



OPEN Relative engagement with sources of climate misinformation is growing across social media platforms

Saverio Storani¹, Max Falkenberg^{2,3}, Walter Quattrociocchi¹ & Matteo Cinelli¹✉

We explore the discourse on climate change across multiple social media platforms, examining the evolution of user engagement with climate-related content and whether this content links to reliable or unreliable news media sources. Through a detailed examination of over 20 million posts on Facebook, Instagram, Twitter, and YouTube over five years (2018–2022), we identify trends in engagement, distinguishing between unreliable and reliable content to assess the impact of misinformation. Further, we investigate the relationships among various discussion topics and their association with information quality, employing a network-based method to quantify the semantic distances between these categories. Our findings reveal diverse trends in engagement that align with global events, suggesting that social media discussions promptly reflect the resonance of real life events concerning climate change such as COP26, the Climate Action Week and climate strikes associated with the Fridays for Future movement. Notably, despite the lower volume of content linking to unreliable sources, we observe significantly greater relative engagement with these sources compared to content from reliable sources on all platforms except Twitter. This highlights a persistent challenge in the online discourse surrounding climate misinformation.

In the contemporary digital environment, social media platforms critically shape public discourse on crucial issues like climate change^{1–3}. Users often select content that aligns with their existing beliefs, typically overlooking opposing viewpoints, fostering echo chambers characterised by homogeneous, like-minded groups^{4–6}. This phenomenon is partly rooted in the core designs of these platforms and user psychology^{7–9}, where the business models of social media companies aim to captivate and retain attention by curating engaging content, thereby intensifying the echo chamber effect¹⁰. This dynamic contributes to increased ideological segregation, influenced by algorithmic curation and user interactions^{11,12}. The challenge of separating inherent human behaviours from digital influences is exacerbated by the limited data available for research¹³. Even when access to data is granted by direct holders such as Meta¹⁴, the intertwining of user interactions with this data makes it challenging to separate natural behaviours from those influenced by the platform's algorithms and design features.

In this study, we focus specifically on how climate change (and climate misinformation) is communicated on social media for two reasons. Firstly, climate change, unlike many other socio-political issues, presents an existential risk to human societies and requires coordinate global action above and beyond what is needed to tackle most other socio-political issues. Achieving this global coordination is particularly challenging because progress in the fight against climate change will only be apparent over a relatively long time period, a timescale which is not compatible with the short-term priorities (often economic) of most policy makers. This sets it apart from other major socio-political issues where the costs and consequences of policy decisions are much more evident in the short term (e.g., the rapid development of Covid-19 vaccines as a result of concerted policy efforts). Secondly, the failure to tackle climate change is primarily a policy failure, not a failure of climate science, since there remains a lack of global consensus as to which policies are correct or necessary to prevent worsening climate change. This lack of policy consensus is what a lot of climate sceptics exploit when they oppose climate action, and counteracting these points requires an understanding of how and why these narratives are appealing to specific political groups and the general public. Building this understanding starts by studying in detail how the issue of climate change is communicated across different contexts and domains, an effort which our study

¹Department of Computer Science, Sapienza University of Rome, Via Regina Elena 295, Rome 00161, Italy.

²Department of Mathematics, City University of London, Northampton Square, London EC1V 0HB, United Kingdom. ³Department of Network & Data Science, Central European University, Vienna, Austria. ✉email: matteo.cinelli@uniroma1.it

aims to contribute to. Many climate sceptic narratives are spread via social media, which is why our study focuses specifically on social media discourses.

While broad scientific consensus recognises human activity as the primary driver of climate change¹⁵, public discourse on social media allows for the (near) unrestricted exchange of views, including views sceptical of climate science which conflict with the scientific consensus, potentially undermining public understanding of climate-related global issues. While this openness introduces a rich diversity of perspectives¹⁶, it also risks generating confusion and discord, especially when scientifically unsupported opinions receive as much emphasis as facts, thereby fostering polarisation^{1,17,18}. Bridging the gap between scientific consensus and diverse public opinions is essential; effective communication and education are vital to achieving this¹⁹. Although traditional media has significantly shaped public opinion in the past²⁰, social media has become a formidable force in disseminating information and transforming the public discourse on climate change through dynamic sharing and the debate of ideas and evidence.

