Proceeding Paper

Can Greek Twitter Provide Insight into the Users' Opinions on the Adverse Effects on Health Caused by Climate Change? †

Vasiliki-Athina Tyrologou, Georgia Merkouriadi, Ioanna Karefyllaki, Nefeli Giannopoulou Marini, Christos Markou [ID], Konstantinos Kesanopoulos *[ID] and Anastasia Barbouni

Special Academic Unit for the Study of Smoking Behavior, Laboratory of Hygiene and Epidemiology, Department of Public and Community Health, School of Public Health, University of West Attica, 11521 Athens, Greece; pch19679187@uniwa.gr (V.-A.T.); pch19679101@uniwa.gr (G.M.); pch19679061@uniwa.gr (I.K.); pch19679029@uniwa.gr (N.G.M.); pch19679097@uniwa.gr (C.M.); abarbouni@uniwa.gr (A.B.)

* Correspondence: kkesanopoulos@uniwa.gr; Tel.: +30-2132010262
† Presented at the 16th International Conference on Meteorology, Climatology and Atmospheric Physics—COMECAP 2023, Athens, Greece, 25–29 September 2023.

Abstract: Climate change is a pressing global issue with significant impacts on human health. This study used a digital surveillance system on Twitter to investigate Greek-speaking users' sentiments and opinions regarding climate change's effects on human health. The study utilized a set of 61 keywords to collect 2.774 tweets from the last decade, which were manually analyzed for content and user type. The analysis revealed that news outlets mainly published informational tweets, whereas personal accounts expressed attitudes, opinions, perceptions, and beliefs. Moreover, the study found that Greek-speaking Twitter users are actively discussing climate change, particularly its consequences on their health.

Keywords: climate change; public health; digital surveillance systems; semantic analysis; Twitter

1. Introduction

Climate change is a topic of global concern, with increasing evidence that human activity has caused significant transformations, leading to environmental alterations and extreme weather events. Earth's temperature has risen by 1.06 °C since 1880 [1], and the five hottest years on record have been since 2015 [2]. The Intergovernmental Panel on Climate Change warns that warming beyond 1.5 °C will result in detrimental impacts such as droughts, floods, heatwaves, and sea-level rises [3]. The US experienced 310 climate disasters between 1980 and 2021, and climate change has been linked to respiratory and cardiovascular disease, premature death, and other health issues [4].

Given the significance of climate change as a worldwide issue, understanding how people perceive it and what influences their choices to mitigate or adapt to it is crucial in determining their level of engagement in climate change mitigation efforts. The use of Internet-Based Surveillance (IBS) systems, based on Infodemiology, enables the monitoring of beliefs and opinions of social media users on Public Health issues, such as climate change [5]. The present study applies a digital surveillance system on Twitter, a social media platform with over 300 million users worldwide [6], frequently used to discuss trending health-related topics, to study their opinions and sentiments. Twitter has emerged as a significant repository of health-related information on the internet, owing to the substantial volume of information exchanged by individuals and official sources. Its real-time nature and worldwide reach make Twitter a valuable resource for researchers studying public health, demonstrating the vast potential of utilizing Twitter for public health surveillance purposes [7].

The objective of this study is to develop an IBS that can be used to collect and analyze the beliefs of Greek Twitter users regarding the correlation of climate change and extreme

weather events and their consequences on human health. Furthermore, the collected data could potentially contribute to the understanding of public perception on climate change and its possible impact on human health and to inform public health stakeholders for adequate policies. The use of IBS systems, such as the one proposed in this study, is increasing due to the valuable epidemiological information they provide and their potential to contribute to the development of effective public health interventions [8–11].

2. Materials and Methods

The study material consisted of 2.774 tweets written in the Greek language (Dataset 2), retrieved from Twitter, using 61 keywords regarding human health and climate change or extreme weather events (Appendix A). The study's total period was 10 years (25 February 2013–25 February 2023). The tweets were retrieved via the Twitter API v.2 (Academic Research Track) and processed through a custom-developed LAMP (Linux, Apache, MySQL, and PHP) architecture application. The combination of the keywords that led to the study dataset underwent several modifications so that the context of the tweets would be as close to the aforementioned topics as possible, i.e., specific keywords may indeed be relevant to the study's topic but were also used in a variety of different contexts resulting in retrieving tweets that ultimately would not be relevant to the study (i.e., the words "φωτιά" or "βροχή", which mean "fire" and "rain" accordingly, were also used in poetic contexts). Alongside the above dataset (Dataset 2), another dataset was created, consisting of 326.886 tweets and based on keywords reflecting only climate change and extreme weather topics (Dataset 1).

