Droughts and Media: when and what do the newspapers talk about the droughts in England?

Inhye Kong[1], Jan Seibert[1], Ross S. Purves[1]

[1]Department of Geography, University of Zurich, Zurich 8057, Switzerland

Correspondence to: Inhye Kong (inhye.kong@geo.uzh.ch)

Abstract. The United Kingdom is traditionally known for its wet climate, but droughts are also a recurring concern. Using newspapers as a medium of public communication, this study explores the timing and content of newspaper articles about droughts in England. We constructed a corpus of more than 800 newspaper articles related to droughts in England for the last 24 years (2000 - 2023), and analysed the temporal alignment of newspaper coverage with hydroclimatic anomalies and seasonality using negative binomial regression model. Our result showed that newspaper coverage of droughts coincides with low precipitation (SPI-3), low groundwater (SGI-1) and spring/summer seasons, but temperature (CET-12) was not a significant factor. The following topic modelling revealed the dominant narratives to be "Drought and hosepipe ban", which concerns of meteorological and hydrological water deficiency and a common measure to restrict water usage in England. Upon the comparison between two major droughts (i.e. spring 2012 vs. summer 2022), we found the summer drought of 2022 to contain summer-related topics (i.e., "Heatwave", "Temperature and hosepipe ban"), harnessing earlier finding about warm seasons to generate more attention. Overall, our findings reveal that newspaper reporting about droughts is influenced by a combination of factors, rather than precipitation alone, with a notable seasonality-based confirmation bias. While this implies a potential mismatch in how droughts are conceptualized by newspapers compared to scientists, we conceive that such bias may underrepresent early signs of drought in cold seasons, potentially undermining public support for preventive measures. Encouragingly, we observed the emergence of water management as a prominent topic during the 2022 drought, reflecting the growing maturity of newspaper narratives in addressing the managerial failures of private water suppliers and calling for appropriate measures to ensure a reliable water supply.

1 Introduction

The United Kingdom (UK) is well known for its wet climate and frequent flooding, but droughts have been commonly recurring throughout history, with a major drought every 5-10 years on average (Dayrell et al., 2022; Marsh et al., 2007; Murphy et al., 2006). Within the last two decades, The National Hydrological Monitoring Programme of the UK has released major drought reports for 2003 (Marsh, 2004), 2004-2006 (Marsh, 2007), 2010-12 (Marsh et al., 2013), and 2018-19 (Turner et al., 2021). England is more susceptible to severe droughts within the UK due to less precipitation and greater water demand from human influences (Tanguy et al., 2021; Tijdeman et al., 2018). This pattern is especially distinctive around southern

England, including the London metropolitan area (Dessai & Sims, 2010; Folland et al., 2015). According to the IPCC report (IPCC, 2023) and scientific works (Guerreiro et al., 2018; Spinoni et al., 2018), climate change is expected to heighten drought-prone conditions in Western and Central Europe, and England is no exception (Dobson et al., 2020; Kay et al., 2023).

Droughts, as an example of extreme weather events, are prolonged periods of imbalance between rainfall and evapotranspiration (Wilhite & Glantz, 1985). They can be further categorised as meteorological droughts, characterised by a lack of precipitation, and hydrological droughts, marked by reduced streamflow, reservoir levels, lake volumes, and groundwater table (Van Loon, 2015). Aside from these hydroclimatic droughts, socioeconomic droughts are defined by the impact of drought on society, including agricultural and economic losses and failures to meet water demands for people and the environment (Van Loon, 2015; Wilhite & Glantz, 1985). Compared to other natural hazards, droughts appear to receive less media and public attention (Brimicombe, 2022). This is primarily due to their being a "creeping phenomenon"—a gradual change, often developing over extended periods, in contrast to the rapid onset of events like floods and wildfires (Kim et al., 2019). In addition, droughts are the result of complex hydrological dynamics involving a lagging effect, resulting in temporal mismatches between rainfall, streamflow, and groundwater recharge (Van Loon, 2015). Deficits in rainfall (i.e., meteorological drought) do not always coincide with or lead to low streamflow or low groundwater (i.e., hydrological drought), and not all such cases result in tangible impacts for society (i.e. socioeconomic droughts).

Given the complex and gradual nature of drought dynamics, it is crucial to analyse how and when droughts are reported in the media, as these factors play a significant role in shaping public perception and response. The timing of media coverage reflects not just the occurrence of droughts but also the point at which their impacts become newsworthy (Caple, 2018). This raises important questions about whether media reporting aligns with the actual progression of drought events. Additionally, the framing and narratives constructed in media coverage deserve close examination (Vargas Molina et al., 2022), as they influence public perception and societal response, from individual behavioural changes (Antwi et al., 2022; Quesnel & Ajami, 2017) to policy initiatives (Hart et al., 2015). Thus, understanding the temporal and narrative dimensions of drought reporting is essential for enhancing public engagement, fostering informed decision-making, and driving effective societal action.

Despite the rise of new and diverse media forms, most prominently as social media, newspapers remain a vital data source for media analysis (Vargas Molina et al., 2022). Written by professional journalists and reporters, newspapers often have large circulations reaching broad and diverse segments of the population (Boykoff, 2008). This characteristic enables newspapers to represent a wide range of societal perspectives, which are often absent in social media or scientific publications. Another advantage of newspapers lies in their coverage of diverse topics and issues, reflecting the multifaceted concerns and dynamics of society (Schmidt et al., 2013). Moreover, advancements in digital transformation have greatly improved access to media archives, providing researchers with extensive historical and contemporary newspaper datasets. These developments facilitate the creation of large-scale newspaper corpora spanning significant temporal ranges, enabling robust media analyses (Dayrell et al., 2022; O'Connor et al., 2022).

Leveraging the access to extensive digital archives of news media, several studies have investigated the temporal alignment between media attention about droughts and hydroclimatic conditions. A common approach involves using the size of the news corpus as a proxy for media attention and comparing it with meteorological and hydrological indices. For example, Linés et al. (2017) used remotely sensed indices, including the Standardized Precipitation Index (SPI), to assess temporal alignments with media reports in Spain. Similarly, Noone et al. (2017) analysed an Irish newspaper corpus and correlated it with SPI-12 over the past 250 years. Dayrell et al. (2022) examined UK newspaper corpora over the last 200 years to explore the correlation between a drought index (i.e., SPI) and media coverage. O'Connor et al. (2022) studied historical newspaper articles (1900–2016) in Ireland, finding SPI-3 to be the best predictor of land-based drought reports. Although Pianta and Sisco (2020) pointed to the influence of temperature anomalies on media coverage of climate change, the role of temperature in drought-related reports remains underexplored. Furthermore, some studies have noted a seasonal bias in media reporting, with increased attention during warmer months (O'Connor et al., 2022; Parsons et al., 2019). However, the extent to which seasonality influences newspaper coverage of droughts warrants further investigation.

Regarding the content of newspapers about droughts, there has been several studies to identify dominant topics, main stakeholders, impacts, and emerging concerns. Dow (2010) utilized qualitative content analysis to analyse decades-long news coverage (1998-2007) in the Carolinas US identifying the impacts of droughts (e.g., agriculture, fires, lawns, recreation, etc.). Wei et al. (2015) also employed qualitative content analysis to explore historical news media reporting of water issues in Australia (1843-2011), revealing institutions and policy initiatives related to different types of water-related events. Osaka et al. (2020) focused on revealing difference in content between nationwide and regional newspapers about Californian droughts between 2011-2017. Painter et al. (2021) collected online news reports of the 2019 summer heatwave across France, Germany, the Netherlands, and the UK, and analysed how climate change was represented in the content. However, most studies heavily rely on a qualitative approach, which poses significant constraints in the scalable application of content analysis. In recent years, natural language processing (NLP) has been implemented to process a large set of text data and elicit key narratives, for example, Bohr (2020) to apply unsupervised text analysis with topic modelling about newspapers discussing climate change in the United States, and Sodoge et al. (2023) to apply pre-trained text classifiers to classify drought impacts from German newspapers (2000-2021).

The first quarter of the 21st century was the period when broad scientific consensus on anthropogenic climate change was reached, and also a period with more frequent, intense, and widespread drought impacts. UK government policies have clarified a need to introduce programs for climate change adaptation, including relevant programs for droughts (e.g., The Third National Adaptation Programme (NAP3) and the Fourth Strategy for Climate Adaptation Reporting). Still, there is limited evidence to understand how people perceive droughts, especially generally humid regions such as England. The research gap addressed in this paper addresses our limited understanding of the link between the volume of newspaper articles and hydroclimatic conditions. Additionally, there is a need for investigation into newspaper narratives surrounding the coverage of droughts, utilizing scalable computational approaches to enhance our understanding and linking these to drought indices.