Our study conducts a comparative analysis across multiple social media platforms to address the outlined challenges, identifying consistent themes in the climate change discourse over five years. We examine interactions with content on Facebook, Instagram, Twitter, and YouTube to discern nuances in user engagement, the evolution of discussion themes, and the influence of sources that frequently share unreliable content on the online debate.

As stated in the 2024 Global Risk Report of the World Economic Forum, misinformation is among the main short term global risks²¹. Climate change and misinformation are, unfortunately, tightly linked and according to the scientific literature, a wide share of misinformation narratives revolve around this topic²² and some of its manifestations (e.g. extreme climate events^{23,24}). Despite the longstanding history of misinformation in political processes, its proliferation in recent years, particularly on social media, has raised significant concerns^{25,26}. Misinformation has been prominently analysed in contexts ranging from elections to natural disasters and climate change, with the COVID-19 pandemic further highlighting its impact^{18,23,27–30}. Our study builds on and advances a growing literature studying the narratives of climate change on social media^{24,31–39}. Within this literature, some studies focus on the discussion around climate change in general, whereas others focus on specific aspects or themes within the wider climate discussion. For example, studies have considered online discussions around climate science^{31–33,35}, climate politics^{1,36}, and climate activism^{40,41}. Across these studies, common findings include (1) that the online discussion around climate change is polarised between groups who are supportive of climate action, and those sceptical of action^{1,3,31,32}, and (2) that engagement with climate narratives has consistently increased over time¹, particularly since the climate protests in 2019 and the emergence of the Fridays for Future movement⁴¹. Studies have also considered individuals' engagement with, and discussion of, pro-climate action behaviours or policies, for instance in relation to veganism⁴², the adoption of electric vehicles⁴³, and concerning controversial actions such as the climate geoengineering⁴⁴.

Historically, the literature on social media and climate change has been heavily biased towards studies focusing on Twitter (now X)², with only a small number of studies considering climate communication on other platforms, for instance on Facebook⁴⁵, YouTube⁴⁶, and more recently on TikTok⁴⁷. Consequently, our understanding of how climate change is communicated online has suffered from a bias in the audience and populations studied. However, perhaps more importantly, the vast majority of studies do not compare findings across platforms (for a rare comparative study, see³⁹). Therefore, it is not clear the extent to which insights can be generalised to social media as a whole, and by extension to the general public. Addressing this gap is a key aim of our study.

In this work, we analyse the dynamics of user engagement with climate change content across various social media platforms, examining both the volume and the nature of engagement, the persistence of specific topics over time, and users engagement with topics spread by unreliable and reliable information sources. We identify trends of engagement that align with significant global events (such as the Climate Action Week or the COP summits), highlighting the pivotal role of social media in raising awareness about environmental issues. Furthermore, the increasing engagement with sources sharing unreliable content across all four platforms underscores the urgent need for strategies aimed at improving the accuracy and quality of online discourse related to climate change.

Results and discussion

Interaction patterns

We gathered data using a keyword search of terms central to the climate change debate, including {*climate change*, *climate crisis*, *climate emergency*, *global warming*} thus focusing on English language content. A more comprehensive explanation regarding the data collection is reported in Methods (see Data collection Section) while a breakdown of the collected data is presented in Table 1.

We start by investigating posting patterns on the four social media platforms⁷, namely Facebook (FB), Instagram (IG), Twitter (TW) and YouTube (YT) obtaining the heavy-tailed distributions of users' activity (i.e. number of posts per account) and engagement per post (i.e. the sum of the interactions received by each post) shown in Fig. 1. The shape of the distributions (displayed on logarithmic axes and common to most social media platforms⁷) shows that while the majority of accounts publish only a small number of posts, a small number of accounts are responsible for a large number of the posts produced. Similar patterns are observed in the case of engagement per post where a small fraction of posts receive the majority of engagement.

