9^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2023, Copenhagen, Denmark

Comparing News Headlines on Heatwaves in the US and Mexico: A BERTopic Analysis

Keywords: heatwaves, headlines analysis, news media, climate change, topic modeling

Extended Abstract

The first decade of the twenty-first century was marked by an unparalleled increase in the frequency and intensity of extreme weather events, a trend that has persisted into the current decade. With record-breaking temperatures, resulting in devastating impacts around the world, heatwaves have captured the public's attention as an important socio-ecological phenomenon.

News media are a central arena in which public understanding of environmental change is continually constructed. Through issue selection, coverage levels, and framing, news media create a context and boundaries through which meanings are produced in public discourse [1]. Given the significant role news media play in shaping public perception and influencing policy responses, numerous studies have assessed how newspapers report climate-related issues. However, existing research on news reporting is still largely characterised by a focus on climate change, and there is little research addressing heatwaves as a subset of climate reporting [2].

With the rapid digitalisation of media, legacy news outlets, which are defined as established publishing and broadcasting institutions that did not originate on the web, are increasingly leveraging digital platforms. Despite the rise of social media, legacy media outlets remain influential due to their longevity, reach, long-term audience building, and editorial quality control. They are still considered reliable and credible sources of news. According to a survey by Reuters digital news report 2017, only a quarter of all respondents thought social media was a trustworthy source of facts, compared to 40% for news media [3].

This study aims to analyse how online news headlines contribute to creating particular representations of heatwaves. Headlines are of particular interest not only because they are able to capture attention quickly, but also because they can reveal social and cultural meanings in society and portray specific societal ideologies [4]. The study has three main research objectives: 1) to identify similarities and differences in news headline reporting of heatwaves in global North and South countries and how it may reflect social, cultural, and political differences; 2) to examine whether and how various causes of heatwaves are represented in the headlines; and 3) to determine the extent to which countries prioritise social, environmental, economic, and political impacts of heatwaves in news headlines. By analysing topics and themes in the headlines, this research can make useful inferences about how heatwaves may be represented, and whether dimensions such as causes, consequences, responsibility, and responses are prioritised. This study contributes to the growing literature on the discursive construction of environmental issues in the media and has implications for communication strategies that drive relevant national and international remedial actions.

The analysis focuses on online news coverage in five countries: India, Mexico, Nepal, the US, and the UK. The selection of these countries was driven by the need to include Global South nations that have experienced heatwaves in the analysis. Additionally, they were chosen due to their varying vulnerabilities, and adaptive capacities, differing socio-cultural contexts and connections to historical and contemporary factors that exacerbate vulnerabilities. National news sources were selected based on their reach and popularity, credibility, and political affiliations. The selection of news sources and keywords was informed by discussions with native speakers, media personnel, and analysis of media landscapes in the countries of interest. A

total of 14,945 articles were collected from 24 news sources covering the period between 2001 and 2022, using NewsAPI. The study began with an initial analysis of online news articles from the US and Mexico, with a total of 548 and 1234 articles respectively. For this analysis, we used the BERTopic topic modelling algorithm, which is considered state-of-the-art due to its use of transformers to encode language and its efficiency in capturing the context of text data compared to other topic models like Latent Dirichlet Allocation (LDA) [5]. BERTopic's modular structure allows for various methods to be used for embedding, reducing dimensionality, and clustering of topics. We used the sentence transformer model for embedding, UMAP for dimensionality reduction, and HDBSCAN method for clustering of topics [6].

The analysis revealed that headlines in both countries primarily focused on meteorological topics (weather forecasts, temperature rise) and event-specific topics (forest fires, record heatwaves in various countries), as well as topics related to climate change, Figure 1. In the US, heatwaves were typically presented within a broader construct, focusing on five main themes: weather reports and forecasts, unusual temperature and fire events, links to climate change, adaptation measures, and political discussions, Figure 2. Headlines drew attention to the abnormality of reported temperatures and explicitly communicated them as "dangerous." Heatwaves in other countries were reported thus construing them as a common problem, despite differing social conditions that often cause weather-related crises. Climate change-centric topics were used to locate causality of heatwaves, while other elements and stressors were unexpressed. Adaptation measures focused on quick fixes, such as "trees," "roofs," and "panels," rather than systemic changes, such as heat-sensitive urban and infrastructure planning. Contrary to expectations, references to health impacts were absent in the headlines.

In contrast, news in Mexico covered a broader range of themes in their headlines, including impacts on health and infrastructure, impacts on energy supply, prices, and water shortages, as well as heatwave warnings and alerts, Figure 3. References to health impacts (deaths) and other types of environmental impacts (fires, floods, droughts) were also present. The reporting of energy price changes and water shortages implied notions of how heatwaves affect day-to-day life. Warning prompts were primarily expressed in phrases such as "alerts" and "warnings." Similar to US headlines, heatwaves and their impacts in other countries were reported.

These findings have important implications for understanding how the discourse of extreme events is shaped and framed in ways that determine responsibility, identify solutions, and select adaptation and mitigation strategies.

References

[1] Nisbet E.C. Newman, T.P. and M.C. Nisbet. Climate change, cultural cognition, and media effects: Worldviews drive news selectivity, biased processing, and polarised attitudes. *Public Understanding of Science*, 27(8):985–1002, 2018.

[2] Ivanova A. Schmidt, A. and M.S. Schäfer. Media attention for climate change around the world: A comparative analysis of newspaper coverage in 27 countries. *Global Environmental Change*, 23(5):1233–1248, 2013.