In this study, we analysed newspaper articles reporting droughts in England, aiming to 1) quantify the influence of hydroclimatic conditions, including not only precipitation but also temperature and seasonality, on the volume of newspaper articles, and 2) examine the content of such newspaper narratives. We hypothesised that newspaper reporting is driven by the conditions of meteorological droughts (i.e., higher temperatures and lower precipitation) and hydrological droughts (i.e. lower groundwater level), alongside the possibility of confirmation bias in seasonality – that is to say, an expectation that drought occurs in warm periods. In our content analysis, we expect to find topics related to the different types of droughts (i.e. meteorological, hydrological, and socioeconomic droughts), and to explore distinctions in topics depending on different background conditions. Therefore, our research questions are summarised as follows:

RQ1. How and to what extent do hydroclimatic conditions and seasonality relate to newspaper reporting in England?

RQ2. What are the dominant narratives in newspaper reporting of droughts, and how do these narratives vary across major drought events?

2 Methodology

To answer the research questions, we build a newspaper corpus about drought reports in England and analyse the corpora a) to compare the size of media coverage to drought-related anomaly indices and seasonality (RQ1), and to explore the media content using natural language processing (RQ2). All the code used in our analysis is available as R-Markdown (see Supplementary Material 1).

2.1 Data collection

2.1.1 Corpus building

The first step in any text-related project is creating a systematic corpus of relevant documents. Among many approaches to building a corpus, we chose a keyword-based approach to collect newspaper articles about drought-related reports in England. We chose only England, as opposed to the whole UK, since England's drier climate and higher population density leave it more susceptible to droughts and their impacts (Folland et al., 2015). Besides, current water management in the UK is delegated to the devolved governments of Northern Ireland, Wales and Scotland (Robins et al., 2017), and we aimed to minimise the influence of such regulatory differences.

To collect newspaper articles, we used the media hub platform Nexis Uni (www.advance.lexis.com). Initially, we retrieved articles that contained the term 'drought' in their headline field and were published between January 2000 and August 2023 (i.e. since August marks the end of one hydrological cycle). However, since the term drought may be used metaphorically (e.g., 'goal drought' in sports), we further refined the search results to retain only those containing the terms 'dry' and 'England' so that we can maximise both precision (i.e., the proportion of articles retrieved whose topic was water shortage in some form) and recall (i.e., the proportion of all articles describing water shortage in some form that we retrieved).

In making a query, we also defined the source of newspapers. We considered both broadsheet and tabloid newspapers to include a wide-spanning political spectrum (i.e., right- and left-wing stances) as well as readership (i.e., both widely-read tabloid journalism and broadsheets, referred to as the quality press) (Boykoff, 2008; Norton & Hulme, 2019). Our corpus was based on five broadsheet sources (i.e. The Guardian, The Independent, The Times, Financial Times, and The Daily Telegraph) and three tabloid sources (i.e. The Sun, Daily Mirror, Daily Mail). When the source of newspapers was not identified, the articles were removed. In the end, we obtained 1,361 drought-related articles.

Having built the corpus, we carried out data-cleaning steps to increase precision (i.e., the proportion of relevant articles from the queried results). We first removed duplicate articles with the same title. Then, we screened the geographic and thematic metadata provided by Nexis Uni and removed those where primary entries when irrelevant (e.g. geographic metadata referring to Wales, Scotland, Northern Ireland, and abroad; subject metadata referring to sport). Nonetheless, we noticed several cases of similar articles remained in the corpus, where, for example, online articles were very slightly edited versions of printed articles or vice-versa, but they were kept in the corpus. After these steps, 836 articles (61%) remained for the analysis, and we grouped them into monthly counts. The process was performed in R Package, version 4.3.2 (R Core Team, 2023)

2.1.2 Hydroclimatic and seasonality metrics

For meteorological and hydrological variables, we considered temperature, precipitation, streamflow, groundwater, and seasonality. For the temperature, we collected Mean Central England Temperature (ºC) (hereafter CET; Parker et al. 1992, https://www.metoffice.gov.uk/hadobs/hadcet/data/meantemp_monthly_totals.txt) from the web archive of the UK Meteorological Office (MET), the United Kingdom's national weather service. To calculate temperature anomalies, we first calculated long-term averages of CET for 1991-2020 (30 years), which served as baselines, and then computed anomalies, i.e., deviations of CET from the long-term baseline. The temperature anomalies were then computed for moving windows of 3, 6, and 12 months (e.g., for an event in January, the window of 3 months includes January and the preceding December and November), to consider the lagging effect in physical processes.

For the precipitation, streamflow, and groundwater, we collected Standardised Precipitation Index (SPI), Standardised Streamflow Index (SSI) and Standardised Groundwater Index (SGI) from the UK Water Resources Portal (https://eip.ceh.ac.uk/hydrology/water-resources/). The SPI is a popular drought index to estimate the probability of precipitation for a given temporal window (e.g., Dayrell et al., 2022; O'Connor et al., 2022; Parsons et al., 2019). SPI and SSI data were provided with the moving windows of 3, 6, 12 months (i.e. SPI-3, SPI-6, SPI-12, SSI-3, SSI-6, and SSI-12). Since these data were at station level, we aggregated them to make an England-level index. Here, we found a strong correlation between SPI and SSI (ranging from 0.82 to 0.90, Supplementary Material 2), thus, we only considered SPI in the analysis, to minimise the effects of multicollinearity. The SGI data was also at station-level, but only with 1-month window, so we prepared an aggregated SGI-1 of England.

Finally, we defined seasons as autumn (SON, September-October-November), winter (DJF, December-January-February), spring (MAM, March-April-May), and summer (JJA, June-July-August), considering September (the onset of autumn) to be the beginning of the hydrological cycle in England.

2.2 Newspaper reporting in relation to hydroclimatic anomalies and seasonality

2.2.1 Statistical regression analysis

To measure the influence of temperature, precipitation, groundwater, and seasonality on the size of newspaper reporting, we used a statistical regression analysis. We selected negative binomial regression, which is optimal when the dependent variable is over-dispersed, i.e. the data does not follow normal distribution and are highly skewed, due to e.g. many zero counts (Green, 2021). Negative binomial regression also allows inclusion of categorical variables into the model, which is necessary for the seasonality. In the regression, we added an additional variable to explain the temporal lagging effect, i.e., the article counts of the preceding month. This lagged variable was meant to capture autoregressive patterns, i.e., when the media attention rises, it can generate a spill-over effect in the following months. As a result, we constructed a model of monthly counts of newspaper articles as a function of CET, SPI, SGI, seasonality, and the lagging effect, where CET, SPI, and SGI refer to the anomalies for temperature, precipitation, and groundwater, and seasonality refers to four classes of seasons where autumn is a baseline. In testing the model performance, we arranged different combinations: there are four options for CET, i.e., 3, 6, 12 months of anomalies, or none; for SPI, likewise, there are four options to choose: SPI of 3, 6, 12 months, or none; for SGI, there were two options: SGI of 1month or none; for seasonality, four categorical values, i.e. SON, DJF, MAM, JJA, were considered altogether or not. The lagging effect was included in all combinations. As such, we tested all combinations except for an empty one (i.e., 4*4*4*2-1 = 63) to find the best-fitting model. The best-fitting model was chosen based on the lowest AIC (Akaike Information Criterion).

2.2.2 Time-series visual plots

We created time-series plots to show how the selective variables from the best-fit model interplay with the monthly count of newspaper articles. The plot is composed of a bar plot to show the count of monthly newspaper articles, along with line plots of 'the most fitting' variables of CET, SPI and SGI from the regression analysis. The colour-blind friendly palette for this visualisation was take from Wong (2011). Additionally, we coloured the background yellow for the months where the conditions are favourable to droughts, i.e. CET > 0, SPI < 0, and SGI < 0, which helped us to find the overlaps between the volume of newspaper reporting and the favourable hydroclimatic conditions.

The second temporal plot delivers more visually oriented information with a heatmap showing the monthly article counts and selected variables out of CET, SPI, SGI, along with text labels and colour gradients. To reduce skewed colour

gradients due to outliers, we applied a winsorisation normalization method to replace the colours for extreme outliers with the colours of statistical endpoints of 1% and 99%.

2.3 Newspaper narratives with topic modelling

Topic modelling is an unsupervised computational method to identify and group topics through natural language processing. The algorithm groups words found in documents in a corpus according to their probability of belonging to a topic (Blei et al., 2010). Topics are reported as groups of words associated with probabilities, and researchers can label them in a post-hoc process. Topic modelling is a common unsupervised approach to exploring the content in a corpus through so-called distant reading, without recourse to reading individual articles. It is often effective in giving an overview of different themes found in the text in case of natural hazards (Zhang et al., 2021).