Tweets were then processed in two different ways. The first consisted of a manual semantic analysis of a representative sample of 25% of the total tweets (695 tweets). To ensure the representativeness through the years, the number of each year's proportion to the original number of tweets was calculated, and tweets were randomly selected using an Excel Add-on. The semantic analysis was performed in order to categorize the sampled tweets into three groups based on their content: informational; tweets with content that expressed attitudes, opinions, perceptions, and beliefs; and irrelevant tweets. The first two groups of tweets were then sorted into five account types: personal, news outlets, public figures, businesses/non-profitable organizations, and other based on specific criteria (Table 1). This process was used in order to investigate the correlation between the tweet content and the tweet's account type.

Table 1. Detailed User Categorization Criteria.

Type of User Categorization	Selection Criteria
News outlets	Pages with exclusively news content. This category does not include people/bloggers/scientists/trolls, regardless of whether they only upload news or not.
Public figures	People with positions of responsibility in a workers' union, political area, scientific association, politicians, singer, dancer, author, etc.
Personal	Personal accounts that have a variety of posts, not just news-related posts (e.g., personal photos, interests and opinions, song lyrics, etc.)
Business/non-profitable organizations	Organizations, radio stations, and all kinds of accounts that promote specific purposes (not personal accounts).
Other	Personal accounts that exclusively upload news (either from their personal news website or others) as well as bloggers.

In addition to the above, an association between high tweet volume dates (peaks in tweets timeline) and tweets' contents was further investigated, in order to explore whether

high tweet volume dates could be used to indicate Twitter users' interest in specific topics or events, especially linked to the Greek topicality.

The analysis was conducted in the totality of the tweets from Dataset 2. In order to identify high tweet volume dates through the 10-year study period, the boxplot technique was used. The technique was deployed in each year separately, taking into account the difference of each year's characteristics, such as the number of tweets. For each year, the extreme outliers would indicate the highest tweet volume dates. Those dates were then isolated, and a manual semantic analysis of their tweets was conducted.

Depending on the discussed tweets' subject, the tweets were associated with current events or topics related to climate change in Greece or worldwide and were then quantified. After the above categorization, the most relevant topics or events were identified. Then, a ratio to see how many tweets were about the same topic or event on that day was calculated.

With the aim of identifying the factors that motivated users to post, tweets were analyzed to determine if there was a connection to the events (recent and/or past events) as well as if there were specific topics that activated their participation.

3. Results

By application of the proposed LAMP methodology for the topics of "Climate change" or "Extreme Weather", a total of 326.886 tweets (Dataset 1) were identified. For this dataset, 149.284 were original tweets, and 177.602 were retweets. Similarly, Dataset 2 was created by using criteria for tweets selection for "Climate change or Extreme Weather" and "Health impact", identifying 2.774 tweets in total, 1.586 of which were original tweets and 1.188 retweets.

As shown in Figure 1, the tweets from Dataset 1 are almost 117.8 times more than those of Dataset 2, depicting that the interest of Twitter users is bigger in topics regarding climate change and extreme weather effects in general than in topics related to their implications on human health. On top of that, while Dataset 1 consists of more retweets than original tweets, Dataset 2 consists of more original tweets than retweets. However, both datasets have similar tweet fluctuations within the study period, starting with lows and reaching highs around 2021 (Figure 1). It is hypothesized that the 2021 peak is related to major current affairs of that time, such as the fire season in Greece and the enormous wildfires in the area of Evia Island and around the country.
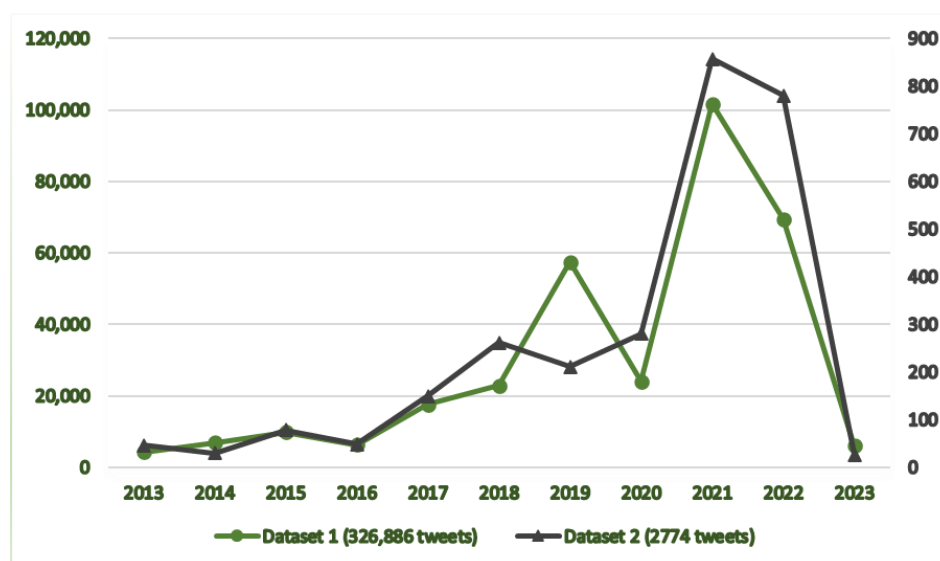

Figure 1. Tweet number fluctuations throughout the study period.

