Trust and Trolling of Weather Forecasters: Analysis of Twitter Conversations related to climate change during UK Heatwaves in 2020 and 2022

Asfee Alam
Department of Computer Science
University of Exeter
Exeter, UK
aa1308@exeter.ac.uk

Michelle Spruce
Department of Computer Science
University of Exeter
Exeter, UK
m.d.spruce@exeter.ac.uk

***Abstract*—This project delves into understanding the attitudes and behaviours of Twitter users towards climate change and weather forecasters during UK heatwaves in 2020 and 2022. The objective is to identify abusive behaviour patterns directed against weather forecasts in order to better understand online discourse dynamics. The research analyses tweet content and attitudes to identify abusive language trends using techniques including Latent Dirichlet Allocation (LDA), VADER Sentiment Analysis, and Profanity-Check. The study identifies dominant topics in climate change conversations during heatwaves by studying major subjects through LDA. Meanwhile, VADER provides insights into sentiments, with a focus on negative attitudes. The use of Profanity-Check detects abusive language, highlighting instances where language usage becomes disrespectful and harmful. The study also looks for changes and patterns in abusive behaviour over the course of the two years. The research has significant implications because the results will reveal Twitter users' attitudes and behaviours towards climate change and weather predictions. These insights will help develop methods to reduce online abuse in the context of the climate change debate. In short, this study offers insightful information about how Twitter users behave and think during heatwaves. The use of advanced methodologies provides a comprehensive perspective of online conversations, ultimately assisting in the development of strategies to combat abusive language. By addressing this crucial issue, the project contributes to a more constructive and meaningful public dialogue on climate change and its impact.**

***Index Terms*—Heatwaves, LDA Topic Modelling, VADER Sentiment Analysis, Profanity-check Abuse Detection.**

I. Introduction

Climate change has emerged as one of the most pressing global challenges in recent times, with rising temperatures leading to increasingly frequent and intense heatwaves. Researchers and policymakers are both very interested in this topic since these prolonged hot weather periods have significant effects on human health, the environment, and the economy [1]. Recent heatwaves in the UK caused intense debate and raised serious questions about any connections between climate change and the occurrence [2].

Addressing climate change requires not only sound scientific evidence but also effective public awareness and understanding. Twitter has emerged as an essential platform for the digital age of information dissemination regarding heatwaves and climate change. However, because of their ability to spread knowledge, these platforms have also given rise to online abuse and trolling. Of particular concern is the disturbing increase in abusive behaviour directed toward weather forecasters who report on heatwaves and climate change on Twitter. These forecasters face unprecedented levels of trolling from individuals denying the existence of climate change or challenging the connection between heatwaves and climate change. Such abusive behaviour not only impacts the mental well-being of forecasters but also undermines public trust in the validity of scientific information related to climate change [3].

Considering this alarming trend, a report by the BBC shed light on the significant rise in trolling faced by weather forecasters on Twitter, especially during the UK heatwave in 2022 [4]. This highlights the urgent need for research that delves into the nature and extent of such trolling, as well as the underlying factors that contribute to its propagation. Understanding this topic is essential for strengthening climate change communication and developing an informed and receptive audience. It also raises fundamental questions about the role of social media in shaping public perceptions of climate change and the responsibility of weather forecasters in effectively communicating climate-related information.

In response to these challenges, this study proposal aims to investigate tweets sent to certain users during the UK heatwaves of 2020 and 2022. The current study uses a multi-dimensional strategy to analyse tweets about climate change and weather forecasts during heatwaves by combining methods like Latent Dirichlet Allocation (LDA), VADER Sentiment Analysis, and Profanity-Check. The use of LDA, a topic modelling technique, will reveal common topics in these conversations. With a focus on negative sentiments, VADER Sentiment Analysis will evaluate how people feel about climate change and weather forecasters in tweets. Profanity-Check will identify abusive language trends, offering insights into the extent of online abuse directed at weather forecasters. By combining these techniques, this study aims to offer a thorough understanding of the interactions, attitudes, and behaviours of Twitter users within this field.

This research offers important insights into reducing the negative effects of online abuse on climate change communication in addition to improving our understanding of how people behave online during climate change conversations. This study intends to improve the discussion on climate change, promote more trust in scientific information, and aid weather forecasters in their crucial position as climate communicators by analysing the dynamics of social media interactions and its impact on public perceptions. In the end, this study aims to close the information gap between scientists and the general people, opening the door to a more educated and civil online discussion of climate change and its effects. Through a deeper understanding of online behaviour during heatwaves, we aspire to promote a collective commitment to tackling climate change and its challenges head-on.

II. BACKGROUND

Due to the diverse ramifications of rising temperatures, climate change has become a top global concern that is receiving unprecedented levels of attention and urgency. Among these effects, the intensification of heatwaves stands out because of their increasing frequency and severity. These prolonged periods of extreme heat present multifaceted challenges to human health, ecosystems, and economies worldwide. Therefore, in the context of climate change debate, policymakers and researchers have increased their efforts to comprehend and minimise the effects of heatwaves [1].