We carried out topic modelling using nouns in English, chosen for their important role in delivering key semantics about subjects (e.g. who is talking) and objects (e.g. what do they talk about). We used the pre-trained model in spaCy library (Benoit & Matsuo, 2020) to process our texts and extract lemmatized nouns as inputs for topic modelling. Lemmatising reduces words to a common root, allowing us to treat singular and plural forms as the same term. We then ran structural topic modelling (STM) using the R package 'stm' (Roberts et al., 2019), which employs latent Dirichlet allocation (LDA) for topic clustering. Since topic modelling is an unsupervised method, it is necessary to experiment to find the best-fitting number of topics, k, and we used two measures: coherence, which measures how well the terms in a topic are grouped in a document, and exclusivity, which captures how well topics split documents into different groups.

After performing topic modelling, we arranged the topics according to the descending order of topic proportions. For each topic, we retrieved the ten keywords with the highest probabilities and ten for the highest frex scoring algorithm (i.e., keyword extraction methods considering both frequent in and exclusive to a topic of interest) to understand the context of topics. After reviewing the keywords, the authors assigned labels that best describe the topics.

Since STM is a mixed membership model, one article can contain more than one topic. For analytic convenience, we assigned one prime topic for each article based on the highest topical probability. Assigning one single topic to each article is helpful in estimating the count and ratio of topics over time. We further extracted entity names and sorted them by frequently mentioned organization names within the sub-corpora.

To take a closer look at the narratives of newspapers, we chose the drought in the first half of 2012 (Drought in spring 2012 hereafter) and the drought in mid-2022 (Drought in summer 2022 hereafter), which were reported as extreme droughts on the one hand by the official reports of the Centre for Ecology & Hydrology (i.e., Major Hydrological Events) and research papers in the last two decades (e.g., the drought of 2012 by Parry et al., 2013; the drought of 2022 by Barker et al., 2024). By making sub-corpora for two droughts, we summarise topical composition using tree map plots (Tennekes, 2023), together with a colour-blind friendly palette 'safe' from 'rcartocolor' in R (Nowosad, 2018). Then we applied close reading (c.f. Dayrell et

al., 2022) to understand how the droughts were represented in context, by looking into headlines and body text corresponding to the topics.

3 Results

3.1 Newspaper reporting trends about droughts

3.1.1 Hydroclimatic anomalies and seasonality effect on newspaper reporting

For the 63 combinations of input variables, we ran the negative binomial regression and compared their AICs. As a result, we found the best AIC (i.e., 883.87) for the following combination: 12-month CET, 3-month SPI, 1-month SGI, and including seasonality. To graphically explore model performance, we plotted the expected results of the dependent variable with the modelled input variables and compared the result with the original dependent variable, i.e. the monthly count of newspaper articles (Supplementary Material 3). Full lists of the tested combinations and their AIC can be found in Supplementary Material 4.

The results of the negative binomial regression analysis are summarized in Table 1. With the dependent variable being the monthly count of newspaper articles, the 12-month CET variable positively associated with the outcome (β=0.395, p=0.067), indicating a marginally significant trend. The coefficient of β=0.395 suggests that for each unit increase in the 12-month CET index, the count of newspaper articles increases by the factor of $e^{0.395}$ or 1.484, an increase of approximately 48.4 %.

Conversely, both the 3-month SPI (β=−0.415, p<0.001) and the 1-month SGI (β=−0.868, p<0.001) were negatively associated with the count of newspaper articles, with strong significance. For each unit decrease in the 3-month SPI, the count of newspaper articles increases by 51.4 %, while a unit decrease in the 1-month SGI corresponds to an increase of 138.3 %.

Seasonal factors also played a substantial role in modelling the monthly count of newspaper articles. The winter season (i.e. DJF, β=0.693, p=0.030) showed a significant positive effect, indicating a 100 % increase in newspaper articles during this season compared to the baseline, i.e. autumn season. Similarly, the spring season (i.e., MAM, β=1.065, p<0.001) and summer season (i.e. JJA, β=1.108, p<0.001) had even stronger positive relationships, with expected increases of 189.3 % and 203 % respectively, suggesting much higher article counts in these periods than others.

The 1-month lag in the dependent variable (β=0.046, p<0.001) was significant, indicating that each additional article in the previous month increases the expected count in the current month by 4.7 %. This demonstrates a modest but persistent autocorrelation effect.

Table 1 Statistical results of negative binomial regression model on the count of newspaper articles

Variables			Coefficients	p-values	Interpretation
Variables	Best-fitting option	Range			

Temperature	12-month CET	(-1.334, 1.458)	0.395	0.067384	+1 unit ≈ +48.4% (.)
Precipitation	3-month SPI	(-2.432, 2.596)	-0.415	0.000733	- 1 unit ≈ +51.4% (***)
Groundwater	1-month SGI	(-1.584, 1.644)	-0.868	3.71e-06	- 1 unit ≈ +138.3% (***)
Autumn	SON (baseline)				
Winter	DJF	binary	0.693	0.030365	when it is winter, ≈ +100.0% (*) to Autumn
Spring	MAM	binary	1.065	0.000502	when it is spring, ≈ +189.3% (***) to Autumn
Summer	JJA	binary	1.108	0.000266	when it is summer, ≈ +203.0% (***) to Autumn
Lagging effect	1-month delay of dependent variable	(0, 191)	0.046	1.20e-12	+1 unit ≈ +4.7% (***)
Intercept			-0.736	0.002569	0.48 (**)
					p-values: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

3.1.2 Temporal dynamics between hydroclimatic anomalies and newspaper reporting

The combined plot of the volume of newspaper articles and meteorological/hydrological anomalies (Figure 1) shows how the reporting of newspaper co-developed with drought-prone conditions. From this plot, we found multiple drought events gathered substantial newspaper reporting, aligning with major drought events declared in governmental reports and scientific articles, e.g., in 2005-2006; 2011; 2017 and 2018/19 (Turner et al., 2021) and most markedly in 2022 (Barker et al., 2024). The plot also showed that most newspaper reporting for major droughts coincided with the drought-prone conditions, i.e. 12-month CET anomalies > 0, 3-month SPI < 0, and 1-month SGI < 0 (see the yellow shades). However, not all drought-prone conditions were reported on in the newspapers.

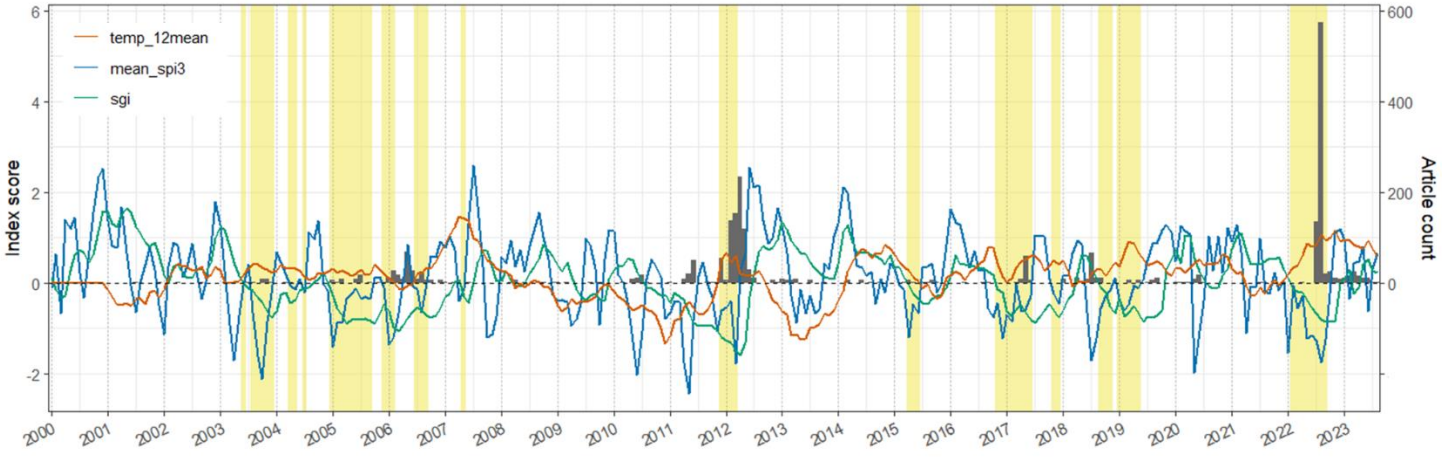

Figure 1 Time-series plot to show the newspaper reporting (i.e., monthly count of newspaper articles) (bar plot) and the hydroclimatic anomalies (i.e., 12-month CET, 3-month SPI, and 1-month SGI) along with the drought-prone conditions in yellow shades.