Trends and volumes of social media content

In this section, we show the evolution in terms of the volume of content posted on each of the social media platforms studied in order to estimate broad trends in the wider climate discussion.

In the first column of Fig. 2, we display the time series for the quantity of content created (i.e., posts on Facebook and Instagram, tweets on Twitter, and videos on YouTube) on the four platforms. The second column displays their total engagement, i.e. the sum of the interactions received by each post; interactions are defined

	Posts	Accounts	Period
Facebook	8,154,674	765,442	2018/01/01–2022/12/31
Instagram	653,447	151,972	2018/01/01–2022/12/31
Twitter	12,718,683	2,595,230	2018/01/01–2021/12/31
YouTube	38,930	9,067	2018/01/01–2022/12/31
Total	21,565,734	3,521,775	

Table 1. Number of posts, accounts and time window of analysis for Facebook, Instagram, Twitter, and YouTube. The term accounts refer to pages and verified profiles on Facebook and Instagram, users on Twitter and channels on YouTube. The time window is one year shorter on Twitter due to unavailability of Twitter API at the time of data collection.

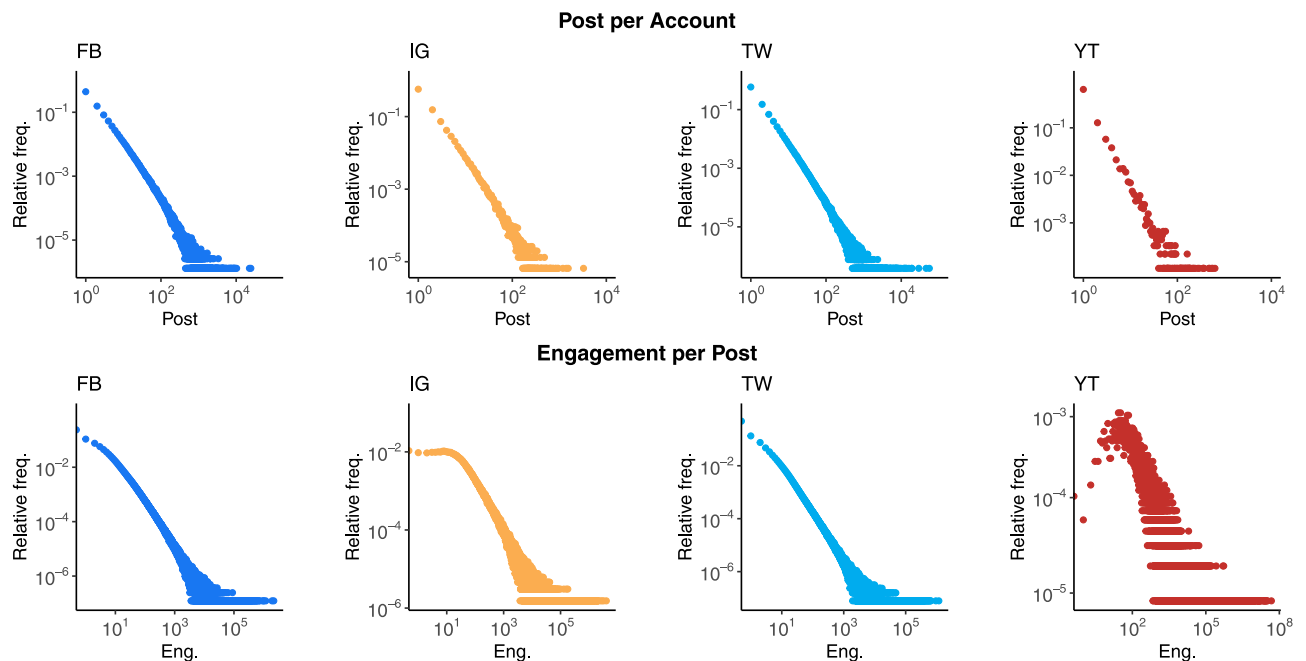


Fig. 1. Posts and engagement distributions shown on a double logarithmic scale. In the first row, we display the distribution of the number of published contents, while in the second row, we show the distribution of engagements for the published contents. In both cases, we present the results for the four analysed platforms.

