Flash droughts and their impacts – using newspaper articles to assess the perceived impacts of rapidly emerging droughts *[UNDER REVIEW]*

Pedro HL Alencar, Jan Sodoge, Eva N Paton, Mariana M de Brito

Technical University Berlin, Institut of Ecology, Ernst-Reuter-Platz 1, 10587 Berlin, Germany

E-mail: `pedro.alencar@campus.tu-berlin.de`

February 2023

Abstract. Flash droughts have attracted increasing interest over the past decade. They are characterized by a rapid depletion of soil moisture resulting from interactions between soil and atmospheric conditions. To this date, there is a lack of consistent flash drought definitions and an understanding of their socio-economic impacts. In this study, we explore the relationship between biophysical flash drought conditions and the perceived impacts of flash droughts in Germany between 2000 and 2022. To measure perceived impacts, we assess the impacts reported in newspaper articles (2000-2022) and online search behavior from Google trends data (2004-2022). To characterize the physical process, we consider root zone soil moisture data. Our results show that flash droughts are becoming increasingly frequent in Germany, occurring on average every second year. Despite a lack of knowledge from the general public regarding the phenomenon of flash droughts, the peaks in interest in drought impacts correspond with the physical occurrence of flash droughts across the country. We identified an average time gap of four weeks between the flash drought onset and the reporting of perceived impacts. This gap is longer than the average duration of flash droughts' onset duration. Consequently, our findings highlight that the consistent monitoring of flash drought conditions and drivers is necessary to guarantee effective preparedness: as impact perception is too slow to allow the adoption of mitigation measures, flash droughts require new schemes for response measures compared to slowly emerging (conventional) drought events. The novel method also allows the consistent and impact-based validation of flash drought identification methods.

Keywords: Digital trace data, time series, validation, trend analysis, correlation analysis
Submitted to: *Environ. Res. Lett.*

1. Introduction: file preparation and submission

In recent years, the scientific community has witnessed an increasing interest in studying flash droughts. Particularly the disastrous 2012 flash drought which devasted central areas in the USA caused increased attention of scholars on a term that was rarely used before (Lisonbee, Woloszyn, and Skumanich 2021; Basara et al. 2019). This event led to staggering losses of over US$30 billion across the country as the rapid progression from wet-to-dry soil moisture hindered timely prevention and impact mitigation (Christian, Basara, Otkin, et al. 2019). This event brought to the public's attention that rapid onset drought conditions pose a new threat to the water-food nexus (Pendergrass et al. 2020), which is estimated to be occurring even more frequently under climate change (Yuan et al. 2023).

The rapid onset is the decisive factor for defining flash droughts. Multiple conflicting definitions have been proposed since its introduction in the academic literature (Alencar and Paton 2022). An initial debate concerned whether flash droughts were short-lived droughts (Mo and Lettenmaier 2016), or intense evolving droughts (Ford and Labosier 2017). The current consensus supports the latter (Christian, Hobbins, et al. 2024). Leading understanding describes flash droughts as emerging rapidly by transitioning from average to high soil moisture conditions to dry conditions (Yuan et al. 2023). Yet, seminal questions remain concerning what should be considered a flash drought. While clear thresholds have been proposed (Christian, Basara, Otkin, et al. 2019; Li et al. 2020; Osman et al. 2021; Yuan et al. 2023), these lack transferability across different case studies (Alencar and Paton 2022). The absence of well-established criteria led to a growing diversity of definitions and monitoring tools for flash droughts. These definitions and tools, however, still lack validation and transferability (Lisonbee, Woloszyn, and Skumanich 2021; Alencar and Paton 2022). Hence, flash drought research would benefit from studies applying the lenses of multiscale and multi-variable frameworks (Liang and Yuan 2021).

In this article, we propose a different perspective for flash drought identification. Like their slower counterparts (i.e., longer and slow-evolving droughts; henceforth addressed as conventional droughts), flash droughts cause impacts on multiple sectors of society, such as agriculture, animal husbandry, energy and water supply, and social activities (Christian, Basara, Hunt, et al. 2020; Walker et al. 2024). While various existing flash drought indexes are centered on hazards (Alencar and Paton 2022), these indexes may not consistently translate into socioeconomic impacts. Therefore, uncertainty persists in determining the conditions that lead to impactful flash droughts. Here, we suggest that understanding society's perception of the impacts of flash droughts can assist us in diminishing these uncertainties.

Here, we present a novel approach to studying flash droughts by considering the link between the biophysical occurrence of flash droughts and their perceived impacts. We assess the impacts based on impacts reported in newspaper articles and internet search behavior of local populations.

Newspapers often serve as real-time recorders of societal responses and reactions to extreme weather events (e.g. Sodoge et al. 2023; Bogdanovich et al. 2024), offering a unique perspective on the social impacts of flash droughts. Similarly, internet searches can approximate public awareness surrounding a specific topic (Kam, Stowers, and S. Kim 2019; Yesuel Kim and Youngchul Kim 2023). This approach can help researchers gain a more comprehensive understanding of the socio-economic implications of flash droughts, enhancing our ability to develop effective strategies for preparedness, mitigation, and adaptation.

To test if digital trace data (e.g. news articles and Google searches) can support the identification of flash drought occurrence, we consider the case of Germany from 2000 to 2022. Germany was chosen as a case study because of its recent history of significant conventional droughts (2003, 2015, 2018-2022) that had widespread impacts on sectors such as agriculture, forestry, livestock, or waterways navigation which gained relevance in public attention (Rakovec et al. 2022; Tijdeman and Menzel 2021; Sodoge et al. 2023). We obtain daily time series for biophysical flash drought occurrence, reported impacts in newspapers, and online search behavior on a district level in Germany. Thereby, we investigate three hypotheses:

- <u>H1</u> – Flash drought is a common phenomenon, and its occurrence has increased over the last few decades in Germany.
- <u>H2</u> – The impacts reported in the media and searched by internet users are a source of risk assessment that can be leveraged to identify when flash droughts impact society.
- <u>H3</u> – There is a delay between flash drought initiation and impact reporting due to latency time for water/plant/soil/society response to dryer conditions.

2. Data and Methods

Our research design comprises multiple steps that involve three distinct datasets: soil moisture, online search behavior measured via Google trends, and reported drought impacts in newspaper articles (Figure 1). For each, we obtained daily time series on a district-level using the statistical regions of Germany defined by the Nomenclature of Territorial Units for Statistics (NUTS) levels 1, 2, and 3. We used the most fine-grained NUTS level available for each dataset for the main analysis and provided results for all NUTS levels in the supplementary material. To explore the relationship between hazard and impact-related indicators, we employed time-series cross-correlation analysis. Subsequent sections provide a detailed description of each step in the research design.