As resulted from the semantic analysis of the 25% sample of Dataset 2, 61.4% of the tweets were informational, 10.5% of them were tweets with content that expressed

attitudes, opinions, perceptions, and beliefs, and 28.1% of the tweets were irrelevant (Table 2). Regarding the account type categorizations (Table 3), an association between the user type and the tweet's content was observed, as originally expected. The analysis showed that 55.3% of the informational tweets were published by news outlets accounts, and 79.5% of the tweets that expressed attitudes, opinions, perceptions, beliefs were posted by personal accounts. On top of that, 97.2% of the news outlets accounts' tweets were original tweets, whereas 73.2% of personal accounts' tweets were retweets.

Table 2. Categorization according to the content of the tweet.

Categories	Number of Tweets
Informational	427 (61.4%)
Tweets with content that expressed attitudes, opinions, perceptions, or beliefs	73 (10.5%)
Irrelevant tweets	195 (28.1%)

Table 3. Categorization according to account type.

Categories	Number of Tweets
Personal	209 (41.8%)
News outlets	246 (49.2%)
Public figures	12 (2.4%)
Business/non-profitable organizations	25 (5.0%)
Other	8 (1.6%)

On the high tweet volume dates analysis on Dataset 2, out of a total of 640 different dates when tweets were posted throughout the 10-year study period, only 53 of them were initially identified as high tweet volume dates (peaks) using the boxplot technique.

According to the semantic analysis of the totality of the tweets for each date, 34 dates out of the 53 (64.1%) consisted of tweets related to the topicality.

Furthermore, in 28 out of these 34 (94%), the content of the tweets of each date is directly correlated to the topic, as well as a specific topicality.

4. Discussion

In the present era, the combination of Infodemiology studies and semantic analysis has emerged as a valuable tool in informing Public Health prevention policies regarding the health impact of climate change [12–14]. By employing the proposed Internet-Based Surveillance methodology on Greek Twitter over the last decade, valuable insights into the Twitter users' perceptions of the climate–health correlation can be captured and subjected to semantic analysis. Twitter users cannot represent all social media users or the general population of society, but monitoring the Greek users' opinions alongside with other social media users' opinions on climate change and human health can improve the representativeness of the public perception.

Moreover, it is observed that Greek Twitter provides primarily informative content, while allowing the users to engage with it, and thus can potentially be a useful tool in providing insight into the attitude/opinions on the adverse effects on health caused by climate change and extreme weather events.

The study results also demonstrate that the proposed methodological model can effectively depict Twitter users' activities and serve as a means to raise awareness about the adverse impacts of climate change and disseminate health-related information. It could also incite more prominent active participation from its users, regarding environmental protection.

Therefore, a real-time monitoring system could prove highly beneficial in capturing people's attitudes, beliefs, and the latest events that influence the Twitter users' opinions.

In essence, the present study captures the interest of Greek Twitter users in the effects of climate change on health as well as their awareness through their activity on the social media platform in relation to climate-related events.

However, it is important to acknowledge the limitations of the proposed methodology. The current model was developed based on Greek language peculiarities; thus, its applicability to other languages and social media platforms requires further evaluation. On top of that, this model should not be considered an accurate representation of the public's opinion due to its limited user base; nonetheless, it can provide information that can be used effectively to take specific actions or make informed decisions by the state. Additionally, the reliance on manual semantic analysis poses limitations in analyzing a large number of tweets, necessitating the future incorporation of automated semantic Artificial Intelligence algorithms validated for this purpose, keeping in mind all of the concerns Artificial Intelligence may pose.