as likes, reactions, comments and shares on Facebook; likes, comments and shares on Instagram; likes, replies and retweets on Twitter; likes, dislikes, comments, shares and views on YouTube. The time series are shown at a monthly resolution, with the data for each social media platform organised by rows and differentiated by colour coding. Given the non-stationary nature of these time series, we employ the Hodrick-Prescott (HP) filter to extract and summarise their trends, potentially revealing non-linear patterns. In a nutshell, the HP filter decomposes a time series into a smooth trend component and a cyclical component by minimising the squared deviations of the time series values from the smoothed trend while applying a penalty function to large trend's variation. For an in-depth explanation of the HP filter and its application to our data, please refer to Methods (see Hodrick-Prescott filter Section).

Upon inspecting Fig. 2, we note that the time series for each platform exhibit distinct peaks corresponding with major international events, demonstrating a widespread engagement with climate change issues. Notably, a significant surge in content creation and user interactions is observed across all platforms around the end of 2019. This surge aligns with critical events: the Climate Action Week starting September 27, 2019, that involved around 7.6 million participants globally, and the climate strikes on November 29, 2019, associated with the Fridays for Future movement⁴¹. Another notable peak in both activity and engagement occurs in November 2021, in coincidence with the 26th Conference of Parties (COP26), highlighting this event's substantial impact on raising awareness and fostering discussion on climate change issues within the digital sphere^{1,40}. To assess whether these events influenced engagement trends, we analysed the average engagement levels before and during the considered events. As shown in Table 2, there is a notable increase in engagement across platforms around these key moments. During Climate Action Week in September 2019, Facebook engagement rose significantly—from 16,735,430 before the event to 38,987,139 during the event. Instagram saw a similarly sharp increase, with engagement jumping from 24,832,920 to 54,882,392. Twitter experienced a more moderate rise, from 8,025,537