In Figure 2, we note that both the drought in 2012, which is more strongly related to long-term groundwater deficiency and that of summer 2022, related to a combination of positive temperature anomalies compounded by a precipitation and groundwater deficit were the subject of most newspaper reporting, and we chose these two events for more in-depth analysis using topic modelling.

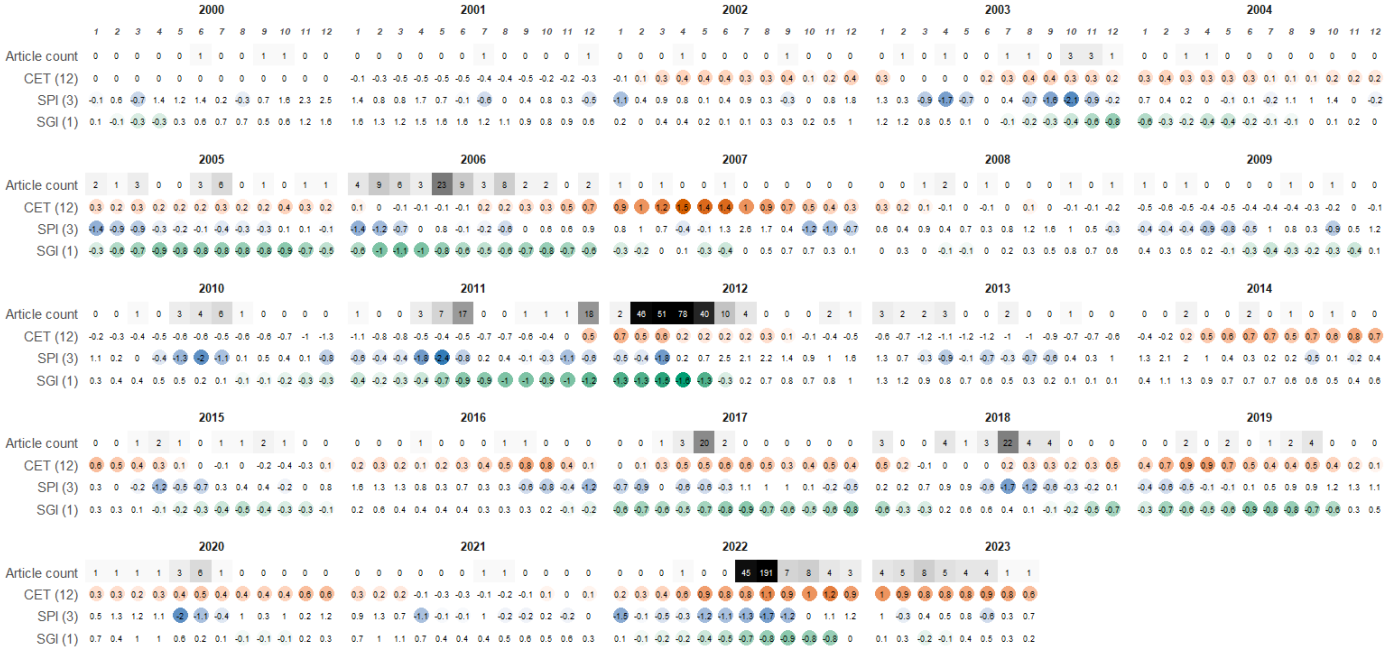

Figure 2 Time-series bubble plot to show the volume of newspaper articles (i.e., monthly count of newspaper articles) (grid in lilac on top) and the hydroclimatic anomalies (i.e., 12-month CET, 3-month SPI, and 1-month SGI) (bubble in colours).

3.2 Newspaper topics about drought

3.2.1 Topic labels in the corpus

According to the coherence and exclusivity test, we confirmed the optimal number of topics, k, to be 15 (see Supplementary material 5). For the 15 topics, we extracted frequent keywords based on probability and frex estimation, together with frequent organization names. As a result, we assigned the topic labels that best describe the keywords and frequent organization names, and summarised the result in Table 1. The order of topics follows the expected probability of topics in the corpus (Supplementary Material 6).

The most common topic in the corpus was "**Drought and hosepipe ban**", which described hosepipe bans, a popular measure for restricting water use in the UK, due to the severity and expanse of drought status around both meteorological (i.e. rainfall) and hydrological (i.e. river, groundwater) domains. Common organizations in this topic included the Environment Agency, the government body responsible for environment protection and management. Additionally, Thames Water and Southern Water, two large companies to supply water to Greater London and beyond, were identified, probably because they imposed the hosepipe ban. Interestingly, we found the topic of "**Temperature and hosepipe ban**" as the 11th most popular topic in the corpus, which also described the declaration of hosepipe ban, but in the context of temperature, not rainfall nor water reserves.

The second most common topic was "**Water shortage**", which contained the shortage in both meteorological (i.e. rainfall, rain) and hydrological (i.e. reservoir) droughts, yet not in the phase of announcing water restrictions, e.g. hosepipe bans. Frequent organizations other than the Environment Agency included the Met Office, which is responsible for weather forecasting in the UK. Interestingly, we found the fourth common topic to be "**Flooding**" seemingly unrelated to droughts. However, this topic addresses how the end of major droughts is reported – often contrast with periods of sustained rainfall and associated flooding (see Parry et al., 2013). Similar to the second topic that described weather conditions, this topic also featured the Environment Agency and the Met Office as major organizations.

The third common topic was "**Water management**", which described current managerial concerns regarding water supply companies, as shown in the keywords of 'leak' and 'leakage'. Accordingly, the topic captured frequent organizations to be major water suppliers, such as 'Thames Water', and the 'Ofwat', i.e. Water Service Regulation Authority in the UK.

The fifth topic was "**Water use impacts** (from hosepipe ban)". This topic described the range of impacts on daily water consumption when it comes to hosepipe bans, such as restricting water use for 'watering the garden' or 'car wash'. As hosepipe bans are announced by water suppliers, common organizations in this topic included 'Thames Water', 'South East Water', and 'Southern Water'. Less commonly, but we found topics that describe the broader impacts of droughts, as shown in "**Agriculture** (Topic #7)", which addressed the impact of droughts on 'crop', 'vegetable', and 'price', by delivering the voices of farmers (i.e. National Farmers Union) as key organization. Furthermore, we found the topics of "**Aqua ecosystems** (Topic #8)" and "**Terrestrial ecosystem**s (Topic #15)", which described the drought impact on 'wildlife', 'wetland', and 'tree', with their vocal organization to be the Environment Agency.

The sixth popular topic was "**Heatwave**", which described 'high temperature', resulting in 'heatwave' and '(wild)fire'. The most frequent organization in this topic was again the Met Office.

We also found discussions related to "**Climate change** (Topic #10)" based around keywords such as 'scientists' to global 'warming' or 'changing climate pattern(s)'. In addition to the Met Office and the Environment Agency, we found the European Environment Agency (EEA) in this topic.

The topic of "**Beavers** (Topic #14) was rare, but interesting given current narratives related to rewilding of this species in the UK (Holmes et al., 2024). We assume these articles link beavers and their activities to drought alleviation, through for example dam building and wetland restoration.

As is common in topic modelling some topics were note assigned labels (Topics of #9, #12, #13) since they were hard to interpret or irrelevant (e.g., Topic #9 seemed to include 'drought' in sport, especially football, and is therefore irrelevant).

Topic compositions for the entire study period can be found in Supplementary Material 7.

Table 2 Summary of 15 topics with their topic labels, probability keywords, frex keywords, and key organization names. The order of topics follows the expected topic proportions in the corpus.