[3] Fletcher R. Kalogeropoulos A. Kleis Nielsen R. Newman, N. and D. Levy. Digital News Report, Reuters Institute & University of Oxford, 2017.

[4] R. Taiwo. Language, ideology and power relations in nigerian newspaper headlines. *Nebula*, 4(1):218–245, 2007.

[5] R. Egger and J. Yu. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify twitter posts. *Frontiers in Sociology*, 7, 2022.

[6] M. Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure. In *arXiv*, 2022.

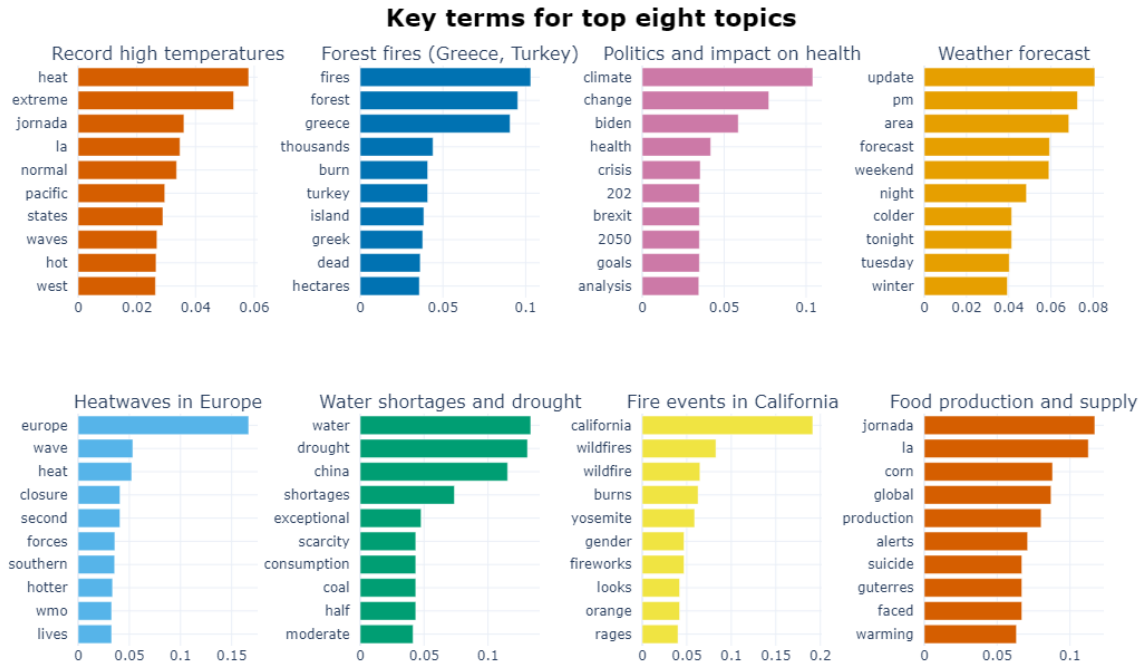

Figure 1: Key terms for top eight topics in headlines from online news in the US and Mexico.

Figure 2: Visualisation of documents inside topics in headlines from online news in the US.

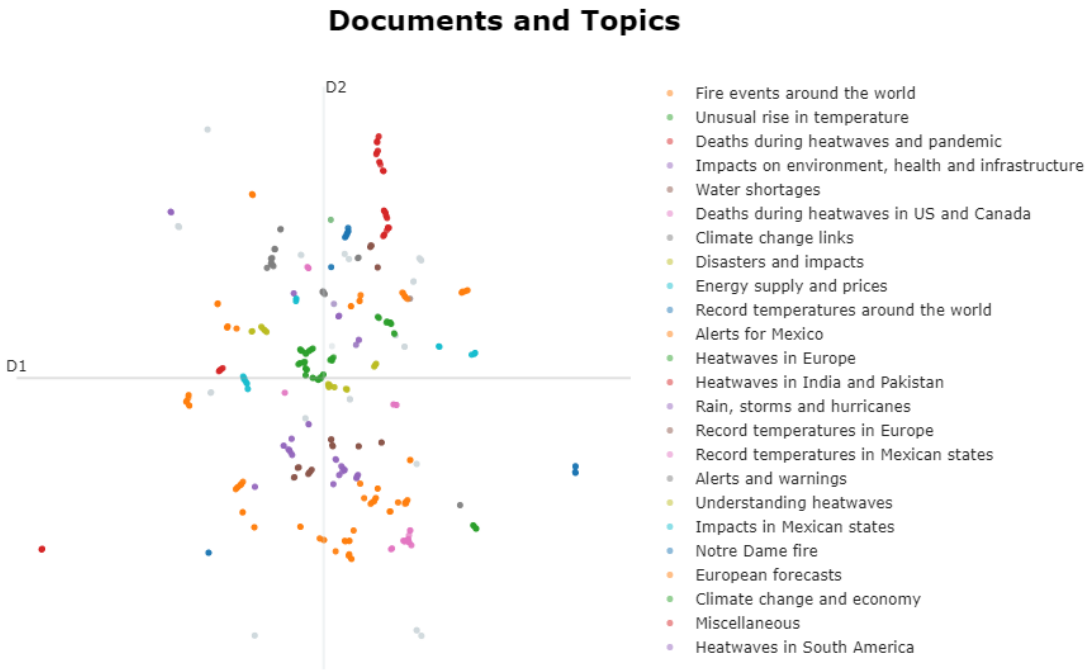

Figure 3: Visualisation of documents inside topics in headlines from online news in Mexico.