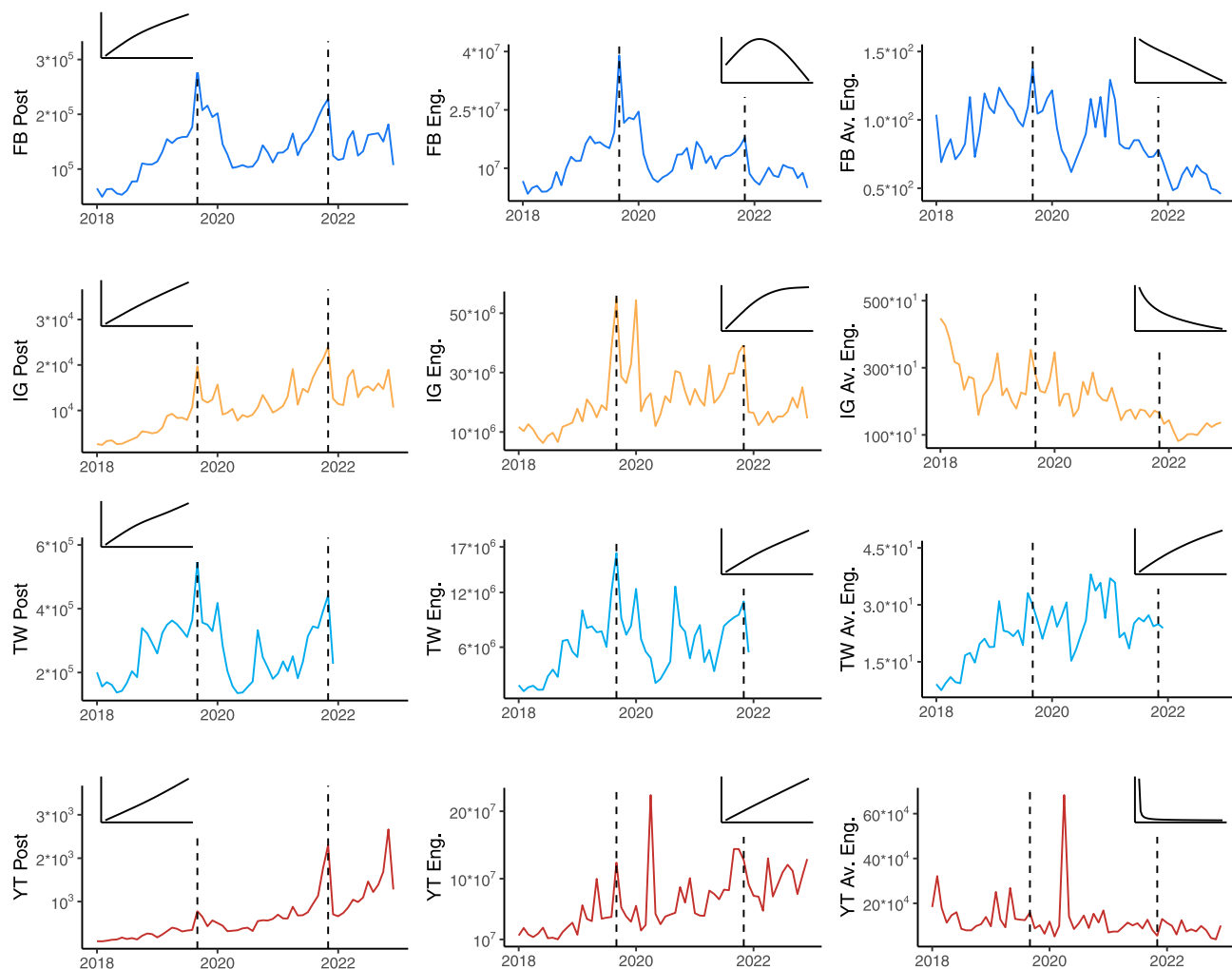


Fig. 2. Trends in the production of, and engagement with, climate related content across four social media platforms. The four social media platforms Facebook (FB), Instagram (IG), Twitter (TW), and YouTube (YT) are colour-coded and represented one per line. Left column: The time series of the number of monthly posts. Middle column: The total engagement (sum of monthly interactions). Right column: The average engagement per post given by the ratio of the quantities in the first and second column. Finally, inset we show the trends estimated using the HP Filter. Dashed lines mark events of particular importance in September 2019 and November 2021: Climate Action Week and the climate strikes associated with the Fridays for Future movement, respectively.

Event month	Platform	Eng. (Event month)	Avg. Eng. (Before)	Avg. Eng. (After)
Sep 2019	FB	38,987,139	16,735,430 (−57%)	22,420,055 (−42%)
	IG	54,882,392	24,832,920 (−55%)	27,548,530 (−50%)
	TW	16,321,216	8,025,537 (−51%)	8,770,479 (−46%)
	YT	123,819,551	41,588,457 (−66%)	44,110,045 (−64%)
Nov 2021	FB	17,829,903	14,322,492 (−20%)	7,061,658 (−60%)
	IG	39,505,208	32,663,626 (−17%)	18,623,670 (−53%)
	TW	11,013,226	9,230,530 (−16%)	6,586,242 (−40%)
	YT	127,572,410	98,092,648 (−23%)	91,732,018 (−28%)

Table 2. Engagement trends on different platforms (FB, IG, TW, YT) around two key climate events. For each event, the engagement values in the event month are compared with the average engagement over the three months preceding and following the event. Percentages are computed with respect to the Event Month.

to 16,321,216, while YouTube saw a prominent surge, increasing from 41,588,457 to 123,819,551. A comparable trend was observed during COP26 in November 2021. Facebook engagement increased from 14,322,492 to 17,829,903. Instagram also saw an increase, from 32,663,626 to 39,505,208. Twitter engagement grew from 9,230,530 to 11,013,226, while YouTube once again experienced a rise, from 98,092,648 to 127,572,410.