2.1. Soil moisture data and flash drought identification

For identifying flash droughts, we used the method suggested by Ford and Labosier (2017) which is based on root-zone soil moisture and was successfully applied in multiple case studies (Otkin, Svoboda, et al. 2018; Lesinger and Tian 2022). For the soil moisture,

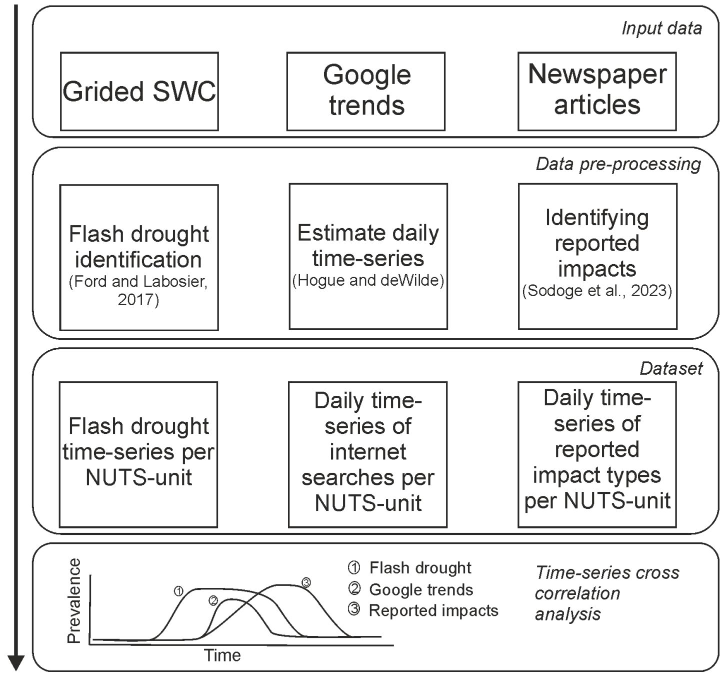


Figure 1: Overview of the research design: based on the three datasets for soil moisture, reported drought impacts, and online search behavior, daily time series are obtained on a district level in Germany. Then, cross-correlation analysis is employed to analyze the relationship between hazard indicators and consequent impacts.

we used data for Germany from Boeing et al. (2022) which covers the period of 1980 to 2022 and has a 1km spatial resolution of daily soil moisture at different soil depths. The data was generated through the mesoscale hydrological model (mHM – Samaniego et al. 2017) and demonstrated a robust correlation ($\rho > 0.7$; p-value < 0.05) with observed soil moisture in TERENO (Bogena et al. 2022) and FluxNET (Pastorello et al. 2020) stations across Germany. The optimal performance of the soil moisture model was observed from April to November, corresponding to the primary flash drought season in Central Europe (Boeing et al. 2022; Alencar and Paton 2022). Following the approach by Ford and Labosier (2017), we focused on data at 40 cm of depth and re-scaled it to a resolution of 4 km to reduce local noise. Then, we computed the accumulated soil moisture values at 5-day intervals (pentads). A flash drought was considered to occur if (1) soil moisture dropped from at or above the 40th percentile to at or below the 20th percentile within 4 pentads, (2) the soil moisture remained below the 30th percentile for at least 3 pentads after the onset to avoid classifying very short events, and (3) the event ended after soil moisture reached the 40th percentile. We computed these criteria using the detailed soil moisture dataset by Boeing et al. (2022) and the method of Ford and Labosier (ibid.) implemented in the R-package fdClassify (Alencar and Paton 2022). Criterion 1 defines the intensification phase, which ends at the flash drought onset. Criterion 2 defines a minimum persistence after onset to avoid over-identification of events (Pendergrass et al. 2020), and criterion 3 establishes the end of the event, indicating the return to normal conditions.

While the spatial resolution of the flash drought dataset corresponds to 4 km, we aggregated the occurrences at NUTS-1, NUTS-2, and NUTS-3 levels. For each NUTS level, the percentage of the area under flash drought conditions was computed at each pentad. At the NUTS-1 level, the three German city-states (Berlin, Bremen, and Hamburg) were joined with the surrounding states (Brandenburg, Niedersachsen, and Schleswig-Holstein, respectively) due to their relatively small size compared to the other states.

2.1.1. Trend analysis To explore temporal patterns in the occurrence of flash droughts in Germany (H1), we assessed the trends in their intensity, duration, and frequency from 1980 to 2022. In this study, we define flash drought intensity as the mean drop in soil moisture percentiles over the intensification phase, measured in percentile.day-1. Duration refers to the period from the start of intensification until the return to normal conditions. Finally, frequency is the expected number of events yearly, ranging from 0 to 1 (i.e., in years with more than one flash drought event, one is counted).

We used the regional Mann-Kendall test to perform this trend analysis and to compute and identify possible changes in flash drought intensity, duration, and frequency (Mann 1945; Helsel et al. 2020). This test evaluates the monotonic trend in the time series data to identify the significance of increasing or decreasing trends. For this research, we used the data on identified flash droughts (see section 2.1) on the NUTS-2 level to assess whether the frequency, duration, and intensity of flash droughts in Germany have increased significantly during the study period (H1). Furthermore, we computed the regional Theil-Sen estimator (Theil 1992; Sen 1968) to estimate the slope of these trends (i.e., the trend direction and magnitude). The Theil-Sen estimator, known for its robustness in assessing rates of changes in time-series, is less sensitive to outliers, and does not rely on assumptions regarding the distribution of data and residuals (Wang and Yu 2005). It is computed as the median of the slope of all lines that connect all pairs of points in the data series (Helsel et al. 2020).

2.2. Digital trace data on perceived impacts and public awareness

2.2.1. Socio-economic impact data To assess the perceived impacts of flash droughts on different socio-economic sectors, we used information on reported socio-economic impacts in newspaper articles. Specifically, we used the approach and data by Sodoge et al. (2023) that consider the frequency of newspaper articles reporting on a specific drought impact by leveraging text-mining tools. Therefore, we examined approximately 50,000 newspaper articles that included terms related to conventional drought (e.g. Dürre* and Trockenheit in German). These were published between 2000 and 2022 in 250 national, regional, and local German newspaper outlets included in the genios.de newspaper aggregator database. Articles were automatically reviewed using text-mining techniques to detect reported impacts in each article and its geographic scope using

NUTS 1, 2, and 3 levels. First, duplicate articles were automatically removed by considering the Jaccard similarity score (Mullen 2020). Then, we categorized the types of impacts using lasso logistic regression models. These models were trained and assessed using a set of 1,800 manually annotated newspaper articles (Madruga de Brito, Kuhlicke, and Marx 2020). Specifically, the following impact types are covered: agriculture (including crop yield losses), energy (including impacts on reduced energy production in nuclear power plants), fire (fire in forests or other areas due to drought conditions), livestock (impacts on livestock farming and animal populations), and social (impacts on places and activities used for recreation, tourism, leisure). To locate impact areas, we identified cities, regions, and districts mentioned in each article. We then selected the regional cluster with the highest frequency of detected locations. Previous work by Sodoge et al. (2023) provides a detailed description of this procedure, including an extensive validation with external indicators and observed impacts such as crop yield losses, forest fire statistics, and precipitation deficit.

The resulting dataset holds daily time-series data for each NUTS unit (levels 1,2,3) of the number of reported impacts. To compare the data to the flash drought measure, we further aggregated the time series into pentads. At the NUTS-1 level, the three German city-states (Berlin, Bremen, and Hamburg) were joined with the surrounding states (Brandenburg, Niedersachsen, and Schleswig-Holstein, respectively) due to their small size compared to the other states.