The increase in heatwaves has sparked concerns regarding possible correlations between these unusual weather events and the broader context of climate change in the United Kingdom. But it's still difficult to explain to the general people the intricate scientific complexities of climate change [5]. Translating these ideas into understandable and relatable words is extremely difficult due to the abstract nature of climate change and the intricate connections between increasing temperatures and particular meteorological occurrences like heatwaves [6]. This communication challenge highlights the need for efficient methods that close the knowledge gap between scientists and the public.

Social media platforms like Twitter have become crucial tools for distributing information and engaging the public in meaningful dialogue in the context of climate change communication [7]. Platforms like Twitter provide a vibrant environment for exchanging ideas, discussing research, and raising awareness of climate-related issues [8]. This digital environment is not without dangers, though. The democratisation of information has made it easier for false information to spread, and the discussion on climate change has evolved into a space for polarized opinions [9]. The "climate change denial" phenomena has acquired popularity in online debates, complicating efforts to encourage educated and productive conversations [10].

In this complex web, weather forecasters have taken on a crucial role in explaining the effects of climate change to the public. They serve as a bridge between research and reality, providing perceptions into how climate change turns into observable weather patterns. However, they have faced opposition in their endeavours. Weather forecasters have noticed an increase in online abuse and trolling, particularly when discussing the connection between heatwaves and climate change on Twitter [4]. This worrying trend has implications for forecasters' psychological health as well as how the public views climate science and its communicators [11].

This rise in online abuse towards forecasters highlights the challenges of using social media for scientific discussions. It underlines the wider issue of online harassment hampering meaningful discourse on essential topics like climate change. Recent studies emphasize the need to understand and address this kind of abuse [12]. Factors like online anonymity and community dynamics play a role in promoting such behaviour. Also, the language people use online can give hints about abusive content [13]. Amidst these challenges, tools like the VADER sentiment analysis have proved useful. VADER is designed to analyse the sentiment of short texts, typical of tweets. It assesses both the type (positive, negative, neutral) and strength of sentiment, giving a detailed view of public opinion [14]. This approach has been effectively used before to understand public feelings about climate change on Twitter [15].

Another useful tool is LDA topic modelling, which identifies key themes in large sets of texts. By examining which words appear together frequently, LDA can find hidden topics and group them in a way that makes sense [16]. When applied to Twitter data, researchers can uncover the main themes of discussions around climate change, shedding light on public concerns [17].

In addition to sentiment analysis and topic modelling, detecting abuse is critical, especially given the challenges forecasters face. The "profanity-check" tool, a machine learning algorithm, is designed to spot abusive content in social media data. It's trained on datasets that have both abusive and non-abusive content, learning to recognize signs of harmful information [18]. This makes it particularly adept at flagging tweets with abusive language or offensive comments.

Combining sentiment analysis, topic modelling, and abuse detection algorithms to analyse social media data relevant to climate change and weather forecasting is feasible, according to the available research. These methodologies bear the potential to augment communication strategies, unravel public attitudes, and unveil offensive content directed at climate change communicators and forecasters. A more informed and respectful exchange of ideas can be fostered by utilising the synergy between these techniques.

III. AIMS & OBJECTIVES

The primary aim of this study is to conduct an in-depth investigation of Twitter conversations related to climate change during heatwaves and develop a comprehensive framework to gain a deeper understanding of the attitudes and behaviours of Twitter users towards tweets posted by weather forecasters on climate change during UK heatwaves. To achieve this overarching aim, the following specific objectives will be pursued:

1) *Data Collection and Pre-processing*: Collect a dataset of Twitter posts related to climate change and heatwaves, ensuring its relevance and representativeness. Implement effective pre-processing techniques to ensure data quality and suitability for analysis.
2) *Sentiment Analysis*: Use sentiment analysis on the collected tweets to determine the sentiment of tweets. Analyse the emotional tone and intensity of sentiments expressed in the posts.
3) *Topic Modelling using Latent Dirichlet Allocation (LDA)*: Use LDA to identify the major themes and topics that appear in Twitter conversations during heatwaves, especially those pertaining to climate change and weather forecasters.
4) *Abuse Detection using Profanity-check*: Implement the profanity-check abuse detection model, to automatically identify instances of abusive language directed towards weather forecasters or climate change communicators within the Twitter dataset.
5) *Correlation Analysis*: Analyse the collected data to look for any connections or correlations between sentiment expressions, recognised discussion themes, and abusive tweet occurrences.

By accomplishing these specific objectives, this research will contribute valuable insights into the complex nature of online climate change communication and public attitudes toward weather forecasters and climate change communicators. Ultimately, this study aims to promote a deeper understanding of climate change challenges and pave the way for improved climate change communication on social media.