To further investigate the relevance of these events we conducted an interrupted time series (ITS) analysis for both the September 2019 and November 2021 events (for details see the Interrupted time series sub-section in Methods). The ITS analyses identify whether events like *Climate Action Week* and *COP26* caused significant changes in engagement trends; and if the effect are temporary or long-term. Table 3 displays the results of the ITS analysis. The significant coefficients for the time variable T and, in particular, its quadratic term T^2 across all platforms indicate that engagement trends over time are nonlinear, suggesting more complex dynamics than a simple linear trend. Specifically, the significance of the quadratic term at the 0.01 level (except for IG and YT at the 0.05 level) highlights acceleration or deceleration patterns over time. The analysis of the event coefficient (D_1 and D_2) reveals that, for many platforms (IG, TW, YT), the event did not produce an immediate and significant impact on engagement. However, this does not imply a complete absence of effect but rather a lack of an abrupt and short-term change. A key finding is the significance of the interaction terms $T \times D_1$ and $T \times D_2$ across all platforms, indicating that the event's effect unfolds over time rather than manifesting immediately. The negative sign of these coefficients suggests a slowdown in the growth rate or a stabilization of engagement levels in the long term. Although Fig. 2 shows visible peaks during the event, the non-significance of the immediate event effect in the model suggests that these variations were temporary and did not lead to a lasting shift in the trajectory of the time series. This highlights how high-profile events may generate visible short-term fluctuations, while their long-term effects become evident only through time-dependent interactions.

However, we would like to clarify that our study does not aim to establish causality between events and platforms trends but rather to highlight the trends observed in the engagement of social media. Many studies provide evidence of the relationship between platforms and user climate awareness, which further supports our observations; see, e.g.^{48–50}. Causality is also difficult to establish because external events such as the Covid-19 pandemic changed social media use trends during the time period we have studied, but unfortunately it is not possible to disentangle how engagement with climate-related events would have changed in the absence of such external factors.

Our analysis reveals diverse trends across social media platforms regarding content publication, engagement, and average engagement. Specifically, we note an overall increase in content creation across all platforms. Regarding total engagement, we note a mix of decreasing trends for Facebook, a plateauing effect on Instagram, and increases on Twitter and YouTube. The most notable differences across social media platforms are observable in the case of average engagement, calculated as the monthly ratio of engagement (i.e. the sum of interactions received by all the posts published in a given month to the sum of monthly posts per platform). Facebook, Instagram, and YouTube exhibit a decline in average engagement, each following a different trajectory (note that in the case of YouTube the same results were obtained also excluding views⁵¹ from the computation of engagement, as shown in Supplementary Information [SI]). Conversely, Twitter shows an increasing trend in

	Dependent variable:			
	Eng			
	(FB)	(IG)	(TW)	(YT)
T	−584,999.600* (312,465.500)	−3,159,103.000 (2,591,998.000)	−93,124.550 (203,032.000)	−3,159,103.000 (2,591,998.000)
T^2	68,171.800*** (13,084.240)	258,622.000** (108,537.800)	27,695.590*** (8,534.626)	258,622.000** (108,537.800)
D_1	5,967,366.000** (2,439,179.000)	18,111,719.000 (20,233,747.000)	324,443.000 (1,567,613.000)	18,111,719.000 (20,233,747.000)
D_2	−2,927,334.000 (2,728,832.000)	−21,720,462.000 (22,636,510.000)	7,515,377.000 (5,642,252.000)	−21,720,462.000 (22,636,510.000)
$T \times D_1$	−4,376,966.000*** (628,057.000)	−12,677,117.000** (5,209,928.000)	−1,831,288.000*** (409,104.300)	−12,677,117.000** (5,209,928.000)
$T \times D_2$	−2,588,012.000*** (590,498.500)	−10,192,103.000** (4,898,369.000)	−6,285,665.000* (3,487,654.000)	−10,192,103.000** (4,898,369.000)
Constant	6,996,407.000*** (2,047,351.000)	27,026,320.000 (16,983,412.000)	2,101,142.000 (1,318,204.000)	27,026,320.000 (16,983,412.000)
Observations	60	60	48	60
R ²	0.653	0.496	0.550	0.496
Adjusted R ²	0.613	0.439	0.484	0.439
F Statistic	16.594***	8.695***	8.347***	8.695***