2.2.2. Public awareness data To infer public awareness of flash droughts, we collected information on spatio-temporal patterns of internet search frequency for the topic of 'drought' using Google trends data. This service provides an index of how many users of Google in a particular day and region search for a specific term (i.e. 'drought') where a higher value (maximum 100) corresponds to the maximum interest for a term. Previous research has shown the capabilities of Google trends data to monitor awareness of droughts from a spatio-temporal perspective (Kam, Stowers, and S. Kim 2019). However, Google does not unveil details and changes in the underlying algorithms. We extracted daily time series for 2004-2022 for each NUTS-1 (only scale available) unit using the PyTrends Python package (see details in Hogue and DeWilde 2023). Again, we aggregated the time series into pentads for comparing the different datasets in subsequent analyses.

2.3. Correlation analysis

To identify if the reported impacts in the news media and the internet search behavior are influenced by the occurrence of flash droughts (hypothesis H2), we performed a cross-correlation analysis between the datasets. This allowed us to assess temporal delays between the occurrence of flash droughts and impact perception (hypothesis H3). Specifically, we employ cross-correlation functions (ccf - Shumway and Stoffer 2019) which can identify delays between two variables, offering insights into whether one variable can predict the other. For this research, these delays represent the asynchronous behaviors expected between peaks of flash drought onset and impact reporting (hypothesis H3). Cross-correlation functions are commonly used to assess the similarity of two series (waveforms) as a function of their relative delay (Eq. 1).

$$\rho_{xy}(i,j) = \frac{\gamma_{xy}(i,j)}{\sqrt{\sigma_x(i,i)\,\sigma_y(j,j)}} \tag{1}$$

Where x and y are time series, γ is the co-variance between x and y, ρ_x and ρ_y are the standard deviations of x and y, and i and j are time indexes. It is important to notice that i and j are related, being $j = i + h$, where h is the delay between the time series. Using equation 1, we can identify the delay between both time series x and y (e.g. flash drought onset and impact reporting) as being the h that maximizes $\rho_{x,y}(i, i+h)$.

3. Results

H1 - *Flash drought is a frequent and intensifying phenomenon in Germany*

Using a high-resolution soil moisture dataset (Boeing et al. 2022) and the method of Ford and Labosier (2017), we assessed the occurrence of flash droughts in Germany from 1980 to 2022 (see Fig. 2). Results show that flash droughts are a frequent phenomena in Germany, with an average frequency of one event every 1.5 years. Particularly, the south and southwestern regions experience the highest frequency of events, with flash droughts occurring almost every year (frequency of approx. 0.87). Conversely, the northeast, including the federal states of Saxony-Anhalt, Brandenburg, and the city of Berlin, holds a lower rate of flash drought occurrence. Here, flash droughts take place for an average of about 55% of the analyzed years.

The temporal trends analysis revealed that flash droughts increased in frequency and affected areas over the last decades (see Fig. 3). Between 1980 and 2022, Germany has experienced flash drought events every year, at least in parts of its territory, with minimum values of affected areas at the range of 10% and a notorious maximum of approximately 99% in the year of 2020. The area affected has consistently increased at a rate of over 9% per decade. Hence, the average yearly affected area by flash droughts in Germany increased from about 40% at the beginning of the 1980s to about 80% in the 2020s. In particular, 2007, 2012, 2020, and 2022 had significant flash drought occurrences across Germany.

Using the regional Mann-Kendall test, we assessed the spatial distribution of trends in flash drought intensity, duration, and frequency (Figure 4). Only positive trends were identified across the NUTS-2 regions in Germany for all three variables. Furthermore, most trends are statistically significant, as indicated by the hatched areas. These findings indicate that flash droughts are not only a frequent (and almost yearly in some regions) phenomenon in Germany (H1) but an intensifying one. Figure 4a shows that flash droughts intensified across the southern portion of Germany, with rates increasing up to 1 percentile.day^{-1}.decade^{-1}. Meanwhile, Figure 4b illustrates changes

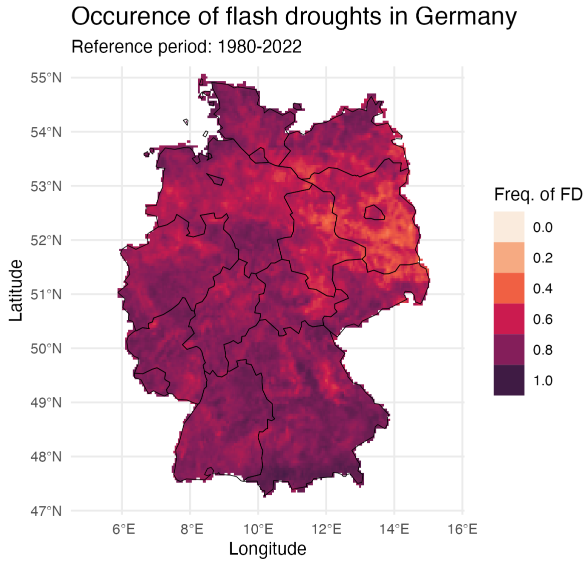


Figure 2: Frequency of flash drought (FD) occurrence in Germany in the period of 1980 to 2022. Events were identified following the Ford and Labosier (2017) approach with data from Boeing et al. (2022).

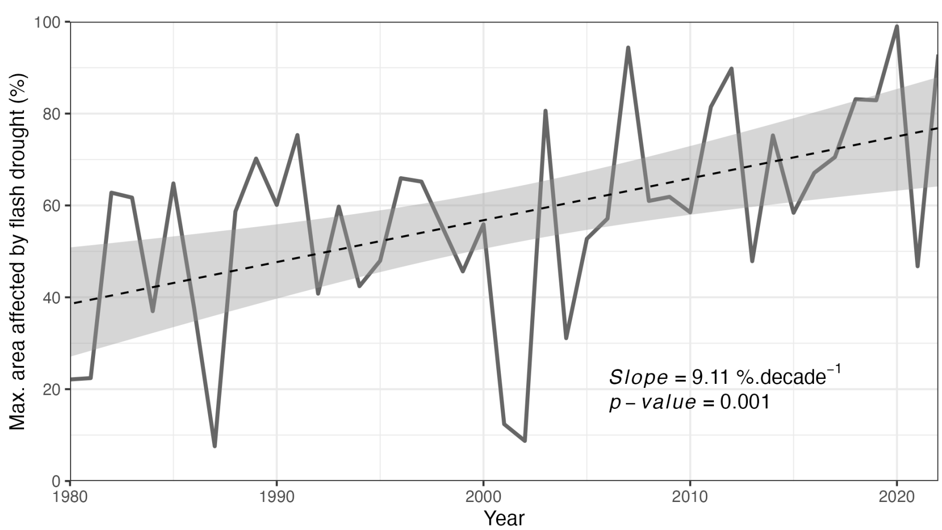


Figure 3: Yearly percentage of areas affected by flash drought in Germany. The flash droughts were identified based on the method of Ford and Labosier (2017) using data from Boeing et al. (2022). Applying the Mann-Kendall test, we identify strong evidence (p-value ¡ 0.01) of an upward trend in the yearly affected area, with an increase of over 9% per decade.

in the duration of flash droughts, showcasing particularly significant trends in the south and west regions. Here, flash droughts have become notably longer over the last decades, with durations increasing up to 30 days.decade^{-1}. These two results combined underscore the heightened hazard posed by flash droughts in the south and southwest of the country, where soil is rapidly drying at a faster pace and persisting for longer durations.

