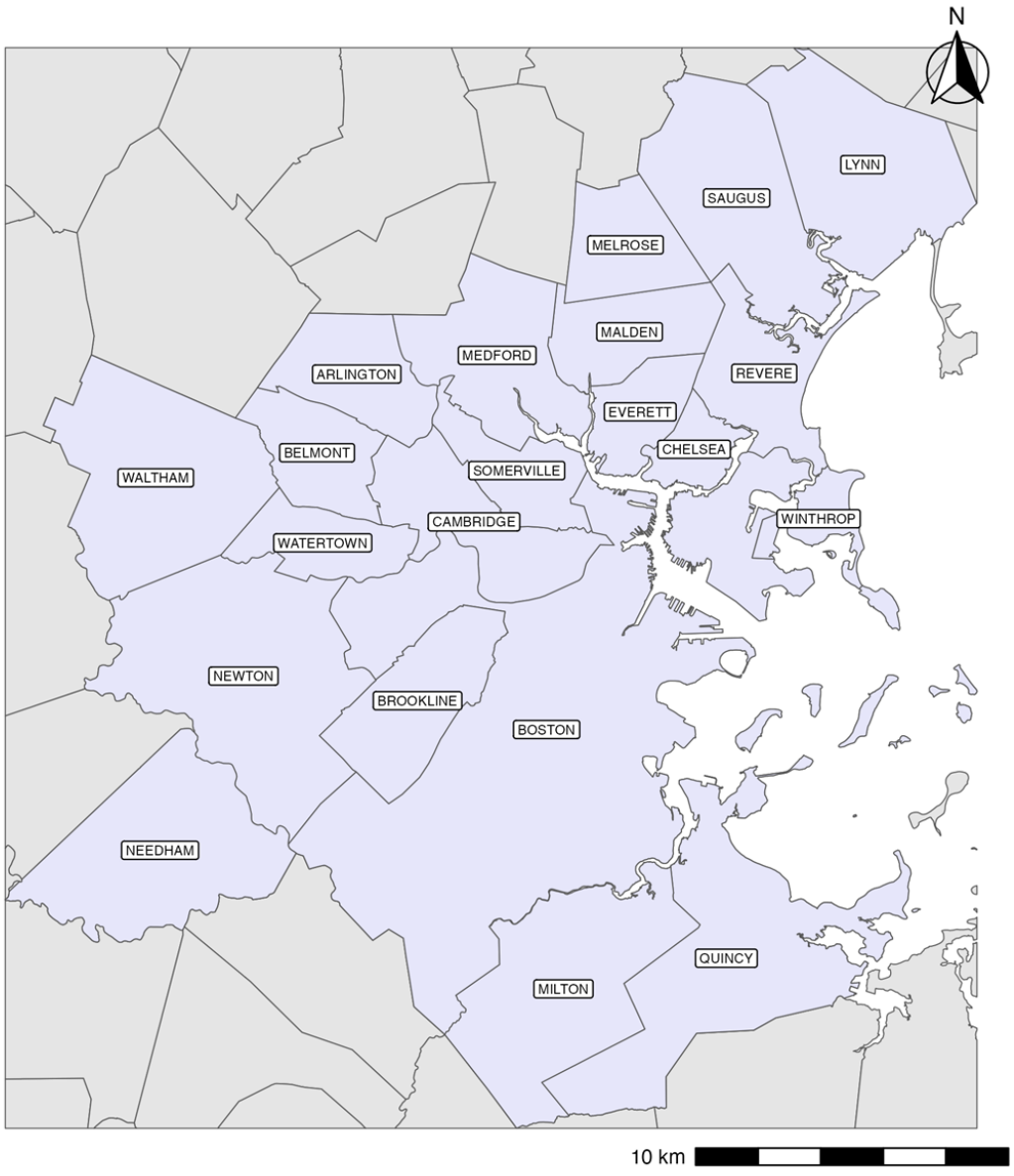
Manual synthesis of municipal documents within a larger metropolitan area is time-consuming. Individual documents may differ in length, format, and online location. These challenges affect the consistency and completeness of identified hazards.

We present a semiautomated framework for acquiring relevant text and synthesizing climate hazards of concerns and outreach strategies used by communities within a given region. We employ API-based Google searches and content analysis, the process by which text-rich documents are converted into text that is computer-legible [2], to assess concerns raised by municipalities across the Greater Boston area [3]. This framework can be used to synthesize climate concerns across other major metropolitan areas.