Table 3. Results of the ITS regression models for platforms engagement (FB, IG, TW, YT). Coefficients are reported with their standard errors in parentheses. Significance levels are indicated as follows: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

average engagement over time. The analysis of average engagement trends reveals a general downtrend across three platforms. Several factors may contribute to this phenomenon including (i) the variability in the volume of posts generated each month; (ii) the significant proportion of accounts posting climate change-related content sporadically (as depicted in Fig. 1, top panel), which could engage a narrowly interested audience, thus diminishing overall average engagement; (iii) the skewed distribution of engagement across all platforms (illustrated in Fig. 1, bottom panel), suggesting that average engagement may not fully capture the nuances of user interaction and that, while many accounts are producing climate content, only few accounts are likely to have a meaningful influence in the debate. Robustness checks to address these complexities and validate the observed trends are presented in SI (section Robustness Checks).

Overall, the evidence suggests that while engagement is likely to be approaching saturation (especially on Facebook and Instagram), content production does not follow this trend. This may indicate that information sources addressing the topic of climate change are encountering difficulties in attracting new individuals to participate in the online climate discussion. Furthermore, the observed trends could be attributed to several factors, including the different demographics of each social media platform⁵², changes in the language used to discuss climate change, or the popularity of new platforms such as TikTok.

To investigate whether trends in average engagement reflect shifts in individual participation in the climate change debate, we analyse accounts' activity patterns. This analysis is essential for evaluating the discussion's ability to attract new participants and includes monthly time series data for both unique and new accounts. Unique accounts are those posting about climate change in a given month, while new accounts are those engaging in the discussion for the first time. Figure 3 displays the observed trends, showing a consistent increase in the number of unique accounts engaged in the climate discussion across all platforms, despite some monthly variations. Even though some data points indicate a decline, the influx of new accounts suggests sustained interest in the topic. The overall growth in terms of unique accounts combined with the presence of new accounts every month indicates that while engagement may be saturating on some platforms, the climate change issue continues to draw new voices, maintaining its relevance and expanding its reach. This being said, it is important to notice the decreasing trends of the new accounts curves on Facebook, Twitter and, to some extent, on Instagram.

Prevalence of content from unreliable sources

In this section, we examine social media engagement with posts referencing URLs to external (i.e. outside of the social media platform) information sources, distinguishing between those frequently spreading misinformation and those deemed reliable. We combine the lists curated by NewsGuard and Media Bias/Fact Check, two independent news rating agencies that assess the credibility and quality of news sources. Our methodology (detailed in Methods, see section Matching sources, and common to a number of previous studies^{18,53}) involves a source-based labelling strategy where posts are categorised as unreliable or reliable based on these assessments. In more detail, unreliable sources are those that have been flagged by NewsGuard and Media Bias/Fact Check for frequently publishing false or misleading information, while reliable sources consistently adhere to journalistic standards and fact-checking protocols. This systematic approach ensures that all content from a flagged source is consistently classified as potentially misleading or trustworthy, aiding in a clearer understanding of misinformation's role in public discourse. Figure 4 presents a comparative analysis of average engagement for content from both sources using a static (left hand side) and dynamic (right hand side) perspective.