In Figure 4c we present the changes in flash drought occurrence in terms of events per year. These events have become more frequent across all regions at an average of 5%.decade^{-1}. The region with the largest increase was in the city of Berlin (11%.decade^{-1}). NUTS 2 regions in the south presented a slightly lower trend in flash drought occurrence compared to their northern counterparts, in contrast to the patterns observed for the other two variables analyzed.

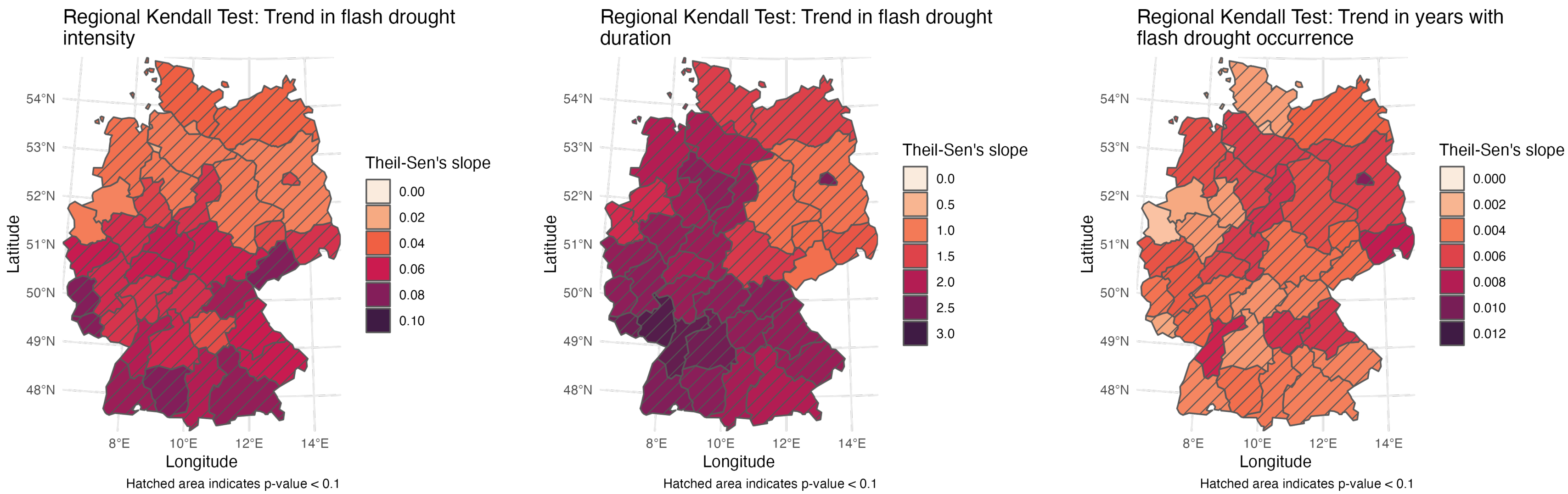
H2 - The blind perception of flash droughts

The term "*Blitzdürre*" (flash drought, in German) is not widespread in the news or is common knowledge of German society. In fact, of the more than 260 million articles in the *genios.de* newspaper aggregator database, only 21 mentioned this term, all after 2023. Comparing the interest for searches using the keywords *Blitzdürre* and *Dürre* (the term for conventional drought, in German), we find that interest in flash droughts is, on average, 14 times lower (Fig. 5). Even during peak periods of the interest for flash drought in the early 2000s, this accounted for only 16% of the searches using the keyword drought. In the last 5 years, searches focusing on flash drought dwindled to as low as 0.5%. Due to the absence of the term flash drought from public discourse, the keywords used to assess impact perception from news articles and public awareness from Google searches focused solely on drought-related terms. Nevertheless, there is a common pattern of flash drought emergence and an increase in drought-related report impact on the media illustrated in Figure 6.

H3 - Relationship between flash drought occurrence, perceived impacts, and public awareness

Comparing the time series of flash drought indices and reported impacts in newspapers, we find that following a rapid increase in areas affected by flash drought, there is a wave of impact reporting. By using the cross-correlation functions (ccf), we assess time lags between the occurrence of flash droughts and the reporting of impacts on the news media (see Fig. 7a).

Results for Germany as a whole show a uniform baseline awareness of droughts (see blue line in Fig. 7a). We interpret this as reflecting the cultural and social perception of droughts, encompassing both the expectation of future events and the recollection of past ones. The baseline value was derived from the ccf computation, where the horizontal axis can be arbitrarily long in both the positive and negative directions. To assess the baseline awareness of drought impacts in the media, we used the ccf values for long delays (above 15 weeks, a period longer than usual flash droughts) and negative



(a) (b) (c)

Figure 4: Trends in *(a)*: intensity (drop in soil moisture percentiles per day during onset phase), *(b)*: duration (total event duration in days), and *(c)*: frequency (years with events) of occurrence of flash droughts in Germany at NUTS-2 level from 1980 to 2022. Most regions presented statistically significant trends in all variables.

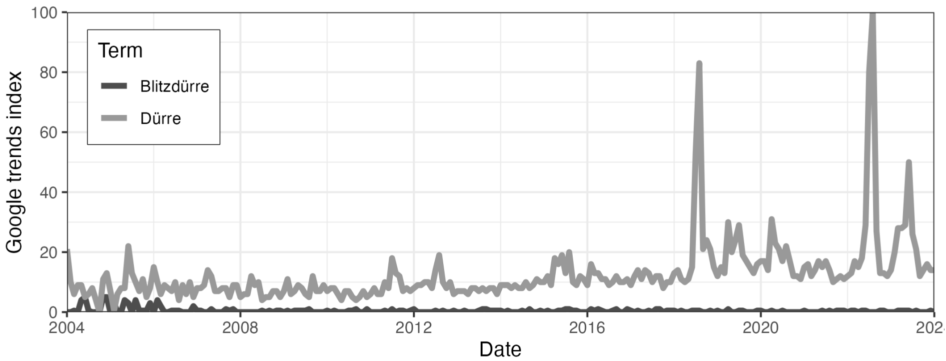


Figure 5: Comparison of searches on Google for the terms "Dürre" (light grey) and "Blitzdürre" (dark grey) in Germany, from 2004 to 2022. Searches in Google Trends are reported as a percentage of the maximum observed in the monitored period.

delays under two weeks (to disregard potential forecast reports). The concept of negative delay results from the mathematical nature of cross-correlation functions that allow the identification of synchronicity independent of causality.

When examining the relationship between the media-reported impacts and flash drought hazards, we found that there is a consistent delay of 4 weeks before impacts peak. This suggests a persistence of interest and reported impacts long after the onset of flash droughts. The 4-week delay can be interpreted as the average time between flash drought initiation and the perception and reporting of socioeconomic impacts. However, there are spatial differences between the changes in impact reporting under flash drought occurrence. In NUTS-2 and NUTS-3 regions, there is a notably higher frequency of impact reporting within the first 2 weeks compared to NUTS-1 regions (see supplementary material). This highlights the effectiveness of localized news sources in monitoring and reporting on the local conditions and impacts of droughts and flash droughts.