IV. EXPERIMENT & DESIGN

The overall workflow of the methodologies and experimental strategy comprises data collection, pre-processing, topic modelling, sentiment analysis and abuse detection. Steps include:

- *Data Collection*: The data collection phase aimed to gather relevant Twitter conversations related to climate change during UK heatwaves in 2020 and 2022, with a specific focus on conversations involving weather forecasters. There were several steps in the procedure to make sure the right information was gathered for later analysis.
  1) *Hashtag Selection*: A list of keywords and hashtags was constructed to observe conversations about climate change during heatwaves. The hashtags "#heatwave," "#heat," "#hot," and "#summer" were chosen to focus on conversations about heatwaves and their effects.
  2) *Twitter Handles*: Additionally, the data collection method included the Twitter accounts of well-known weather forecasters, including @MetMattTaylor and @metoffice. The Met Office's Twitter account and meteorologist Matt Taylor's Twitter account were chosen due to their substantial contributions to UK weather forecasting and climate communication. While Matt Taylor, a reputable meteorologist, covers climate issues on his account, the Met Office is a reliable source for weather and climate information. This made sure that the dataset included the conversations these forecasters had and their observations.
  3) *Timeframe Selection*: For the years 2020 and 2022, the data collection timeline was established to include the months of July and August, which correspond to the summer season when heatwaves occurred. This broad span was then narrowed to a 15-day period for both years from July 15 to July 31.
  4) *Twitter API*: Relevant tweets based on the chosen hashtags and Twitter accounts within the predetermined time period were gathered using the Twitter API. The API made it easier to retrieve metadata such as timestamps, user information, and tweet text.

  The choice of keywords, hashtags, and Twitter handles was motivated by the objective to capture a comprehensive dataset of climate change-related conversations during heatwaves. The inclusion of weather forecasters' handles aimed to specifically target discussions involving experts in the field. The temporal focus on the summer months and the subsequent 15-day period ensured that data was gathered during periods when heatwave events were likely to occur.
- *Data Pre-processing*: Data pre-processing was essential in getting the gathered Twitter dataset ready for further analysis. The pre-processing pipeline included a number of procedures designed to standardise the data, reduce noise, and improve the text quality for efficient analysis.
  1) *Retweet Removal*: Retweets, which are duplicate posts of existing tweets, were removed from the dataset. This step ensured that only unique content was retained for analysis, eliminating redundancy.
  2) *Text Standardization*: To guarantee consistency across the collection, textual data was standardised. All text was converted to lowercase to remove variations in case. This step was crucial for subsequent text analysis, as it treated identical words in different cases as the same.
  3) *3. Removal of URLs, Links, Punctuations, Emojis, Hashtags, and Mentions*: To reduce clutter and draw attention to the tweets' actual content, URLs, links, and special characters were removed from the text. Emojis, hashtags, and mentions were also removed to ensure that only text data was subjected to subsequent analyses.
  4) *Stopword Removal*: The tweets were cleaned up of stopwords, which are often used words that don't significantly add to the meaning of a text. This

process was designed to reduce noise and improve the relevance of the text.

5) *Lemmatization*: Lemmatization, a text normalisation procedure that reduces words to their root or base forms, was applied to the tweets. This step ensured that different forms of the same word were treated as identical, thus enhancing the accuracy of subsequent analyses.
6) *Removal of Short words*: The dataset was cleared of words with fewer than two characters. This stage was designed to remove extremely brief and maybe irrelevant terms that could have an impact on the quality of the analysis.

The data pre-processing methods were chosen to improve data quality and consistency while eliminating unnecessary elements. Removing retweets, URLs, and special characters simplified the analysis. Lemmatization and stopword removal enhanced word relevance for sentiment analysis, topic modelling, and abusive language detection. These steps resulted in a refined and consistent dataset for various text analyses.

- *VADER Sentiment Analysis*: Sentiment analysis is crucial to understanding the general attitudes and behaviours of Twitter users towards climate change and weather forecasts while examining Twitter conversations about climate change during heatwaves. The Valence Aware Dictionary and sentiment Reasoner (VADER) has been chosen due to its suitability for analyzing short and informal text data commonly found on social media platforms like Twitter.
According to the lexicon-based sentiment analysis theory, VADER evaluates the polarity and intensity of sentiment represented in a given text by using a pre-built lexicon of terms and their associated sentiment scores [14]. The dictionary is made to handle slang, emoticons, and other informal terms frequently found in social media posts [20]. This strategy offers a rapid and effective way to identify sentiment trends and fits well with the nature of Twitter conversations.

1) *Lexicon-based Sentiment Analysis*: VADER performs sentiment analysis by assigning a sentiment score to each word in a text based on its appearance in the lexicon. The sentiment scores include positive, negative, neutral, and compound scores. The positive, negative, and neutral scores show the percentages of each sentiment category, and the compound score shows the text's total sentiment intensity.
2) *VADER Analysis*: VADER sentiment analysis was applied to the cleaned and pre-processed text data. The VADER SentimentIntensityAnalyzer was used to determine sentiment scores, including the compound score.
3) *Sentiment Labelling*: Compound sentiment scores were categorized into labels: scores above 0.05 as "Positive," below -0.05 as "Negative," and those in between as "Neutral."

Using VADER helps us figure out how people feel about climate change during heatwaves on Twitter. By looking at the sentiment scores and labels, we can see what most people think and how they behave. Also, it works quickly and gives us a number that shows how strong the emotions are. This helps us understand the challenges and opportunities for talking about climate change effectively.

- *LDA Topic Modelling*: Latent Dirichlet Allocation (LDA) topic modelling is used to try and find latent themes and patterns in the Twitter conversations about climate change during heatwaves. LDA is a probabilistic generative model that discovers underlying topics through analysing the frequency of words across documents [16].