## Methods

The 21 municipalities in the Greater Boston region (Figure 1) contain 1.6 million individuals across municipalities that range in population size (18,558 to 689,326) and median household income (US $68,331 to US $154,398). Common environmental hazards include urban heat islands and flooding [4]. We analyzed municipal reports addressing extreme weather resilience published between 2017 and 2024.

**Figure 1.** Map of the Greater Boston Inner Core.




We searched for climate resilience and hazard mitigation plans manually (Figure S1 in Multimedia Appendix 1) and via a Massachusetts government repository [5]. We then used Google's Programmable Search Engine API [6], which provides search results without bias from user data [7]. We created queries combining community name and terms related to reporting, for example, "Boston" AND "climate" AND ("hazard" OR "resilience" OR "vulnerability") AND "report."

These queries resulted in 751 documents, which we filtered for relevance using Python (version 3.13.13) and R (version 4.5.1; R Foundation for Statistical Computing) scripts. We excluded reports where "Massachusetts" was not the most common US state mentioned, the most common town was not a Massachusetts town, and the words "community" or "climate" did not appear. We reviewed the remaining reports and kept 123. We took a convenience sample of 1 report per town and manually created a dictionary of key terms related to hazards of concern and strategies for community outreach (see Tables S1 and S2 in Multimedia Appendix 1). We then tokenized and normalized each PDF, calculated the frequency of each dictionary-based term, and calculated the total number of hazard or outreach terms within each document.

We then calculated metrics to compare the frequency of referenced hazards and outreach efforts across towns with varying quantities of climate documents and across documents with different lengths. We assessed the percentage of documents that contained a specific hazard or outreach term, calculated the proportion of terms related to each term, and calculated summary statistics of these proportions in the full region and within each municipality.

### Ethical Considerations

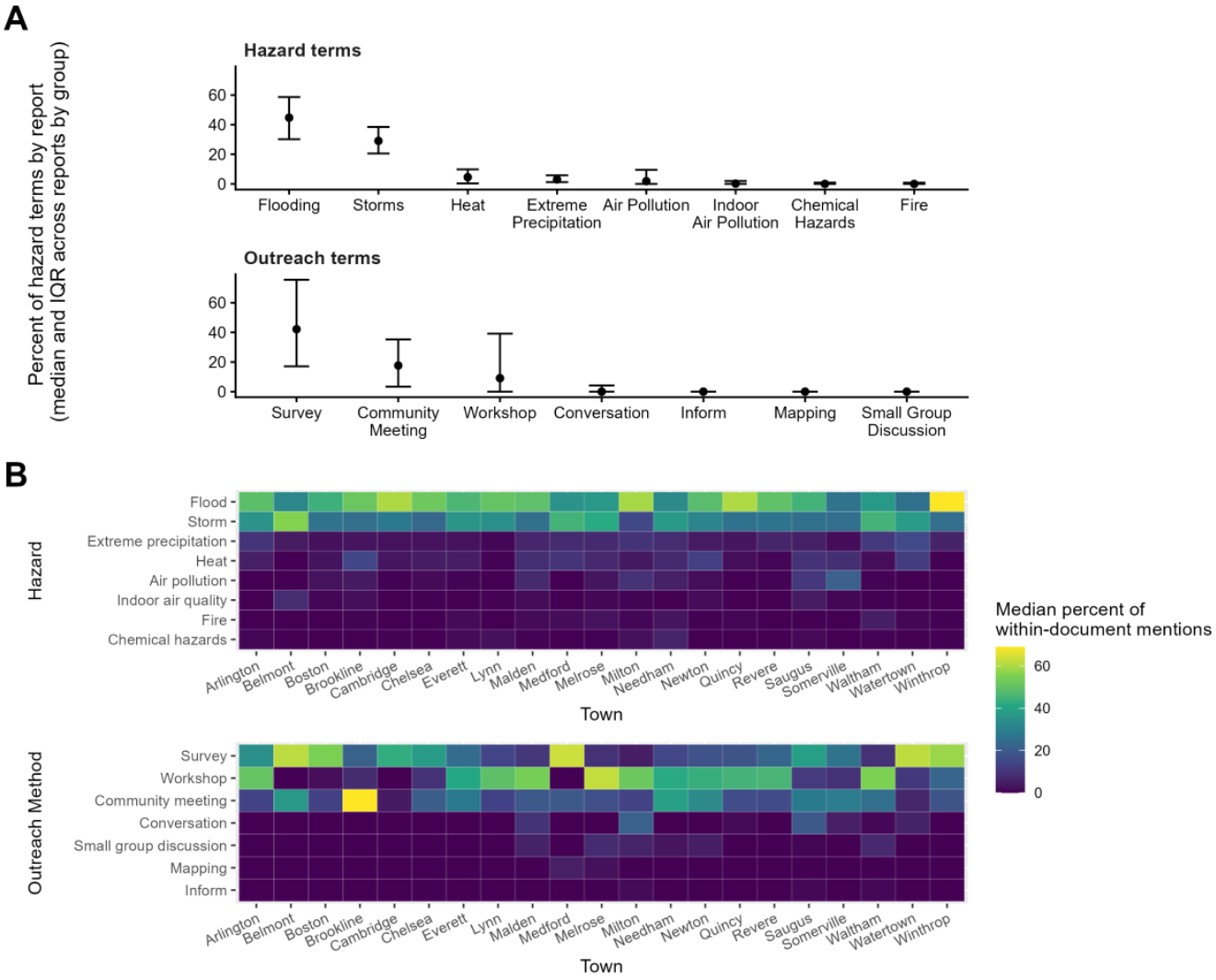
This study used publicly available reports to characterize priorities and trends across communities. Because these data were not gathered from individuals and involve no direct interaction with human participants, this study does not require human subjects research approval through the Boston University Institutional Review Board.