#	Topic labels	Probability-based keyword (prob), frex keywords (frex), and frequent organization names (org). *In descending order of frequency*
1	Drought and hosepipe ban	• (prob) drought, water, ban, hosepipe, company, level, area, rainfall, river, part • (frex) restriction, status, groundwater, ban, hosepipe, drought, environment, agency, area, level • (org) the Environment Agency, Thames Water, Southern Water
2	Water shortage	• (prob) year, cent, per, water, rainfall, reservoir, month, level, rain, drought • (frex) per, cent, average, reservoir, mm, capacity, rainfall, figure, month, fear • (org) the Environment Agency, the Met Office, Thames Water
3	Water management	• (prob) water, company, supply, year, leak, drought, reservoir, leakage, plan, government • (frex) leak, target, leakage, meter, bill, metering, regulator, plan, industry, consumption • (org) Ofwat, Thames Water, the Environment Agency
4	Flooding	• (prob) rain, flood, flooding, drought, weather, week, water, warning, downpour, area • (frex) flooding, flood, downpour, flash, wind, subsidence, tornado, claim, road, sandbag • (org) the Environment Agency, the Met Office
5	Water use impacts from hosepipe ban	• (prob) water, ban, hosepipe, plant, garden, order, drought, car, people, company • (frex) drip, order, butt, sprinkler, lawn, watering, hose, can, car, swimming • (org) Thames Water, South East Water, Southern Water, the Environmental Agency
6	Heatwave	• (prob) temperature, day, week, heatwave, weather, water, fire, heat, condition, area • (frex) thunderstorm, temperature, barbecue, heatwave, fire, firefighter, blaze, wildfire, crew, high • (org) the Met Office, Thames Water, Southern Water
7	Agriculture	• (prob) farmer, crop, drought, food, price, year, weather, potato, condition, vegetable • (frex) potato, wheat, harvest, livestock, barley, grower, crop, yield, vegetable, carrot • (org) NFU (National Farmers Union), the Environmental Agency, EU
8	Aquatic	• (prob) water, river, drought, wildlife, year, fish, level, wetland, environment, bird

	ecosystems	• (frex) trout, vole, wetland, fish, salmon, stream, wildlife, lapwing, chalk, oxygen • (org) the Environment Agency, Thames Water, RSPB, WWF
9	n/a (frex)	• (prob) garden, year, plant, drought, flower, gardener, thing, goal, grass, day • (frex) prairie, goal, dahlia, trophy, game, show, rose, defender, euphorbia, verbascum • (org) Chelsea, RHS, Arsenal, Tottenham
10	Climate change	• (prob) climate, change, year, weather, drought, scientist, pattern, summer, world, warming • (frex) scientist, climate, pattern, change, warming, study, emission, extreme, strawberry, world • (org) the Met Office, the Environment Agency, EEA (European Environment Agency)
11	Temperature and hosepipe ban	• (prob) water, ban, customer, hosepipe, company, drought, temperature, area, day, supply • (frex) customer, leakage, bottle, property, bin, litre, health, wildfire, fire, amber • (org) Thames Water, the Environment Agency, Government
12	n/a (frex)	• (prob) water, summer, week, drought, day, rain, weather, people, temperature, heat • (frex) ladybird, standpipe, street, championship, memory, bath, swarm, trial, clock, umpire • (org) Thames Water, the Environment Agency
13	n/a (prob)	• (prob) year, cent, per, price, market, time, school, sale, month, beach • (frex) jellyfish, fever, market, share, pollen, sufferer, investor, school, analyst, beach • (org) Straight, the Met Office, the Environment Agency
14	Beavers	• (prob) year, water, beaver, drought, people, project, infrastructure, population, reservoir, country • (frex) beaver, cave, fund, bog, dam, peat, project, component, developer, debt • (org) Thames Water, Ofwat
15	Terrestrial ecosystems	• (prob) tree, drought, year, leave, summer, plant, water, autumn, flower, robinia • (frex) robinia, bluebell, tree, pseudoacacia, ivy, leave, leaf, blossom, blackberry, sapling • (org) the Environment Agency, Labour

3.2.2 Close reading for spring 2012 and summer 2022

For the droughts of spring 2012 and summer 2022, we visualized topic compositions using treemaps (Figure 3) to allow us to explore differences between topic use for these two events. The most distinctive difference between the spring 2012 and summer 2022 was the dominant topics: for the drought in spring 2012, "**Flooding**" was the most visible, whereas the drought in summer 2022 highlighted the topics of "**Drought and hosepipe ban**". For the topics in summer 2022, we found

two distinctive topics: first, newspaper attention to the drought in summer 2022 was attributed to the summer season, as highlighted by topics of "Heatwave" and "Temperature and hosepipe ban"; second, we observed the emergence of "Water management" as a prominent topic in the summer of 2022.

For the most dominant "**Flooding (Topic #4)**" topic in spring 2012, the newspaper articles positioned the flooding in contrast to the precedent droughts, e.g. "From drought to flood warnings in a week [Article 475, April 2012]", "April showers lead to floods but there's no end to drought [Article 77, April 2012]", and "How drought can lead to flooding [Article 559, May 2012]". This topic also appeared in the drought of summer 2022, "UK weather: warning of floods from thundery showers after drought [Article 867, August 2022]", but not as strong as spring 2012.

"**Drought and hosepipe ban** (Topic #1)" was the most prominent topic in summer 2022, yet it appeared in spring 2012 as well. The topic describes the onset of drought-prone conditions and the (possibility of) declaration of hosepipe ban, e.g., "Crippling drought hits south and east of England [Article 5, February 2012]" and "Drought fears after England suffers driest spell since '76 [Article 254, July 2022]". Water suppliers often appeared in the headlines related to this topic, e.g. "Southern Water announces hosepipe ban over drought fears [Article 764, July 2022]". An accompanying topic, "**Water use impacts (Topic 5)**" elaborates what and how people should be prepared when it comes to the drought and hosepipe ban, "Survive the drought and keep your garden going until the rain arrives [Article 787, March 2012]", "Can I wash my car during a hosepipe ban? When you CAN use your hose during drought [Article 1042, August 2022]".

Meanwhile, "**Temperature and hosepipe ban** (Topic #11)", which attributed high temperatures as a trigger for drought and hosepipe ban, appeared exclusively for summer 2022, with the headlines of "Drought expected to be declared amid heatwave with UK hotter than Caribbean [Article 251, August 2022]". In a similar context of summer season narratives, "**Heatwave (Topic #6)**" was visible in the summer 2022, implicating high temperatures as the driver of droughts, heatwave, and wildfire, as shown in the headlines, e.g., "Britons take to the beach as the country basks in 35C heat this weekend as drought is officially declared [Article 1036, August 2022]" and "Now we face triple threat of heatwave, fire and drought [Article 1231, August 2022]".

The topic of "**Water management (Topic #3)**" focussed on responsibility during the drought of summer 2022. Articles often pointed out the failure in managerial bodies to handle droughts, "Ministers accused of having no plan to deal with drought [Article 635, August 2022]", "Why is Britain not better prepared for this drought? [Article 953, August 2022]". Furthermore, the narrative deepens with discussions around water suppliers for their managerial failure to handle water inefficiency, e.g., "Water firms fail on leak targets as drought looms [Article 908, August 2022]".

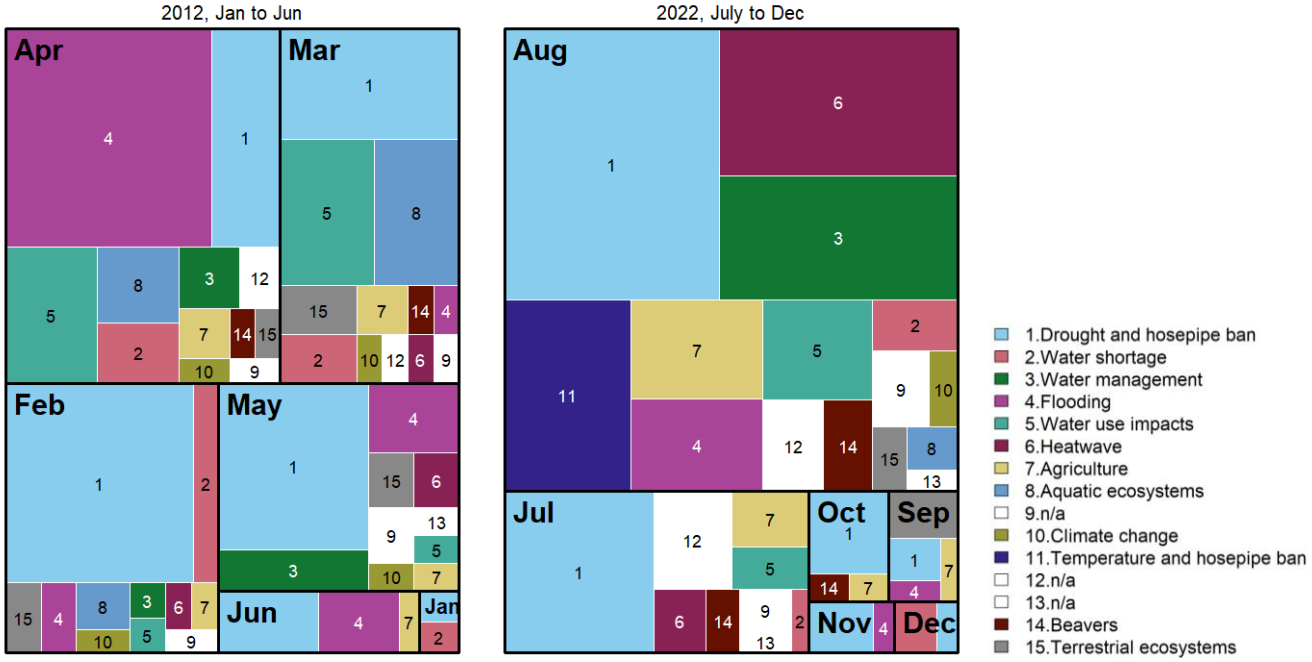

Figure 3 Treemaps comparing topic compositions for droughts in spring 2012 and summer 2022. The counts of monthly articles are 2, 46, 51, 78, 40, 10 for January to June, 2012, and 45, 191, 7, 8, 4,3 for July to December, 2022.