1) *Topic Range and Parameters*: The LDA topic modelling process commences with selecting an appropriate range of topics. The number of topics to be considered ranges from a minimum value to a maximum value in steps determined by the step size. Additionally, the alpha and beta hyperparameters are set, with 'auto' and 'symmetric' as the chosen options, representing the Dirichlet prior parameters for the per-document topic distribution and the per-topic word distribution, respectively.
2) *Coherence Calculation*: To determine the optimal number of topics and hyperparameters, the coherence score is computed for each combination. Coherence measures the interpretability and semantic relevance of topics. For each set of parameters, the coherence value is calculated based on the 'c_v' coherence metric. The coherence values for different topic numbers, alpha, and beta values are recorded to assist in finding the best configuration [21].
3) *Optimal Model*: The best model is selected using the combination of themes and hyperparameters with the highest coherence score. The most important and relevant topics discussed on Twitter will be revealed by this approach.
4) *Perplexity and Coherence Score Calculation*: The performance of the chosen optimal model is evaluated further. To evaluate how well the model fits the data, the perplexity metric is calculated. A lower perplexity value indicates a better model fit. The coherence score is computed using the Coherence Model class with the 'c_v' coherence metric. A higher coherence score suggests that the topics are more semantically coherent and interpretable [22].

We choose LDA because it is effective at identifying different topics in active conversations regarding climate change on Twitter [17]. It helps us in identifying significant themes that might be connected to trolling. In this way, when people discuss climate change and weather forecasters during heatwaves, it is evident what they are referring to.

- *Profanity-check-based Abuse Detection*: Profanity-Check is a Python library is used to identify and categorize text that contains offensive, vulgar, or inappropriate language. It utilizes a trained model that has learned from a diverse range of text data containing offensive language. The model assigns a probability score to each text input, indicating the likelihood that the text contains profane or abusive language [18]. This score can help differentiate between mild instances of potentially offensive language and more explicit and inappropriate content.
  1) *Profanity Detection*: The Profanity-Check-based approach involves implementing two primary functions: 'detect_profanity' and 'profanity_score'. These functions are applied to detect the presence and intensity of offensive language within the text.
  2) *detect_profanity Function*: The 'detect_profanity' function uses the Profanity-Check library's 'predict' function. This function takes an array of text entries and assigns a binary label: 1 for offensive language and 0 for non-offensive language.
  3) *profanity_score Function*: In addition to binary classification, the 'profanity_score' function utilizes the 'predict_prob' function from Profanity-Check. This function assigns a probability score to each text entry, indicating the likelihood of the text being offensive.

  We picked the Profanity-Check library because it's really good at finding offensive words in the Twitter data. Since online abuse is a big issue, this tool helps us quickly find and measure how much offensive language is there. It's a straightforward way to figure out if text is offensive or not. This matches our goal of checking how much abusive behavior happens when people talk about climate change on Twitter.

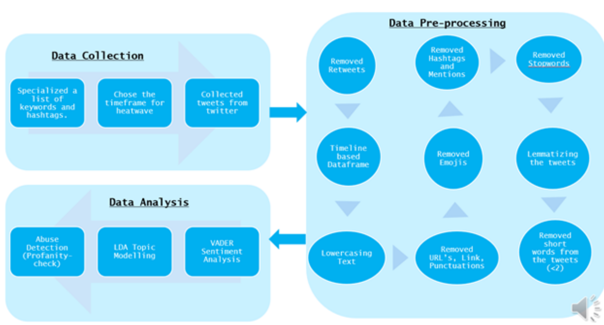


Fig. 1. Flowchart of Experiment and Method

V. RESULTS

- *Tweet Volume Analysis*: In 2020, the graph in Fig 2 depicting the daily tweet volume highlighted July 31st as the day with the highest number of tweets, exceeding 9,000 tweets. Remarkably, this day also coincided with the hottest day of the year in the UK. Similarly, in 2022, Fig 3 shows July 19th emerged as the day with the highest tweet volume, with approximately 40,000 tweets. Notably, this day corresponded to the hottest day in the UK for that year. The analysis of tweet volume on this day suggested a strong correlation between extreme temperatures and increased online activity as individuals turned to social media to discuss the heatwave and its implications.