Using the online search behavior data from Google Trends, we find similar trends concerning public awareness in response to the occurrence of flash droughts (see Fig. 7b). Overall, the value of the baseline awareness aligns with our previous observation (Fig. 7a). However, the peak delay is shorter (3 weeks) compared to the response of reported impacts in newspaper articles (4 weeks). This shorter timeframe indicates that the general community engaging in online searches has a quicker perception of the impacts of flash droughts. One possible explanation for this is the time required between perception and reporting on the news, as there is an intrinsic interval necessary to collect, verify, and write the news compared to online search behavior.

4. Discussion

Flash droughts are becoming increasingly frequent under climate change and require different adaptation schemes due to their quick emergence (Yuan et al. 2023; Christian,

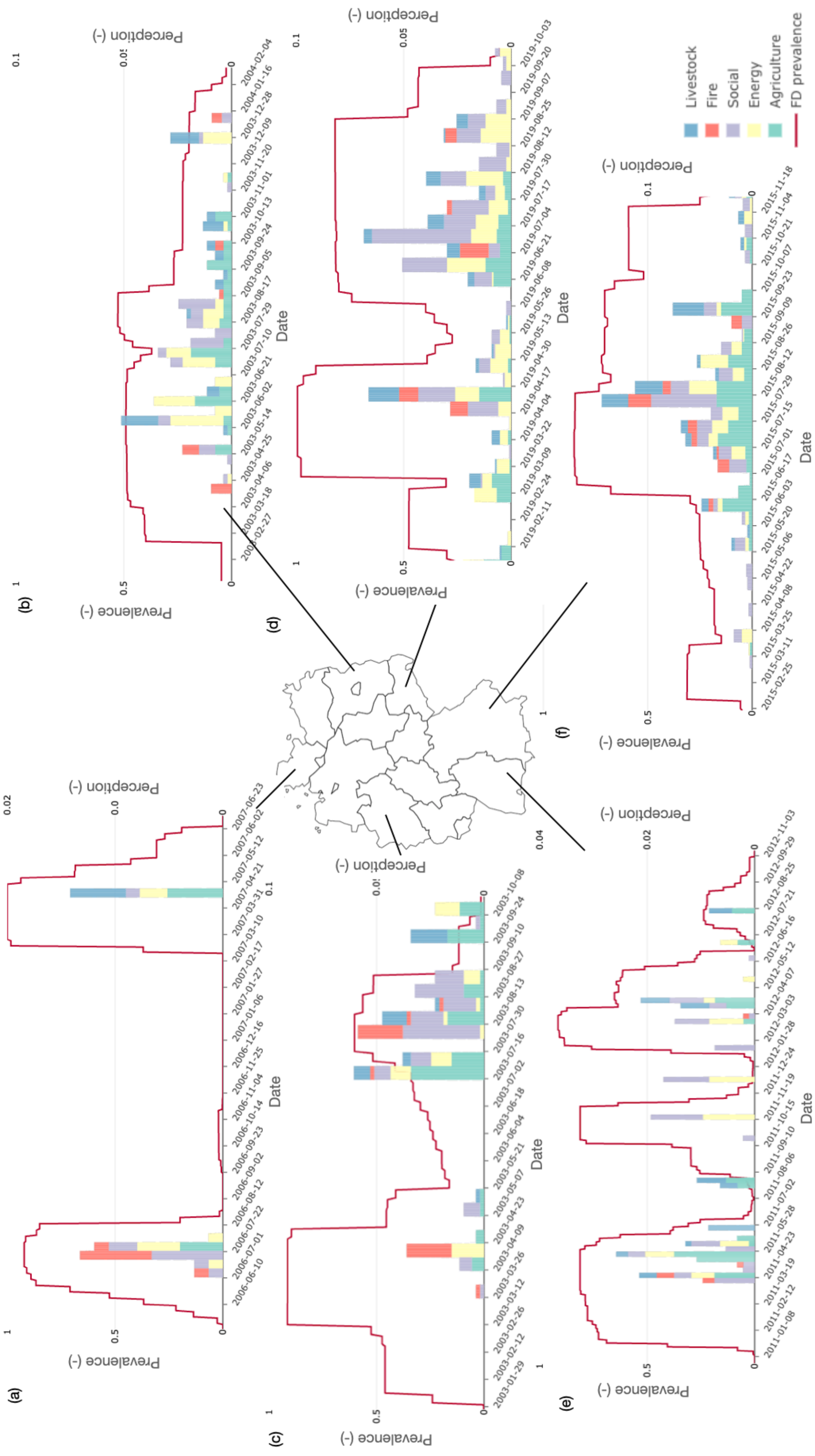


Figure 6: Examples of synchronicity between flash drought occurrence and impact reporting. These examples show the patterns of synchronicity at multiple locations (a: Schleswig-Holstein; b: Brandenburg; c: Nordhein-Westfalen; d: Sachsen; e: Baden-Wüttemberg; f: Bayern). *Prevalence* indicates the percentage of the area affected by flash droughts, and *Perception* the reporting of drought-related impacts according to news articles. For more examples, visit the open application PerceptionFD, used to generate these figures.

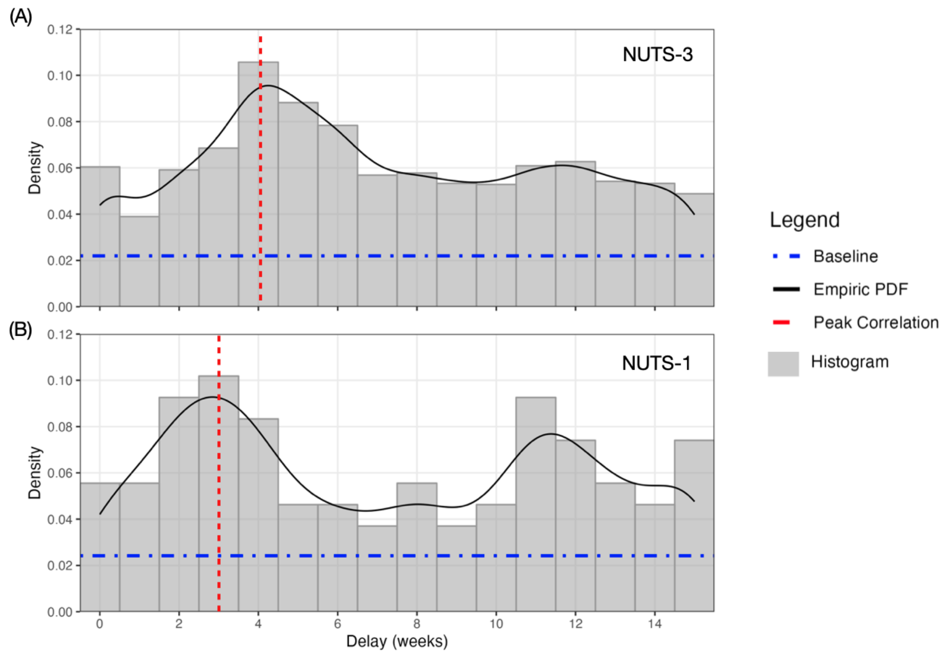


Figure 7: Results of the cross-correlation function between A: Impact reporting in news articles and flash drought occurrences (NUTS-3) and B: searches on Google and flash drought occurrences (NUTS-1). We observe a baseline of correlation, which reflects noise and pre-existing awareness or discussions regarding drought impacts in the media. The red line indicates the peak in the delay of flash drought and impact reporting. Graphs for all NUTS levels can be found in supplementary material.