4 Discussion

4.1 How do hydroclimatic conditions relate to newspaper reporting?

Our findings confirmed that newspaper reporting about droughts was related to the conditions when the precipitation (SPI-3) was low, the groundwater table (SGI-1) was low, and temperature (CET-12) was high, and particularly when the season is spring or summer. Our results are novel in eliciting combinations of hydroclimatic variables and seasonality, which explain the volume of newspaper reports, and also in calculating the scale of contribution of explainable variables on the number of articles, specifically one unit decrease of SPI-3 or SGI-1 resulted in ~50 per cent or ~140 per cent of newspaper article being more published.

While we confirmed that such drought-prone conditions are positively correlated to newspaper attention, our findings also suggest that precipitation anomalies alone are insufficient to explain the volume of newspaper attention. This is intriguing to show the perception gap of drought between media and scientists, as hydrologists and meteorologists prioritise precipitation deficits when defining droughts (e.g., the drought reports from UK CED). Similarly, previous studies examining long-term newspaper reporting of droughts have predominantly focused on precipitation-related indices, such as the SPI (Dayrell et al., 2022; O'Connor et al., 2022). Meanwhile, we confirm the weak role of temperature anomalies in driving newspaper reporting. Such finding aligns with the work of Linés et al. (2017) who also found a temperature proxy, i.e. LST, to be an insignificant

index in explaining newspaper reports about drought. This is interesting because higher temperatures are associated with higher potential evapotranspiration, increasing the possibility and the magnitude of hydrological droughts. Indeed, high temperature anomalies tended to garner newspaper reporting about climate change (Pianta & Sisco, 2020). However, such a weak contribution of temperature on newspaper reporting about droughts highlights the precipitation deficit to be more critical factor to lead the perception of droughts.

A further interesting finding was concerning seasonality. While seasonality is often overlooked in related works, our findings confirmed that given identical hydroclimatic conditions, spring and summer seasons resulted in almost 200 per cent increase (189%, 203%) in newspaper articles, compared to autumn as a baseline. Similar results have been reported by a few researchers, for example O'Conner et al. (2022) noted that seasonality was a meaningful variable in driving the volume of newspaper reporting of droughts, with summer seasons leading to more articles with only a modest SPI-3 deficit. Parsons et al. (2019) suggested that perceived impact of droughts, especially agricultural droughts in warm seasons of July and August, corresponded to more documents in a corpus of farming periodicals in the UK. Such a seasonality bias, although it can be potentially reflecting public interest, newsworthiness, or the scale of actual damages, highlights the risk that early signs of drought during colder seasons may be overlooked by newspapers, thereby failing to generate support for proactive actions or preventive measures.

With a broader picture of temporal dynamics, two peaks of newspaper reporting about droughts (i.e. spring 2012, summer 2022) overlapped with drought-prone conditions. However, such conditions did not always generate noticeable newspaper reporting as few articles were found 2003, 2005, 2017, and 2018-2019. Interestingly, these years of drought-prone conditions are confirmed by the UKCEH to publish reports about unusual hydrological events, for the years of 2003 (Marsh, 2004), 2004-2006 (Marsh, 2007), 2010-12 (Marsh et al., 2013), and 2018-19 (Turner et al., 2021). In other words, our analysis of drought-prone conditions corresponded to the scientific acknowledgement of drought conditions by UKCEH, but media did not always pay attention to such occasions. Such a selective nature of media attention was also identified by Dayrell et al. (2022) from their text corpora about droughts over the last 200 years in the UK.

This mismatch, or more precisely the selective nature of newspaper reporting about droughts, may be explained by the newsworthiness of events: even if the meteorological conditions are favourable for droughts, they may not generate media coverage for those not making newsworthy impacts yet. As explained earlier, droughts are complex, lagged processes with precipitation anomalies linked to deficits in streamflow or groundwater and in the end, the water shortage. Thus, the observation of 'scientific' droughts does not immediately imply direct impacts on people, and thus incipient droughts rarely get media coverage (Dayrell et al., 2022). For example, the drought of 2004-2006 is observed as an unusual drought event by UKCED (Marsh, 2007), but the actual impact of the drought was minimal (Tanguy et al., 2021).

We discuss two more limitations in our approach that may also contribute to such mismatches: temporal windows and spatial granularity. As temporal windows, we considered 3, 6, 12 months to differentiate the long-term effects of low precipitation (SPI) and high temperature (CET). However, a fixed-window approach may not be suitable to explain all the fluctuations of newspaper reporting, because droughts events, or more specifically newspaper attention to droughts, are

uniquely situated: it can be intensified as a result of the duration of drought-prone conditions or it can be extinguished from the condition of just preceding months, especially the rainfall. For example, 2003 saw severe drought across continental Europe, with the UK also experiencing high temperature and low precipitation anomalies, marked by the second lowest February-October rainfall in 83 years. Nonetheless, English newspaper coverage in this period is not visible, likely because heavy rainfalls in the precedent winter, i.e., November and December of 2002, filled up reservoirs and recharged groundwater resources (Marsh, 2004). Similar patterns can also be found in 2005 and 2019, again leading to limited newspaper reporting with respect to incipient drought conditions. Conversely, the two peaks of newspaper reporting in spring 2012 and summer 2022 were accompanied by a years-long duration of drought-prone conditions, i.e., before spring 2012, there had been persistently low precipitation since mid-2010, leading to a deficit in groundwater, soil moisture, and leaving major reservoirs unfilled (Marsh et al., 2013). Likewise, for the summer of 2022, low precipitation anomalies commenced in the autumn season of 2021 and persisted until the summer of 2022, and hotter temperature anomalies resulted in severe droughts, calling for substantial newspaper reporting.

Another limitation of our study is its spatial granularity. When processing the indicator metrics, we aggregated the variables nationally (i.e., across England), obscuring regional variations or temporal transitions over geography. Although in principle it would be possible to assign news articles to smaller regions using methods from geographic information retrieval (Acheson & Purves, 2020), in practice this would further stratify our corpora, making statistical modelling challenging. To consider geographic variance better, which is important because droughts are geographically heterogeneous (Tanguy et al., 2021), a larger corpus would be necessary (Parsons et al., 2019). To overcome this limitation, it would also be required to prepare a complete set of documents that can be assigned to finer scales, e.g. catchments, corresponding to the scale of meteorological stations that generate hydroclimatic metrics (O'Connor et al., 2022).

4.2 Which topics are represented in newspapers?

Using the unsupervised method of computational text analysis, topic modelling, we identified 12 meaningful topics in drought-related newspaper corpus and visualised their compositions over time. By carrying out close reading we zoomed into the detail of newspaper narratives for two major drought events in the newspaper: spring 2012 and summer 2022. This multi-scalar method allowed us to take advantage of both methods, i.e., computational topic modelling elicits the overarching composition of topics and close reading complement the in-depth context of these topics in the corpus. As a result, this study enabled to distinguish between the newspaper narratives between two major droughts in England.

From the topical composition, we found the topic "**Drought and hosepipe ban (Topic #1)**" was present in almost all months. The topic encompasses meteorological and hydrological droughts in, e.g. rainfall and river, and further reports on both potential and actual declarations of hosepipe ban, a common management measure to reduce water usage in England, imposed by water companies. This highlights on the one hand, the role of media in informing readers about drought conditions that may pose significant impacts on daily lives, and on the other hand, in emphasizing readers concerning water supply and restrictions

when it comes to droughts. According to Stahl et al. (2016), a major concern related to droughts in the UK was public water supply, which contrasts with other countries where different concerns, e.g., drinking water quality or agriculture and livestock farming, emerged. Stahl et al. (2016) also revealed that 'local and regional water supply shortage and problems (drying up of springs/wells, reservoirs, streams)' and 'bans on domestic and public water use (e.g., car washing, watering the lawn/garden, irrigation of sports fields, filling of swimming pools)' consisted of major water concerns in the UK. These findings are very much in accord with the results of our topic modelling.