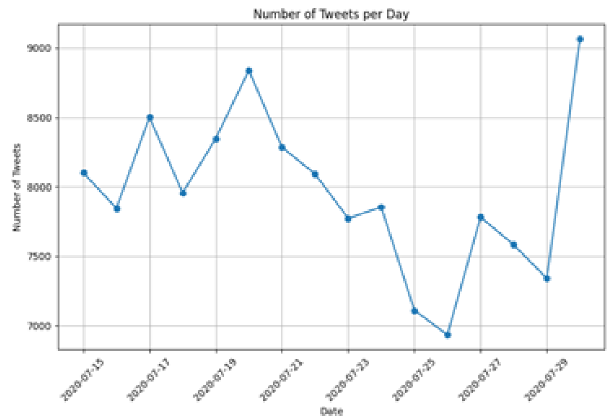


Fig. 2. Line plot showing number of tweets per day for 2020

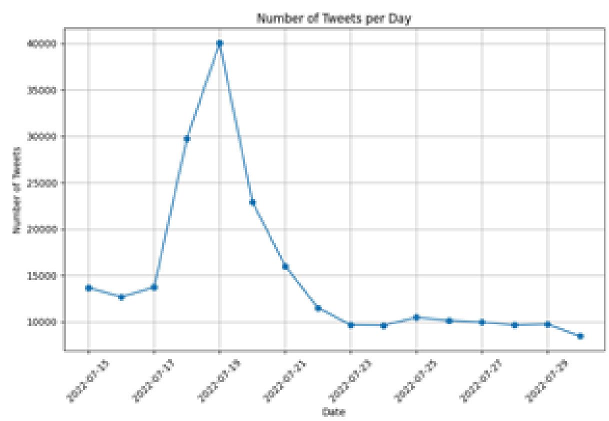


Fig. 3. Line plot showing number of tweets per day for 2022

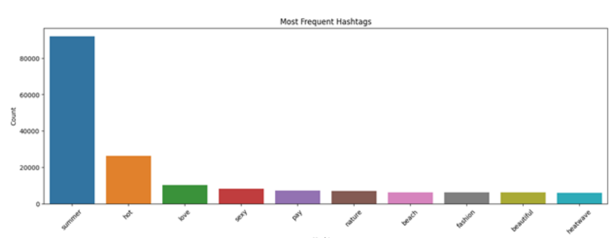


Fig. 4. Bar graph showing most frequent hashtags in 2020

- *Prevalent Hashtags*: By examining the most used hashtags graph from Fig 4, we found that in 2020, hashtags such as #summer were used in over 80,000 tweets, and #hot was present in over 20,000 tweets. This correlation

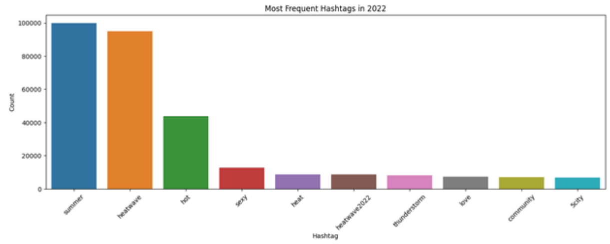


Fig. 5. Bar graph showing most frequent hashtags in 2022

of hashtags with the hottest day of the year (July 31st) showed that people tend to talk about weather when it's extremely hot.

- Fig 5 shows the popular hashtags in 2022 that offered a different viewpoint. Both the hashtags #summer and #heatwave were used in roughly 100,000 tweets, respectively. From the above graphs, we suggested that 2022 might have experienced even higher temperatures than 2020, a hypothesis confirmed by news reports from the UK.
- The LDA model highlighted key topics during heatwaves in 2020 and 2022, with an interactive pvLDAvis notebook aiding interpretation shown in Fig 6. Common topics emerged, including "heat" and "temperature" reflecting extreme conditions. "Heatwave" and "summer" were prevalent, discussing ongoing heatwaves within the season. Expressive terms like "sexy," "cool," and "hot" indicated emotional responses. The visualization revealed diverse themes beyond "heatwave," capturing emotions, weather terms, and environmental discourse, emphasizing heatwave's impact on Twitter discussions.

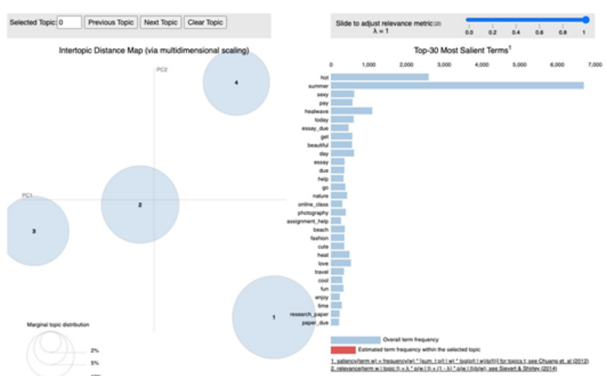
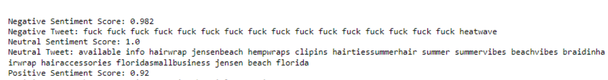


Fig. 6. LDA model output showing topics extracted from the tweets.



```
Negative Sentiment Score: 0.982
Negative Tweet: fuck fuck fuck fuck fuck fuck fuck fuck fuck fuck fuck fuck fuck fuck fuck fuck heatwave
Neutral Sentiment Score: 1.0
Neutral Tweet: available info hairwrap jensenbeach hempwraps clipins hairtiessummerhair summer summervibes beachvibes braidinha
irwrap hairaccessories floridasmallbusiness jensen beach florida
Positive Sentiment Score: 0.92
Positive Tweet: happy partying partying hotgirls partying
```

Fig. 7. Positive, Negative and Neutral Tweets after performing VADER Sentiment Analysis

- In 2022, the sentiment distribution indicated that around 51.8% of tweets were classified as positive, reflecting optimistic or favorable language. About 14.6% were categorized as negative, revealing pessimistic or critical language, and 33.6% were labelled as neutral. In 2020, the sentiment distribution was slightly different, with roughly 59% of tweets categorized as positive, 9.6% as negative, and 31.3% as neutral.