Hobbins, et al. 2024). Against this background, we investigated the occurrence of flash droughts in Germany along with their societal awareness (Kam, Stowers, and S. Kim 2019) and perceived impacts (Sodoge et al. 2023). Using these digital trace data, our study shows the delayed response of social perception to flash droughts as reported in the media and according to online search behavior. Expanding on the results of our study, we now contextualize our findings within the structure of our hypothesis presented in the Introduction.

Hypothesis 1 - Flash drought is a common phenomenon, and its occurrence has increased over the last few decades in Germany

Our analysis revealed that flash droughts have become increasingly frequent in Germany, occurring once every two years on average (Figure 2). Notably, they are more prevalent in the southern and western portions of the country. In contrast, northeast Germany (e.g. Berlin and Brandenburg) with sandier soils and lower plant-available water (Kruse 2015; Gebauer et al. 2022) exhibit faster variations of soil moisture in terms of percentiles (Alencar and Paton 2022). This characteristic leads to a lower rate of flash drought intensification, which might justify the lower occurrence of flash droughts in the region. It is important to note that this does not necessarily imply that regions like Brandenburg experience less rapid drying such as flash droughts (Shah et al. 2022). Rather, it suggests a limitation in the chosen method in identifying these local events (Alencar and Paton 2022). Impact-based and multivariate indices and analysis have been shown as possible solutions to overcome the limitations of more traditional methods of flash drought identification (Shyrokaya et al. 2023; Mukherjee and Mishra 2022; Zhang et al. 2022).

Furthermore, the trend analysis showed that, beyond being a common phenomenon, flash droughts are becoming more intense, longer, and more frequent across Germany, with positive trends in all three variables observed in all regions (NUTS-2). These trends, however, are not uniform and can be linked to different factors. Specifically, flash drought intensity has increased in the southern region, agreeing with previous reports (Brunner et al. 2023; Yuan et al. 2023), in mountainous regions. Results also indicate that western and southern regions in Germany have experienced a rise in the duration of flash droughts. These variations in trends might be a consequence of differences in regional climate. The western and southern regions are under a temperate oceanic climate (Cfb - Köppen-Geiger classification; Beck et al. 2018) while the eastern regions are under a colder humid continental climate (Dfb). Moreover, east Germany also has sandier soils with lower organic matter content and plant-available water (Kruse 2015; Gebauer et al. 2022). Nevertheless, the exact cause of different trends in flash drought features is beyond the scope of this paper, and more investigation is required.

Hypothesis 2 – The impacts reported in the media and searched by internet users are a source of risk assessment that can be leveraged to identify when flash droughts impact society

While prior studies focused on hazard-centric indicators (Ford and Labosier 2017; Pendergrass et al. 2020; Noguera, Domínguez-Castro, and Vicente-Serrano 2021; Ho, Buras, and Tuo 2023) to identify flash droughts, our research demonstrates the effectiveness of utilizing digital trace data to track both flash droughts and their societal impacts.

We found the use of the term "flash drought" is not widespread in the German public discourse. Consequently, terms associated with drought were employed to evaluate impact perception and public awareness. It is noteworthy that, despite the distinct nature and drivers of droughts and flash droughts, they are often perceived as a single phenomenon, as evidenced by the positive cross-correlation values (Figure 7). Additionally, negative correlations between event occurrence and impact perception and public awareness may be attributed to the semantic blending of flash and conventional droughts in public discourse. This blending results from impacts related to drought that are examined for synchronicity with flash drought events, as well as flash droughts that eventually become longer (conventional) droughts (Basara et al. 2019). Additionally, while previous research utilized Google Trends and newspaper data to enhance the modeling of social awareness in response to droughts (Kam, Stowers, and S. Kim 2019;

Yesuel Kim and Youngchul Kim 2023), this study pioneers the modeling of flash droughts using innovative digital trace data and relating these results with the hazard occurrence. Thus, while the modeling for larger, less-fine-grained temporal resolutions of heatwaves has been shown in previous works (Polt et al. 2023), we thereby highlight how also shorter drought hazards – e.g. flash droughts – can be monitored for very short-term responses. These results underscore the value of digital trace data in comprehending the impacts of flash droughts, unraveling the distinct characteristics that differentiate them from more gradually unfolding droughts.

Hypothesis 3 – There is a delay between flash drought initiation and impact reporting due to latency time for water/plant/soil/society response to dryer conditions

Using high-resolution soil moisture data and digital trace data at multiple levels of spatial aggregation (NUTS1, 2, and 3), we found a correlation between flash drought occurrence and impacts reporting ($r = \tilde{0}.65$) and Google searches ($r = \tilde{0}.61$). Analysis of these digital trace data revealed the existence of a consistent delay between flash drought occurrence and the reporting of impacts on the media (4 weeks) as well as to searches on Google (3 weeks). Hence, these correlations and observed delays prove our hypothesis H3 and further prove hypothesis H2. These delays are, however, too long to be actionable, since they are longer than the usual onset duration for flash droughts (up to 20 days – Ford and Labosier 2017; Pendergrass et al. 2020; Lisonbee, Woloszyn, and Skumanich 2021). The consequences of this finding are two-fold. First, it reinforces the necessity of early-warning and early-action systems to inform stakeholders about the occurrence of flash droughts (Sehgal, Gaur, and Mohanty 2021; Otkin, Woloszyn, et al. 2022). Second, it provides a consistent metric to identify impactful flash droughts, that can be used to validate hazard identification methods, monitoring, and early warning systems (Alencar and Paton 2022).

While multiple attempts at flash drought monitoring and forecasting can be found in recent literature (Mukherjee and Mishra 2022; Ho, Buras, and Tuo 2023; Christian, Hobbins, et al. 2024), these efforts often lack the capability to identify impactful events and validate their metrics and indices with observed societal consequences, leading to misidentification and misinformation (Alencar and Paton 2022). This study provides a novel technique to bridge this gap and allow the production of more reliable methods for flash drought identification and monitoring based on validated events. Currently, Germany does not monitor flash droughts, despite the German drought monitor providing daily assessments of drought conditions (Zink et al. 2016). Moreover, there is no tool in place for alert or early warning, despite the frequent occurrence of flash droughts and their significant impacts on multiple sectors, particularly agriculture.

5. Conclusions

Flash droughts are becoming increasingly frequent under climate change and require different adaptation strategies due to their quick emergence. Against this background, we investigated the occurrence of flash droughts and their relationship with societal awareness in Germany. By analyzing digital trace data on the perception of drought impacts, our study shows the delayed response of social perception to flash droughts both in newspaper media as well as in online searches.