## Results

### Hazards of Concern

Hazard frequencies were similar across communities (Figure 2A). Flooding was mentioned proportionately the most within reports, accounting for a median of 44.7% (IQR 30.2%-58.7%) of mentioned hazards per report. Next most common were storms (29.1%, IQR 20.5%-38.4%), heat (4.5%, IQR 0.3%-9.8%), and extreme precipitation (3%, IQR 1.2%-5.8%).

**Figure 2.** (A) Medians and IQRs of the per-report percentage of hazard terms represented by all reports in the Greater Boston Inner Core. (B) A heatmap representing the median per-report percentage of hazard and outreach terms across municipalities within the Greater Boston Inner Core.




Flooding and storms were also mentioned most often across reports, with each hazard appearing in 97.6% of reports. Extreme precipitation and heat were the third (86.2%) and fourth (79.7%) (see Table S3 in Multimedia Appendix 1). Air pollution, indoor air, chemical hazards, and fire were discussed less frequently.

### Reported Outreach Strategies

Surveys had the highest mentions by proportion within reports (Figure 2B), with a median proportion of 42.2% (IQR 17.1%-75.5%) of all outreach terms. Community meetings amounted to 17.6% of per-report mentions (IQR

3.3%-32.3%), followed by workshops with 9.1% (IQR 0%-39.2%).

Surveys were also mentioned most often across reports, appearing in 94.3% of reports. Community meetings and workshops were also frequently mentioned, appearing in 78% and 65.9% of reports, respectively (see Table S3 in Multimedia Appendix 1).

Outreach strategies varied by community. For example, the IQR of per-report percent of survey terms was 17.1% to 75.5%, and for community meeting it was 3.3% to 35.3%.

# Discussion

## Principal Findings

Semiautomated methods captured and summarized municipal text addressing climate resilience in the Greater Boston Inner Core. Concerns related to climate hazards were consistent across the study area and predominantly cited flooding and storms, despite variation in population size, median income, and proximity to water. Though helpful for policy prioritization, homogeneous climate concerns may signify gaps in acknowledgment of other key hazards—for example, heat [8]. Reported outreach strategies were less consistent across communities, which may indicate opportunities for coordination across cities and towns.

## Comparison With Existing Studies

This study provides an approach to surveying past initiatives that is more scalable than a manual literature review and leverages content analysis as an effective means of analyzing municipal broadcasts. It builds upon Divay and Micheau's [9] content analysis of citizen engagement and recognition by focusing on a policy area. It shares content analysis methodology with the assessment by Tang et al [10] of smart city archetypes and adds a semiautomated approach to document retrieval and analysis.

## Limitations

While our Google search methodology is generalizable, search criteria and restrictions may have impacted our sample. Additionally, our dictionary approach may miss references to hazards or outreach and does not account for hazard co-occurrence. While our qualitative evaluation of 21 initial climate resilience reports helped us anticipate document content, we may have missed terminology or conversations unique to specific municipalities. Additionally, the frequency of community outreach terms does not reflect the breadth, depth, or impact of outreach and may not represent on-the-ground engagement. Finally, we cannot infer the rationale behind highlighted hazards without additional policy context.