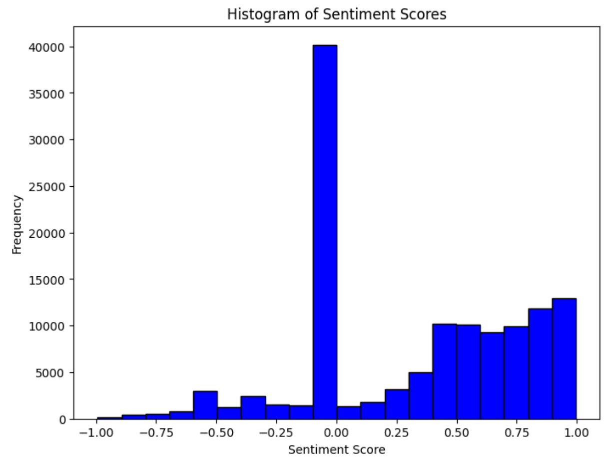


Fig. 8. Sentiment Score Distribution for year 2020

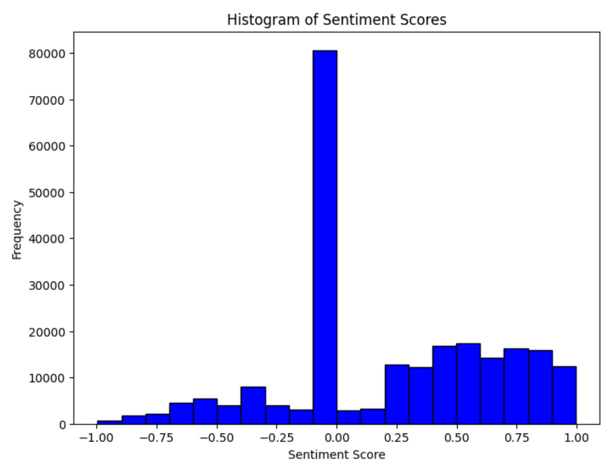


Fig. 9. Sentiment Score Distribution for year 2022

- Comparing the two years, the sentiment distribution analysis reveals that sentiment was more negative in 2022 compared to 2020, showing a 5% difference (Fig 8 and Fig 9). This observation aligns with the news and the perception that the 2022 heatwave was more severe than that of 2020. The higher proportion of negative tweets raises concerns about the presence of abusive and critical language directed towards weather forecasters and climate communicators.
- *Abuse Detection*: With the help of Fig 10 and Fig 11, the analysis of profanity scores associated with tweets directed at the Met Office and MetMattTaylor Twitter handles in both the years 2020 and 2022 reveals a concerning pattern of abusive language and negativity.

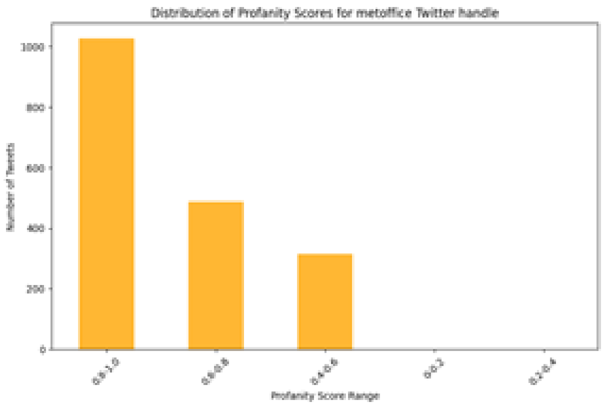


Fig. 10. Distribution of Profanity Scores for MetOffice

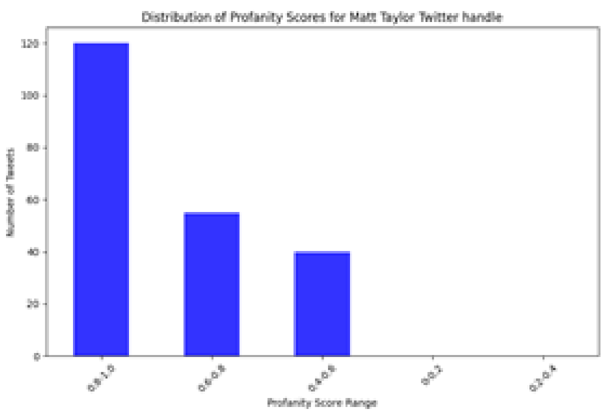


Fig. 11. Distribution of Profanity Scores for Mat Taylor

The distribution of profanity scores clearly indicates a high range of profanity scores for negative tweets directed at the Met Office and MetMattTaylor Twitter handles in both 2020 and 2022. This suggests that a significant proportion of the discourse aimed at these professionals involves the use of profane and offensive language. This analysis is confirmed when a wordcloud was formed. Upon a detailed analysis of the generated wordclouds, a consistent pattern emerges. The wordclouds as shown in Fig 12 and Fig 13, vividly depict the prevalence of abusive and disrespectful language used by Twitter users towards weather forecasters. The presence of terms like "fucking," "idiot," "shit,", "crap", "stupid" etc, underscores the hostile and derogatory tone employed in the tweets. These terms not only reflect strong emotional reactions from users but also highlight the frequent use of coarse language to express dissatisfaction or disagreement with the forecasters' insights.