We found that flash drought occurrence in Germany has been increasing over the past decades, making flash droughts a common phenomenon. Synchronicities were observed between the reporting of drought-related impacts in newspaper media and the incidence of flash droughts in Germany. Yet the societal awareness, measured by online search behavior and reported impacts in newspapers, is dampened by a delay of, on average, four weeks compared to the emergence of the flash drought. These findings carry two key implications: (1) Digital trance data coupled with high-quality soil moisture data, can successfully assist modelers in accurately validating designed metrics for flash drought identification, (2) While digital trace data serves as a rich source of information to pinpoint when and where impactful flash droughts happen, it alone is insufficient to provide actionable information, due to inheriting delays in the publication of data. With the increasing threat posed by flash droughts in a changing climate, it is urgent to design information systems capable of integrating forecast assessments with potential impacts in vulnerable areas.

Acknowledgements

This research was funded through the Einstein Research Unit "Climate and Water under Change" from the Einstein Foundation Berlin and Berlin University Alliance (ERU-2020-609).

Declarations

- The authors are members of the Einstein Research Unit "Climate and Water under Change". The Einstein Research Unit "Climate and Water under Change" is funded by the Einstein Stiftung Berlin and the Berlin University Alliance
- Conflict of interest/Competing interests: The authors have no conflict of interests
- Code and Data availability: All code and data are available at https://github.com/pedroalencar1/FlashDroguht_NLP
- Authors' contributions: PHLA contributed with conceptualization, data acquisition, processing, and analysis, script and manuscript writing. JS contributed with data acquisition, processing, and analysis, and manuscript writing. ENP contributed with results interpretation and manuscript writing. MMdeB contributed with conceptualization, results interpretation and manuscript writing.

References

Alencar, Pedro Henrique Lima and Eva Nora Paton (Aug. 2022). "How do we identify flash droughts? A case study in Central European Croplands". In: *Hydrology Research* 53.9, pp. 1150–1165. DOI: 10.2166/nh.2022.003.

Basara, Jeffrey B. et al. (July 2019). "The evolution, propagation, and spread of flash drought in the Central United States during 2012". In: *Environmental Research Letters* 14.8, p. 084025. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ab2cc0.

Beck, Hylke E. et al. (Oct. 2018). "Present and future Köppen-Geiger climate classification maps at 1-km resolution". In: *Scientific Data* 5.1. ISSN: 2052-4463. DOI: 10.1038/sdata.2018.214.

Boeing, Friedrich et al. (Oct. 2022). "High-resolution drought simulations and comparison to soil moisture observations in Germany". In: *Hydrology and Earth System Sciences* 26.19, pp. 5137–5161. DOI: 10.5194/hess-26-5137-2022.

Bogdanovich, Ekaterina et al. (Jan. 2024). "Official heat warnings miss situations with a detectable societal heat response in European countries". In: *International Journal of Disaster Risk Reduction* 100, p. 104206. ISSN: 2212-4209. DOI: 10.1016/j.ijdrr.2023.104206.

Bogena, Heye Reemt et al. (Mar. 2022). "COSMOS-Europe: a European network of cosmic-ray neutron soil moisture sensors". In: *Earth System Science Data* 14.3, pp. 1125–1151. ISSN: 1866-3516. DOI: 10.5194/essd-14-1125-2022.

Brunner, Manuela I. et al. (Jan. 2023). "Hydrological Drought Generation Processes and Severity Are Changing in the Alps". In: *Geophysical Research Letters* 50.2. ISSN: 1944-8007. DOI: 10.1029/2022gl101776.

Christian, J.I., Jeffrey B. Basara, E.D. Hunt, et al. (Sept. 2020). "Flash drought development and cascading impacts associated with the 2010 Russian heatwave". In: *Environmental Research Letters* 15.9, p. 094078. DOI: 10.1088/1748-9326/ab9faf.

Christian, J.I., Jeffrey B. Basara, Jason A. Otkin, et al. (May 2019). "A Methodology for Flash Drought Identification: Application of Flash Drought Frequency across the United States". In: *Journal of Hydrometeorology* 20.5, pp. 833–846. DOI: 10.1175/jhm-d-18-0198.1.

Christian, J.I., M. Hobbins, et al. (Jan. 2024). "Flash drought: A state of the science review". In: *WIREs Water*. ISSN: 2049-1948. DOI: 10.1002/wat2.1714.

Ford, Trent W. and Christopher F. Labosier (Dec. 2017). "Meteorological conditions associated with the onset of flash drought in the Eastern United States". In: *Agricultural and Forest Meteorology* 247, pp. 414–423. DOI: 10.1016/j.agrformet.2017.08.031.

Gebauer, Anika et al. (Jan. 2022). "Topsoil Texture Regionalization for Agricultural Soils in Germany—An Iterative Approach to Advance Model Interpretation". In: *Frontiers in Soil Science* 1. ISSN: 2673-8619. DOI: 10.3389/fsoil.2021.770326.

Helsel, Dennis R. et al. (2020). *Statistical methods in water resources*. DOI: 10.3133/tm4a3.

Ho, Sarah, Allan Buras, and Ye Tuo (Nov. 2023). "Comparing Agriculture-Related Characteristics of Flash and Normal Drought Reveals Heterogeneous Crop Response". In: *Water Resources Research* 59.11. ISSN: 1944-7973. DOI: 10.1029/2023wr034994.

Hogue, J and B DeWilde (2023). *PyTends: pseudo API for Google Trends*. Online. URL: https://pypi.org/project/pytrends/.

Kam, Jonghun, Kimberly Stowers, and S. Kim (Apr. 2019). "Monitoring of Drought Awareness from Google Trends: A Case Study of the 2011–17 California Drought". In: *Weather, Climate, and Society* 11.2, pp. 419–429. ISSN: 1948-8335. DOI: 10.1175/wcas-d-18-0085.1.

Kim, Yesuel and Youngchul Kim (Sept. 2023). "Global regionalization of heat environment quality perception based on K-means clustering and Google trends data". In: *Sustainable Cities and Society* 96, p. 104710. ISSN: 2210-6707. DOI: 10.1016/j.scs.2023.104710.

Kruse, Klaus (2015). *Bodendaten der BGR und der neue Bodenatlas Deutschland*. Tech. rep. Bundesanstalt für Geowissenschaften und Rohstoffe.

Lesinger, Kyle and Di Tian (Sept. 2022). "Trends, Variability, and Drivers of Flash Droughts in the Contiguous United States". In: *Water Resources Research* 58.9. ISSN: 1944-7973. DOI: 10.1029/2022wr032186.

Li, Jun et al. (Nov. 2020). "A new framework for tracking flash drought events in space and time". In: *CATENA* 194, p. 104763. ISSN: 0341-8162. DOI: 10.1016/j.catena.2020.104763.

Liang, Miaoling and Xing Yuan (2021). "Critical role of soil moisture memory in predicting 2012 central USA flash drought". In: *Frontiers in Earth Science* 9, p. 46.

Lisonbee, Joel, Molly Woloszyn, and Marina Skumanich (Feb. 2021). "Making sense of flash drought: definitions, indicators, and where we go from here". In: *Journal of Applied and Service Climatology* 2021.1, pp. 1–19. DOI: 10.46275/joasc.2021.02.001.