Repeated droughts and hosepipe ban in England seemed to have resulted in more discussion of water management, as shown by the topic "**Water management (Topic #3)**", which became much visible in the drought of summer 2022. Such a pattern can be ascribed to massive newspaper reporting in August 2022, scoring 191 articles for a single month. At the same time, close reading alluded that repeated droughts and hosepipe ban, at least in the last two decades, have galvanized the newspaper to redirect the attention to the water management scheme by the government and water suppliers. The results suggest growing criticism towards water supply companies being unable to handle metering or leakage issues, compounded with the privatization of water supply in England (Bayliss et al., 2021).

Topics related to "**Heatwave (Topic #6)**" and "**Temperature and hosepipe ban (Topic #11)**" seem to point to some confirmation bias concerning newspaper reporting. Both topics were very prominent in the summer of 2022 and through close reading, we concluded that these topics gravitated to typical summer narratives, e.g., relating high temperatures and heatwaves to drought conditions. Our negative binomial regression modelling also confirmed a seasonality bias – with spring and summer seasons resulting in increased newspaper reporting.

One important limitation of our content analysis lies in the nature of our corpus, which is restricted to a single genre: newspapers. Despite the advantages of newspapers as a data source (Boykoff, 2008), this genre has experienced significant declines in readership given the rise of online news platforms (Kim et al., 2019) and social media platforms such as Twitter/X (Antwi et al., 2022). Extending the corpus to incorporate more genres of writing about drought, e.g., by incorporating resources including the UK Drought Inventory (e.g., Parsons et al., 2019), and the European Drought Impact Inventory (EDII) (e.g., Stahl et al., 2016), as well as social media could provide valuable additional perspectives.

5 Conclusion

In this study, we constructed a corpus of newspaper articles reporting on droughts in England to analyse: (a) how newspaper reporting correlates with hydroclimatic conditions and seasonality and (b) the topics that emerge in newspaper narratives. Our statistical analysis confirmed that newspaper reporting is strongly influenced by low groundwater levels (SGI) and low precipitation levels (SPI-3). Interestingly, temperature (CET-12) had a negligible impact on garnering newspaper reporting, while seasonality, particularly warm seasons such as spring and summer, played a critical role in generating sizeable newspaper articles. Using topic modelling we could zoom into newspaper narratives on two severe drought events in spring 2012 and summer 2022. The results showed that newspaper narratives about droughts often gravitated towards drought reports

on meteorological and hydrological droughts together with hosepipe ban. On top of that, the summer drought of 2022 was interesting to reveal summer-related topics, underscoring the seasonality bias of droughts. This study contributes to further our understanding of how newspaper reporting dynamics are intertwined with hydroclimatic conditions and the perception of seasonality, and how the narratives fluctuated over the last 24 years. In the end, we suggested the potential mismatch between scientific droughts and media droughts, which calls for attention considering the role of media to affect public attitude and policy interventions.

Code availability

The R code lines we applied for the analysis were archived in a Markdown file, which can be found in the Supplementary Material 1.

Supplementary Materials

All the supplementary materials mentioned in the manuscript, including the R Markdown, can be found in the online version of publication.

Author contribution

All the authors (IK, JS, RP) conceptualised and developed the research questions. IK took a lead on implementation, including data collection, formal analysis, and visualisation, upon the communication with all authors. JS and RP prepared the original draft, and all the authors (IK, JS, RP) contributed the edits, reviews, revisions for the final draft.

Competing interests

One of our co-authors is a member of the editorial board of Hydrology and Earth System Sciences (HESS).

References

Acheson, E. and Purves, R. S.: Extracting and modeling geographic information from scientific articles, PLoS One, 16, 1–19, https://doi.org/10.1371/journal.pone.0244918, 2021.

Alexander, L. V. and Jones, P. D.: Updated Precipitation Series for the U.K. and Discussion of Recent Extremes, Atmos. Sci. Lett., 1, 142–150, https://doi.org/10.1006/asle.2001.0025, 2000.

Antwi, S. H., Rolston, A., Linnane, S., and Getty, D.: Communicating water availability to improve awareness and implementation of water conservation: A study of the 2018 and 2020 drought events in the Republic of Ireland, Sci. Total Environ., 807, 150865, https://doi.org/10.1016/j.scitotenv.2021.150865, 2022.

Barker, L. J., Hannaford, J., Magee, E., Turner, S., Sefton, C., Parry, S., Evans, J., Szczykulska, M., and Haxton, T.: An appraisal of the severity of the 2022 drought and its impacts, Weather, 1–12, https://doi.org/10.1002/wea.4531, 2024.

Bayliss, K., Mattioli, G., and Steinberger, J.: Inequality, poverty and the privatization of essential services: A 'systems of provision' study of water, energy and local buses in the UK, Compet. Chang., 25, 478–500, https://doi.org/10.1177/1024529420964933, 2021.

Benoit, K. and Matsuo, A.: spacyr: Wrapper to the 'spaCy' 'NLP' Library_. R package version 1.2.1. (https://CRAN.R-project.org/package=spacyr), 2020. (Last access on 30 May, 2024)

Blei, D., Carin, L., and Dunson, D.: Probabilistic Topic Models, IEEE Signal Process. Mag., 27, 55–65, https://doi.org/10.1109/MSP.2010.938079, 2010.

Bohr, J.: Reporting on climate change: A computational analysis of U.S. newspapers and sources of bias, 1997–2017, Glob. Environ. Chang., 61, https://doi.org/10.1016/j.gloenvcha.2020.102038, 2020.

Boykoff, M. T.: The cultural politics of climate change discourse in UK tabloids, Polit. Geogr., 27, 549–569, https://doi.org/10.1016/j.polgeo.2008.05.002, 2008.

Brimicombe, C.: Is there a climate change reporting bias? A case study of English-language news articles, 2017-2022, Geosci. Commun., 5, 281–287, https://doi.org/10.5194/gc-5-281-2022, 2022.

Caple, H.: News Values and Newsworthiness, Oxford Res. Encycl. Commun., 1–21, https://doi.org/10.1093/acrefore/9780190228613.013.850, 2018.

Dayrell, C., Svensson, C., Hannaford, J., McEnery, T., Barker, L. J., Baker, H., and Tanguy, M.: Representation of Drought Events in the United Kingdom: Contrasting 200 years of News Texts and Rainfall Records, Front. Environ. Sci., 10, 1–18, https://doi.org/10.3389/fenvs.2022.760147, 2022.

Dessai, S. and Sims, C.: Public perception of drought and climate change in southeast england, Environ. Hazards, 9, 340–357, https://doi.org/10.3763/ehaz.2010.0037, 2010.

Dobson, B., Coxon, G., Freer, J., Gavin, H., Mortazavi-Naeini, M., and Hall, J. W.: The Spatial Dynamics of Droughts and Water Scarcity in England and Wales, Water Resour. Res., 56, https://doi.org/10.1029/2020WR027187, 2020.

Dow, K.: News coverage of drought impacts and vulnerability in the US Carolinas, 1998-2007, Nat. Hazards, 54, 497–518, https://doi.org/10.1007/s11069-009-9482-0, 2010.

Folland, C. K., Hannaford, J., Bloomfield, J. P., Kendon, M., Svensson, C., Marchant, B. P., Prior, J., and Wallace, E.: Multi-annual droughts in the English Lowlands: A review of their characteristics and climate drivers in the winter half-year, Hydrol. Earth Syst. Sci., 19, 2353–2375, https://doi.org/10.5194/hess-19-2353-2015, 2015.

Green, J. A.: Too many zeros and/or highly skewed? A tutorial on modelling health behaviour as count data with Poisson and negative binomial regression, Heal. Psychol. Behav. Med., 9, 436–455, https://doi.org/10.1080/21642850.2021.1920416, 2021.

Guerreiro, S. B., Dawson, R. J., Kilsby, C., Lewis, E., and Ford, A.: Future heat-waves, droughts and floods in 571 European cities, Environ. Res. Lett., 13, https://doi.org/10.1088/1748-9326/aaaad3, 2018.

Hart, P. S., Nisbet, E. C., and Myers, T. A.: Public attention to science and political news and support for climate change mitigation, Nat. Clim. Chang., 5, 541–545, https://doi.org/10.1038/nclimate2577, 2015.

Holmes, G., Rowland, G., and Fox, K.: Eager about beavers? Understanding opposition to species reintroduction, and its implications for conservation, People Nat., 6, 1524–1537, https://doi.org/10.1002/pan3.10674, 2024.

IPCC: Sections. In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 35-115, https://doi.org/10.59327/IPCC/AR6-9789291691647, 2023.