## Conclusions

As weather events become more extreme and communities work in parallel to develop action plans, community coalitions may use semiautomated techniques to identify and synthesize publicly available municipal reports related to climate hazards. With limited resources available for climate mitigation, pooling concerns could lead to shared strategies that benefit multiple communities. Additional research may explore how variation in policy discussion is tied to municipal characteristics such as population density, income, or proximity to natural environments.

### Funding

Research funding included the United States Environmental Protection Agency (EPA) under assistance agreement RD-84048001 to Boston University, until the assistance agreement was terminated in May 2025 for no longer effectuating agency priorities. Additional support was provided by the Boston University Research Continuity Fund and CHDS Supplemental Research Award.

### Data Availability

All processing scripts are available on GitHub [11]. Data will be made available upon request.

### Conflicts of Interest

None declared.

### Multimedia Appendix 1.

1Flowchart of document selection process; search terms used for hazards analysis; search terms used for outreach content analysis; and number of reports containing each hazard and outreach term.
[DOCX File (Microsoft Word File), 148 KB-Multimedia Appendix 1.]

### References

1. McDowell G, Ford J, Jones J. Community-level climate change vulnerability research: trends, progress, and future directions. Environ Res Lett. Mar 1, 2016;11(3):033001. [doi: 10.1088/1748-9326/11/3/033001]
2. Grimmer J, Stewart BM. Text as data: the promise and pitfalls of automatic content analysis methods for political texts. Polit Anal. 2013;21(3):267-297. [doi: 10.1093/pan/mps028]
3. About MAPC. Metropolitan Area Planning Council. URL: https://www.mapc.org/aboutus/ [Accessed 2026-11-12]
4. Massachusetts priority climate action plan. Office of Climate Innovation and Resilience and Department of Transportation; 2024. URL: https://www.epa.gov/system/files/documents/2024-03/massachusetts-pcap.pdf [Accessed 2026-11-05]

5. Municipal Vulnerability Preparedness (MVP) Program Planning Reports. Mass.gov; URL: https://www.mass.gov/info-details/municipal-vulnerability-preparedness-mvp-program-planning-reports [Accessed 2025-03-06]
6. Overview | Programmable Search Engine. Google Dev; URL: https://developers.google.com/custom-search/docs/overview [Accessed 2024-07-31]
7. Piasecki J, Waligora M, Dranseika V. Google Search as an additional source in systematic reviews. Sci Eng Ethics. Apr 2018;24(2):809-810. [doi: 10.1007/s11948-017-0010-4] [Medline: 29249022]
8. Nori-Sarma A, Cartwright BM, Zanobetti A, Pastwa A, Stucky N, Willis MD. Impact of the summer 2025 heatwaves on emergency department visits in the USA. Nat Health. 2026;1(11):1380-1388. [doi: 10.1038/s44360-026-00194-y]
9. Divay G, Micheau M. Recognizing citizens in municipal management: an exploratory study based on a content analysis of municipal websites in the province of Quebec. Int Rev Adm Sci. Dec 2017;83(4):773-788. [doi: 10.1177/0020852315608251]
10. Tang Z, Jayakar K, Feng X, Zhang H, Peng RX. Identifying smart city archetypes from the bottom up: a content analysis of municipal plans. Telecomm Policy. Nov 2019;43(10):101834. [doi: 10.1016/j.telpol.2019.101834]
11. GitHub - cmilandoacresnlp. GitHub. URL: https://github.com/cmilando/acresNLP [Accessed 2026-11-05]

*Edited by Christos Argyropoulos; peer-reviewed by Alessandra Colocci, Jin-Ding Lin; submitted 08.May.2026; final revised version received 17.Sep.2026; accepted 14.Oct.2026; published 12.Nov.2026*

*Please cite as:*
*Cesare N, Milando CW, James AM, Khemani M, Scammell MK, Ge B, Echevarría-Ramos M, Nori-Sarma A, Levy JI*
*Addressing Climate Hazards and Resilience in the Greater Boston Inner Core: Semiautomated Content Analysis of Municipal Reports*
*JMIR Public Health Surveill 2026;12:e100771*
*URL: https://publichealth.jmir.org/2026/1/e100771*
*doi: 10.2196/100771*

© Nina Cesare, Chad W Milando, Allison M James, Muskaan Khemani, Madeleine K Scammell, Beverly Ge, Mariangelí Echevarría-Ramos, Amruta Nori-Sarma, Jonathan I Levy. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.