Madruga de Brito, Mariana, Christian Kuhlicke, and Andreas Marx (Oct. 2020). "Near-real-time drought impact assessment: a text mining approach on the 2018/19 drought in Germany". In: *Environmental Research Letters* 15.10, 1040a9. ISSN: 1748-9326. DOI: 10.1088/1748-9326/aba4ca.

Mann, Henry B. (July 1945). "Nonparametric Tests Against Trend". In: *Econometrica* 13.3, p. 245. ISSN: 0012-9682. DOI: 10.2307/1907187.

Mo, Kingtse C. and Dennis P. Lettenmaier (Apr. 2016). "Precipitation Deficit Flash Droughts over the United States". In: *Journal of Hydrometeorology* 17.4, pp. 1169–1184. DOI: 10.1175/jhm-d-15-0158.1.

Mukherjee, Sourav and Ashok Kumar Mishra (Jan. 2022). "A Multivariate Flash Drought Indicator for Identifying Global Hotspots and Associated Climate Controls". In: *Geophysical Research Letters* 49.2. ISSN: 1944-8007. DOI: 10.1029/2021gl096804.

Mullen, Lincoln (2020). *textreuse: Detect Text Reuse and Document Similarity.* https://docs.ropensci.org/textreuse, https://github.com/ropensci/textreuse.

Noguera, Iván, Fernando Domínguez-Castro, and Sergio M. Vicente-Serrano (Jan. 2021). "Flash Drought Response to Precipitation and Atmospheric Evaporative Demand in Spain". In: *Atmosphere* 12.2, p. 165. DOI: 10.3390/atmos12020165.

Osman, Mahmoud et al. (Feb. 2021). "Flash drought onset over the contiguous United States: sensitivity of inventories and trends to quantitative definitions". In: *Hydrology and Earth System Sciences* 25.2, pp. 565–581. DOI: 10.5194/hess-25-565-2021.

Otkin, Jason A., Mark Svoboda, et al. (May 2018). "Flash Droughts: A Review and Assessment of the Challenges Imposed by Rapid-Onset Droughts in the United States". In: *Bulletin of the American Meteorological Society* 99.5, pp. 911–919. DOI: 10.1175/bams-d-17-0149.1.

Otkin, Jason A., Molly Woloszyn, et al. (Oct. 2022). "Getting ahead of Flash Drought: From Early Warning to Early Action". In: *Bulletin of the American Meteorological Society* 103.10, E2188–E2202. ISSN: 1520-0477. DOI: 10.1175/bams-d-21-0288.1.

Pastorello, Gilberto et al. (July 2020). "The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data". In: *Scientific Data* 7.1, p. 225. ISSN: 2052-4463. DOI: 10.1038/s41597-020-0534-3. URL: https://doi.org/10.1038/s41597-020-0534-3.

Pendergrass, Angeline G. et al. (Mar. 2020). "Flash droughts present a new challenge for subseasonal-to-seasonal prediction". In: *Nature Climate Change* 10.3, pp. 191–199. DOI: 10.1038/s41558-020-0709-0.

Polt, Kelley De et al. (Sept. 2023). "Quantifying impact-relevant heatwave durations". In: *Environmental Research Letters* 18.10, p. 104005. ISSN: 1748-9326. DOI: 10.1088/1748-9326/acf05e.

Rakovec, Oldrich et al. (Mar. 2022). "The 2018–2020 Multi-Year Drought Sets a New Benchmark in Europe". In: *Earth's Future* 10.3. ISSN: 2328-4277. DOI: 10.1029/2021ef002394.

Samaniego, Luis et al. (Sept. 2017). "Toward seamless hydrologic predictions across spatial scales". In: *Hydrology and Earth System Sciences* 21.9, pp. 4323–4346. ISSN: 1607-7938. DOI: 10.5194/hess-21-4323-2017.

Sehgal, Vinit, Nandita Gaur, and Binayak P. Mohanty (Aug. 2021). "Global Flash Drought Monitoring Using Surface Soil Moisture". In: *Water Resources Research* 57.9. ISSN: 1944-7973. DOI: 10.1029/2021wr029901.

Sen, Pranab Kumar (Dec. 1968). "Estimates of the Regression Coefficient Based on Kendall's Tau". In: *Journal of the American Statistical Association* 63.324, pp. 1379–1389. ISSN: 1537-274X. DOI: 10.1080/01621459.1968.10480934.

Shah, Jignesh et al. (May 2022). "Increasing footprint of climate warming on flash droughts occurrence in Europe". In: *Environmental Research Letters* 17.6, p. 064017. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ac6888.

Shumway, Robert H. and David S. Stoffer (May 2019). *Time Series: A Data Analysis Approach Using R*. Chapman and Hall/CRC. ISBN: 9780429273285. DOI: 10.1201/9780429273285.

Shyrokaya, Anastasiya et al. (Oct. 2023). "Advances and gaps in the science and practice of impact-based forecasting of droughts". In: *WIREs Water*. ISSN: 2049-1948. DOI: 10.1002/wat2.1698.

Sodoge, Jan et al. (Dec. 2023). "Text-mining uncovers the unique dynamics of socio-economic impacts during multi-year drought". In: DOI: 10.5194/nhess-2023-228.

Theil, Henri (1992). "A Rank-Invariant Method of Linear and Polynomial Regression Analysis". In: *Henri Theil's Contributions to Economics and Econometrics*. Springer Netherlands, pp. 345–381. ISBN: 9789401125468. DOI: 10.1007/978-94-011-2546-8_20.

Tijdeman, Erik and Lucas Menzel (Apr. 2021). "The development and persistence of soil moisture stress during drought across southwestern Germany". In: *Hydrology and Earth System Sciences* 25.4, pp. 2009–2025. DOI: 10.5194/hess-25-2009-2021.

Walker, David W. et al. (Jan. 2024). "Flash Drought Typologies and Societal Impacts: A Worldwide Review of Occurrence, Nomenclature, and Experiences of Local Populations". In: *Weather, Climate, and Society* 16.1, pp. 3–28. ISSN: 1948-8335. DOI: 10.1175/wcas-d-23-0015.1.

Wang, Xueqin and Qiqing Yu (Sept. 2005). "Unbiasedness of the Theil–Sen estimator". In: *Journal of Nonparametric Statistics* 17.6, pp. 685–695. ISSN: 1029-0311. DOI: 10.1080/10485250500039452.

Yuan, Xing et al. (Apr. 2023). "A global transition to flash droughts under climate change". In: *Science* 380.6641, pp. 187–191. DOI: 10.1126/science.abn6301.

Zhang, Yu et al. (May 2022). "A new multi-variable integrated framework for identifying flash drought in the Loess Plateau and Qinling Mountains regions of China". In: *Agricultural Water Management* 265, p. 107544. ISSN: 0378-3774. DOI: 10.1016/j.agwat.2022.107544.

Zink, Matthias et al. (July 2016). "The German drought monitor". In: *Environmental Research Letters* 11.7, p. 074002. ISSN: 1748-9326. DOI: 10.1088/1748-9326/11/7/074002.