- In the analysis involving the hashtags #heat, #heatwave, and #summer, the outcomes were consistent with words like "fuck", "ass", "hell", "die" with those observed for the MetMattTaylor and Met Office Twitter handles. The similarities in the results across many entities serve as additional evidence of the prevalence of abusive language

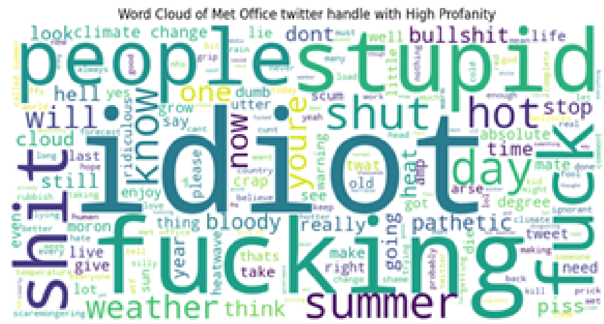


Fig. 12. Wordcloud of Met Office

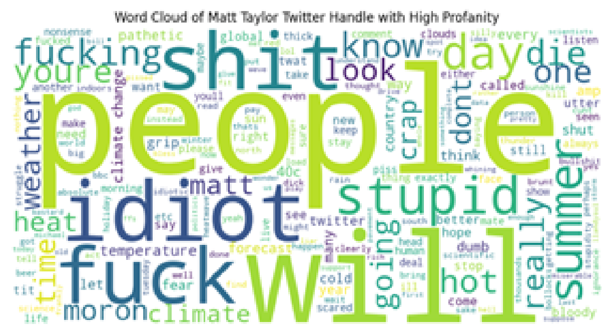


Fig. 13. Wordcloud of Mat Taylor

and unfavourable attitudes in conversations about climate change, heatwaves, and weather forecasts.

VI. Discussion

We discovered some intriguing trends when we examined tweets regarding climate change during the UK heatwaves in 2020 and 2022. During severe weather, such as heatwaves, more people tweet. In addition, our study demonstrated that there can be strong opinions on both sides of the climate change debate in online discussions [19].

According to our research, users tweeted more negative tweets in 2022 than they did in 2020. Other experts have also discussed this rise in insults and rude remarks directed at weather forecasts and climate change [3]. We noticed several unpleasant tweets, which is in line with what the BBC discovered. They said that weather forecasters received a lot of online abuse, particularly during the 2022 UK heatwave [4].

Despite the depth of our investigation, the approaches we used have certain limitations. When used with Twitter data, LDA topic modelling, for instance, has limitations. As previously indicated, some tweets simply include hashtags, emoticons, and mentions. It is difficult to identify underlying subjects while reading content with such little linguistic value. Similar to this, our sentiment analysis model, VADER, has trouble detecting nuances like irony and sarcasm. Given Twitter's informal language and frequent use of irony, this limitation may result in potential inaccuracies. Additionally, even while the profanity-check algorithm is good at highlighting explicit language, it might miss other forms of online abuse. Due to its reliance on a pre-defined list, it may overlook context-specific or new offensive terms.

We can make some modifications in the near future to enhance our research. Instead of using an existing model, we might build our own to look for offensive words in tweets. The type of tweets, such as whether they are abusive, discriminatory, or mean, could be distinguished using this new technique.

A more comprehensive picture of public mood might be obtained by extending the analysis' time scope beyond the initial 15 days. We can identify patterns, track the evolution of sentiments, and determine the impact of outside events on online behaviour by analysing data over the course of a full year or over a number of years. To identify regional differences in opinions and discussions, geospatial analysis of tweets can be used. This provides insights into how geographical factors affect online climate discourse.

VII. CONCLUSION

The analysis of climate change-related tweets during the UK heatwaves of 2020 and 2022 has given a thorough overview of online climate discourse, highlighting major trends and opinions that are widely held in general. The study, which highlighted the public's sensitive reactivity to important weather occurrences like heatwave, particularly showed a correlation between extreme temperatures and increased online engagement.

The sentiment analysis, facilitated by VADER, was particularly enlightening, illuminating a shift in public sentiment from 2020 to 2022. This shift towards a more negative sentiment in 2022 emphasizes the growing intensity of climate phenomena and, possibly, increasing public concern or dissatisfaction with how these issues are being addressed or communicated.

Furthermore, the application of Latent Dirichlet Allocation (LDA) topic modeling provided a structured overview of the prevailing themes in these Twitter conversations. Beyond the predictable discussions on heat and temperature, the emergence of colloquial expressions and real-time observations offers a vivid portrayal of how climate events are experienced and articulated by the public in real-time.

Perhaps one of the most critical findings of this study was the evident rise in abusive language directed towards weather forecasters and climate communicators, as indicated by our Profanity-check results. The digital realm, while democratizing information dissemination, also presents challenges in the form of trolling and online abuse. The findings of this study eloquently demonstrate this problem, highlighting the necessity of promoting respectful and productive online conversation.

In conclusion, this research has highlighted crucial areas of concern and interest in addition to mapping the digital terrain of climate discourse during significant heatwave events. Understanding public opinion and discourse is crucial as the world struggles with multiple issues caused by climate change like heatwave. This work makes a substantial contribution to this understanding by its rigorous methodology and insights, laying the way for future climate communication efforts that are more informed and more effective.

VIII. DECLARATION

Declaration of Originality I am aware of and understand the University of Exeter's policy on plagiarism, and I certify that this assignment is entirely original with the exception of the citations, and that I have adhered to best practices for academic writing.