Kay, A. L., Bell, V. A., Davies, H. N., Lane, R. A., and Rudd, A. C.: The UKSCAPE-G2G river flow and soil moisture datasets: Grid-to-Grid model estimates for the UK for historical and potential future climates, Earth Syst. Sci. Data, 15, 2533–2546, https://doi.org/10.5194/essd-15-2533-2023, 2023.

Kim, S., Shao, W., and Kam, J.: Spatiotemporal patterns of US drought awareness, Palgrave Commun., 5, 1–9, https://doi.org/10.1057/s41599-019-0317-7, 2019.

Linés, C., Werner, M., and Bastiaanssen, W.: The predictability of reported drought events and impacts in the Ebro Basin using six different remote sensing data sets, Hydrol. Earth Syst. Sci., 21, 4747–4765, https://doi.org/10.5194/hess-21-4747-2017, 2017.

Marsh, T. The 2004-2006 drought in southern Britain, Weather, 62, 2007.

Marsh, T., Cole, G., and Wilby, R.: Major droughts in England and Wales, 1800-2006, Weather, 62, 87–93, https://doi.org/10.1002/wea.67, 2007.

Marsh, T.: The UK drought of 2003 - an overview, Weather, 59, 2004. (https://nora.nerc.ac.uk/id/eprint/510184/1/N510184CR.pdf)

Marsh, T.J., Parry, S., Kendon, M.C., and Hannaford, J.: The 2010-12 drought and subsequent extensive flooding. Centre for Ecology & Hydrology, 2013. (https://nora.nerc.ac.uk/id/eprint/503643/1/N503643CR.pdf)

Murphy, C., Wilby, R. L., Matthews, T., Horvath, C., Crampsie, A., Ludlow, F., Noone, S., Brannigan, J., Hannaford, J., McLeman, R., and Jobbova, E.: The forgotten drought of 1765–1768: Reconstructing and re-evaluating historical droughts in the British and Irish Isles, Int. J. Climatol., 40, 5329–5351, https://doi.org/10.1002/joc.6521, 2020.

Noone, S., Broderick, C., Duffy, C., Matthews, T., Wilby, R. L., and Murphy, C.: A 250-year drought catalogue for the island of Ireland (1765–2015), Int. J. Climatol., 37, 239–254, https://doi.org/10.1002/joc.4999, 2017.

Norton, C. and Hulme, M.: Telling one story, or many? An ecolinguistic analysis of climate change stories in UK national newspaper editorials, Geoforum, 104, 114–136, https://doi.org/10.1016/j.geoforum.2019.01.017, 2019.

O'Connor, P., Murphy, C., Matthews, T., and Wilby, R. L.: Relating drought indices to impacts reported in newspaper articles, Int. J. Climatol., 43, 1796–1816, https://doi.org/10.1002/joc.7946, 2023.

Osaka, S., Painter, J., Walton, P., and Halperin, A.: Media representation of extreme event attribution: A case study of the 2011–17 california drought, Weather. Clim. Soc., 12, 847–862, https://doi.org/10.1175/WCAS-D-19-0050.1, 2020.

Painter, J., Ettinger, J., Doutreix, M. N., Strauß, N., Wonneberger, A., and Walton, P.: Is it climate change? Coverage by online news sites of the 2019 European summer heatwaves in France, Germany, the Netherlands, and the UK, Clim. Change, 169, https://doi.org/10.1007/s10584-021-03222-w, 2021.

Parker, D. E., Legg, T. P., and Folland, C. K.: A new daily central England temperature series, 1772–1991, Int. J. Climatol., 12, 317–342, https://doi.org/10.1002/joc.3370120402, 1992.

Parsons, D. J., Rey, D., Tanguy, M., and Holman, I. P.: Regional variations in the link between drought indices and reported agricultural impacts of drought, Agric. Syst., 173, 119–129, https://doi.org/10.1016/j.agsy.2019.02.015, 2019.

Parry, S., Marsh, T., and Kendon, M.: 2012: From drought to floods in England and Wales, Weather, 68, 268–274, https://doi.org/10.1002/wea.2152, 2013.

Quesnel, K. J. and Ajami, N. K.: Changes in water consumption linked to heavy news media coverage of extreme climatic events, Sci. Adv., 3, 1–10, https://doi.org/10.1126/sciadv.1700784, 2017.

R Core Team (2023). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Roberts, M. E., Stewart, B. M., and Tingley, D.: Stm: An R package for structural topic models, J. Stat. Softw., 91, https://doi.org/10.18637/jss.v091.i02, 2019.

Robins, L., Burt, T. P., Bracken, L. J., Boardman, J., and Thompson, D. B. A.: Making water policy work in the United Kingdom: A case study of practical approaches to strengthening complex, multi-tiered systems of water governance, Environ. Sci. Policy, 71, 41–55, https://doi.org/10.1016/j.envsci.2017.01.008, 2017.

Schmidt, A., Ivanova, A., and Schäfer, M. S.: Media attention for climate change around the world: A comparative analysis of newspaper coverage in 27 countries, Glob. Environ. Chang., 23, 1233–1248, https://doi.org/10.1016/j.gloenvcha.2013.07.020, 2013.

Sodoge, J., Kuhlicke, C., and de Brito, M. M.: Automatized spatio-temporal detection of drought impacts from newspaper articles using natural language processing and machine learning, Weather Clim. Extrem., 41, 100574, https://doi.org/10.1016/j.wace.2023.100574, 2023.

Spinoni, J., Vogt, J. V., Naumann, G., Barbosa, P., and Dosio, A.: Will drought events become more frequent and severe in Europe?, Int. J. Climatol., 38, 1718–1736, https://doi.org/10.1002/joc.5291, 2018.

Stahl, K., Kohn, I., Blauhut, V., Urquijo, J., De Stefano, L., Acácio, V., Dias, S., Stagge, J. H., Tallaksen, L. M., Kampragou, E., Van Loon, A. F., Barker, L. J., Melsen, L. A., Bifulco, C., Musolino, D., De Carli, A., Massarutto, A., Assimacopoulos, D., and Van Lanen, H. A. J.: Impacts of European drought events: Insights from an international

database of text-based reports, Nat. Hazards Earth Syst. Sci., 16, 801–819, https://doi.org/10.5194/nhess-16-801-2016, 2016.

Tanguy, M., Haslinger, K., Svensson, C., Parry, S., Barker, L. J., Hannaford, J., and Prudhomme, C.: Regional Differences in Spatiotemporal Drought Characteristics in Great Britain, Front. Environ. Sci., 9, https://doi.org/10.3389/fenvs.2021.639649, 2021.

Tennekes, M.: _treemap: Treemap Visualization_. R package version 2.4-4, 2023. https://CRAN.R-project.org/package=treemap (last access on 30 May, 2024).

Tijdeman, E., Hannaford, J., and Stahl, K.: Human influences on streamflow drought characteristics in England and Wales, Hydrol. Earth Syst. Sci., 22, 1051–1064, https://doi.org/10.5194/hess-22-1051-2018, 2018.

Turner, S., Barker, L. J., Hannaford, J., Muchan, K., Parry, S., and Sefton, C.: The 2018/2019 drought in the UK: a hydrological appraisal, Weather, 76, 248–253, https://doi.org/10.1002/wea.4003, 2021.

UK CEH (UK Center for Ecology & Hydrology). (n.d.) Major Hydrological Events (https://nrfa.ceh.ac.uk/occasional-reports) and Drought Inventory (https://www.ceh.ac.uk/our-science/projects/drought-inventory) (last access on 18 November, 2024).

Vargas Molina, J., Paneque Salgado, P., and Augusto Breda Fontão, P.: Drought-related media analysis from Andalusia and São Paulo, Environ. Hazards, 21, 174–197, https://doi.org/10.1080/17477891.2021.1932712, 2022.

Van Loon, A. F.: Hydrological drought explained, Wiley Interdiscip. Rev. Water, 2, 359–392, https://doi.org/10.1002/WAT2.1085, 2015.

Wei, J., Wei, Y., Western, A., Skinner, D., and Lyle, C.: Evolution of newspaper coverage of water issues in Australia during 1843–2011, Ambio, 44, 319–331, https://doi.org/10.1007/s13280-014-0571-2, 2015.

Wilhite, D. A. and Glantz, M. H.: Understanding: The drought phenomenon: The role of definitions, Water Int., 10, 111–120, https://doi.org/10.1080/02508068508686328, 1985.

Wong, B.: Color blindness, Nat. Methods, 8, 441, https://doi.org/10.1038/nmeth.1618, 2011.

Zhang, T., Shen, S., Cheng, C., Su, K., and Zhang, X.: A topic model based framework for identifying the distribution of demand for relief supplies using social media data, Int. J. Geogr. Inf. Sci., 35, 2216–2237, https://doi.org/10.1080/13658816.2020.1869746, 2021.