Appendix A

Datasets' keywords

Health effects keywords: υγεια, επιπτωση, επιπτωσεις (health, effect, effects)

Climate change keywords: πλημμυρα, ξηρασια, πυρκαγια, φαινομενο του θερμοκηπιου, τρυπα του οζοντος, κλιματικη αλλαγη, κλιματικη κριση, αλλαγη του κλιματος, υπερθερμανση, ακραια θερμοκρασια, πλημμυρες, ξηρασιες, πυρκαγιες, ακραιες θερμοκρασιες, πλημμυρας, ξηρασιας, πυρκαγιας, φαινομενου του θερμοκηπιου, τρυπας του οζοντος, κλιματικης αλλαγης, κλιματικης κρισης, αλλαγης του κλιματος, υπερθερμανσης, ακραιας θερμοκρασιας, πλημμυρων, ξηρασιων, πυρκαγιων, ακραιων θερμοκρασιων (flood, drought, fire, greenhouse effect, ozone hole, climate change, climate crisis, overheating, extreme temperature, floods, droughts, fires, extreme temperatures and other grammatical variations of the above words in Greek)

Extreme weather keywords: καταιγιδα, καυσωνας, τυφωνας, κυκλωνας, ακραιο καιρικο φαινομενο, καταιγιδες, καυσωνες, τυφωνες, κυκλωνες, ακραια καιρικα φαινομενα, καταιγιδας, καυσωνα, τυφωνα, κυκλωνα, ακραιου καιρικου φαινομενου, καταιγιδων, τυφωνων, κυκλωνων, ακραιων καιρικων φαινομενων (storm, heat wave, hurricane, cyclone, extreme weather phenomenon, storms, heat waves, hurricanes, cyclones and other grammatical variations of the above words in Greek)

References

1. Global Climate Report. 2022. Available online: https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/202213 (accessed on 15 May 2023).
2. Watts, N.; Amann, M.; Arnell, N.; Ayeb-Karlsson, S.; Beagley, J.; Belesova, K.; Boykoff, M.; Byass, P.; Cai, W.; Campbell-Lendrum, D.; et al. The 2020 report of The Lancet Countdon on health and climate change: Responding to converging crises. *Lancet* **2021**, *397*, 129–170. [CrossRef] [PubMed]
3. Ipcc. Special Report—Global Warming of 1.5 °C. Intergovernmental Panel on Climate Chnage. Available online: https://www.ipcc.ch/sr15/ (accessed on 15 May 2023).
4. Haines, A.; Patz, J.A. The Health Effects of Climate Change. *Am. J. Nurs.* **2022**, *122*, 14. [CrossRef] [PubMed]
5. Eysenbach, G. Infodemiology and Infoveillance: Framework for an Emerging Set of Public Health Informatics Methods to Analyze Search, Communication and Publication Behavior on the Internet. *J. Med. Internet Res.* **2009**, *11*, e1157. [CrossRef] [PubMed]

6. Dixon, S. Twitter: Number of Worldwide Users 2019–2024. Statista. 2022. Available online: https://www.statista.com/statistics/303681/twitter-users-worldwide/ (accessed on 12 May 2023).

7. Jordan, S.E.; Hovet, S.E.; Chun-Hai Fung, I.; Liang, H.; Fu, K.W.; Tse, Z.T.H. Using Twitter for Public Health Surveillance from Monitoring and Prediction to Public Response. *Data* **2019**, *4*, 6. [CrossRef]

8. Hartley, D.M.; Nelson, N.P.; Arthur, R.R.; Barboza, P.; Collier, N.; Lightfoot, N.; Linge, J.P.; van der Goot, E.; Mawudeku, A.; Madoff, L.C.; et al. An overview of Internet biosurveillance. *Clin. Microbiol. Infect.* **2013**, *19*, 1006–1013. [CrossRef] [PubMed]

9. Zeraatkar, K.; Ahmadi, M. Trends of infodemiology studies: A scoping review. *Health Inf. Libr. J.* **2018**, *35*, 91–120. Available online: https://onlinelibrary.wiley.com/doi/full/10.1111/hir.12216 (accessed on 11 May 2023). [CrossRef] [PubMed]

10. O'Shea, J. Digital disease detection: A systematic review of event-based internet biosurveillance systems. *Int. J. Med. Inform.* **2017**, *101*, 15–22. [CrossRef] [PubMed]

11. Mavragani, A. Infodemiology and Infoveillance: Scoping Review. *J. Med. Internet Res.* **2020**, *22*, e16206. [CrossRef] [PubMed]

12. Loureiro, M.L.; Alló, M. How has the COVID-19 pandemic affected the climate change debate on Twitter? *Environ. Sci. Policy* **2021**, *124*, 451–460. [CrossRef]

13. Smirnov, O.; Hsieh, P.H. COVID-19, climate change, and the finite pool of worry in 2019 to 2021 Twitter discussions. *Psychol. Cogn. Sci.* **2022**, *119*, e2210988119. [CrossRef]

14. Reyes-Menendez, A.; Sauara, J.R.; Alvarez-Alonso, C. Understanding #WorldEnvironmentDay User Opinions in Twitter: A Topic-Based Sentiment Analysis Approach. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2537.