Declaration of Ethical Concerns There are no moral concerns raised by this work. There are no human or animal subjects involved, and no personal information about human subjects has been handled. Additionally, no acts that could compromise security or safety have been carried out.

REFERENCES

[1] Global warming of 1.5 oc, https://www.ipcc.ch/sr15/ (accessed Aug. 18, 2023).

[2] State of the UK climate 2021 - Kendon - Wiley Online Library, https://rmets.onlinelibrary.wiley.com/doi/10.1002/joc.7787 (accessed Aug. 18, 2023).

[3] The social media life of climate change ... - wiley online library, https://wires.onlinelibrary.wiley.com/doi/10.1002/wcc.569 (accessed Aug. 18, 2023).

[4] M. Thomas, "UK Heatwave: Weather Forecasters report unprecedented trolling," BBC News, https://www.bbc.co.uk/news/uk-62323048 (accessed Aug. 18, 2023).

[5] The online disinhibition effect — Cyberpsychology & Behavior, https://www.liebertpub.com/doi/10.1089/1094931041291295 (accessed Aug. 18, 2023).

[6] Consensus on consensus: A synthesis of consensus estimates ... - iopscience, https://iopscience.iop.org/article/10.1088/1748-9326/11/4/048002 (accessed Aug. 18, 2023).

[7] D. Brossard and M. C. Nisbet, "Deference to scientific authority among a low information public: Understanding U.S. opinion on Agricultural Biotechnology," OUP Academic, https://academic.oup.com/ijpor/article/19/1/24/752714 (accessed Aug. 18, 2023).

[8] C. Mora et al., "Global risk of deadly heat," Nature News, https://www.nature.com/articles/nclimate3322 (accessed Aug. 18, 2023).

[9] E. M. Cody, A. J. Reagan, L. Mitchell, P. S. Dodds, and C. M. Danforth, "Climate change sentiment on Twitter: An unsolicited public opinion poll," PLOS ONE, https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0136092 (accessed Aug. 18, 2023).

[10] "Apa PsycNet," American Psychological Association, https://psycnet.apa.org/record/2017-57700-001 (accessed Aug. 18, 2023).

[11] C. Llorente, G. Revuelta, M. Carrió, and M. Porta, "Scientists' opinions and attitudes towards citizens' understanding of science and their role in public engagement activities," PloS one, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6853295/ (accessed Aug. 18, 2023).

[12] P. Nauroth, M. Gollwitzer, J. Bender, and T. Rothmund, "Social identity threat motivates science-discrediting online comments," PLOS ONE, https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0117476 (accessed Aug. 18, 2023).

[13] The development and Psychometric Properties of liwc2015, https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageM (accessed Aug. 18, 2023).

[14] Vader: A parsimonious rule-based model for sentiment analysis of social ..., https://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf (accessed Aug. 18, 2023).

[15] Anatomy of news consumption on Facebook — PNAS, https://www.pnas.org/doi/10.1073/pnas.1617052114 (accessed Aug. 18, 2023).

[16] Latent dirichlet allocation - Journal of Machine Learning Research, https://jmlr.org/papers/volume3/blei03a/blei03a.pdf (accessed Aug. 18, 2023).

[17] H. Jelodar et al., “Latent dirichlet allocation (LDA) and topic modeling: Models, applications, a survey - multimedia tools and applications,” SpringerLink, https://link.springer.com/article/10.1007/s11042-018-6894-4 (accessed Aug. 18, 2023).

[18] T. Davidson, D. Warmsley, M. Macy, and I. Weber, “Automated hate speech detection and the problem of offensive language,” Proceedings of the International AAAI Conference on Web and Social Media, https://ojs.aaai.org/index.php/ICWSM/article/view/14955 (accessed Aug. 18, 2023).

[19] Full article: The role of (social) media in political polarization: A ..., https://www.tandfonline.com/doi/full/10.1080/23808985.2021.1976070 (accessed Aug. 18, 2023).

[20] “Simplified text processing¶,” TextBlob, https://textblob.readthedocs.io/en/dev/ (accessed Aug. 18, 2023).

[21] M. R. L. University et al., “Exploring the space of topic coherence measures: Proceedings of the Eighth ACM international conference on web search and data mining,” ACM Conferences, https://dl.acm.org/doi/10.1145/2684822.2685324 (accessed Aug. 18, 2023).

[22] Automatic evaluation of Topic Coherence - ACL Anthology, https://aclanthology.org/N10-1012.pdf (accessed Aug. 18, 2023).

[23] Author links open overlay panelApoorva Upadhyaya et al., “Towards sentiment and temporal aided stance detection of climate change tweets,” Information Processing amp; Management, https://www.sciencedirect.com/science/article/pii/S0306457323000626 (accessed Aug. 18, 2023).

[24] K. K. Zander, J. Rieskamp, M. Mirbabaie, M. Alazab, and D. Nguyen, “Responses to heat waves: What can twitter data tell us? - natural hazards,” SpringerLink, https://link.springer.com/article/10.1007/s11069-023-05824-2 (accessed Aug. 18, 2023).

[25] Inoculating the public against misinformation ... - wiley online library, https://onlinelibrary.wiley.com/doi/10.1002/gch2.201600008 (accessed Aug. 18, 2023).