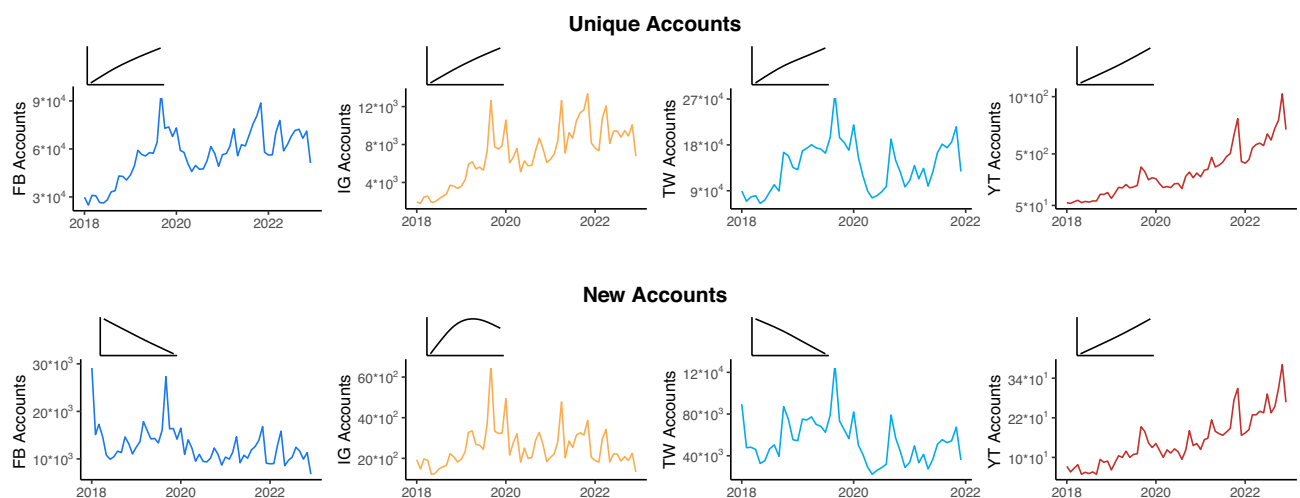


Fig. 3. Unique and new accounts posting about climate change. The top panel displays monthly time series of unique accounts for Facebook (FB), Instagram (IG), Twitter (TW), and YouTube (YT). The bottom panel displays the monthly time series of new accounts for Facebook (FB), Instagram (IG), Twitter (TW), and YouTube (YT). By new accounts we refer to those who have never posted content related to climate change in the previous months. Finally, above each plot, we find the theoretical trends estimated using the Hodrick-Prescott Filter (HP Filter).

The bar charts in the left hand side of Fig. 4 show that content linked to unreliable sources receive, on average, higher engagement than content linked to reliable sources on all platforms except Twitter. Possible reasons behind this observation may include the fact that posts relying on unreliable sources are often characterised by sensationalistic and provocative nature, or even due to the popularity of certain counter-narratives within platform-specific communities.

The comparisons are statistically significant and align with the average engagement values, highlighting a consistent pattern across platforms. To assess these differences, we performed Wilcoxon rank-sum tests—standard, bootstrapped, and log-transformed—which all yielded highly significant results (p -values $< 10^{-3}$). These findings underscore a robust divergence in user engagement with reliable versus unreliable content, thus offering a critical perspective on how misinformation competes with factual content on social media platforms (see SI for full tables and robustness checks).

The bar chart outcomes in Fig. 4 can be partially explained by the different features, algorithms, and user bases of each platform. For instance, YouTube’s engagement-driven algorithm tends to amplify content that triggers strong emotional responses, while Twitter historically prioritized real-time updates within follower networks. These design choices influence not only what content is surfaced, but also how users interact with it. Although it is difficult to disentangle algorithmic amplification from shifts in user behaviour or platform policies in observational studies, these factors likely interact and co-shape engagement dynamics. Moreover, past research has shown that false or questionable content often appears more novel or surprising, which may attract more user interaction⁵⁴. This could help explain the increasing engagement with questionable sources on several platforms. The exception of Twitter, where reliable sources received comparatively more engagement, may relate to the platform’s user base. For instance, climate-related discussions on Twitter often involve communities of experts, activists, and journalists, potentially fostering greater interaction with scientifically sound information. These trends may also reflect self-selection effects or norms around information credibility within specific online communities. Similar findings have been found also for general news content which is not climate specific. For example, a study on Facebook found higher level of engagement with unreliable content related to politics⁵⁵, while analysing Twitter⁵⁶ has shown substantial difference in the production of engaging content by the pro-vaccine anti anti-vaccine groups, the former receiving substantially more engagement. To further investigate the nature of engagement across content reliability, we present a set of contrasting examples of climate-related posts labelled as either “reliable” or “unreliable” across Facebook, Instagram, YouTube, and Twitter. While only a small subset

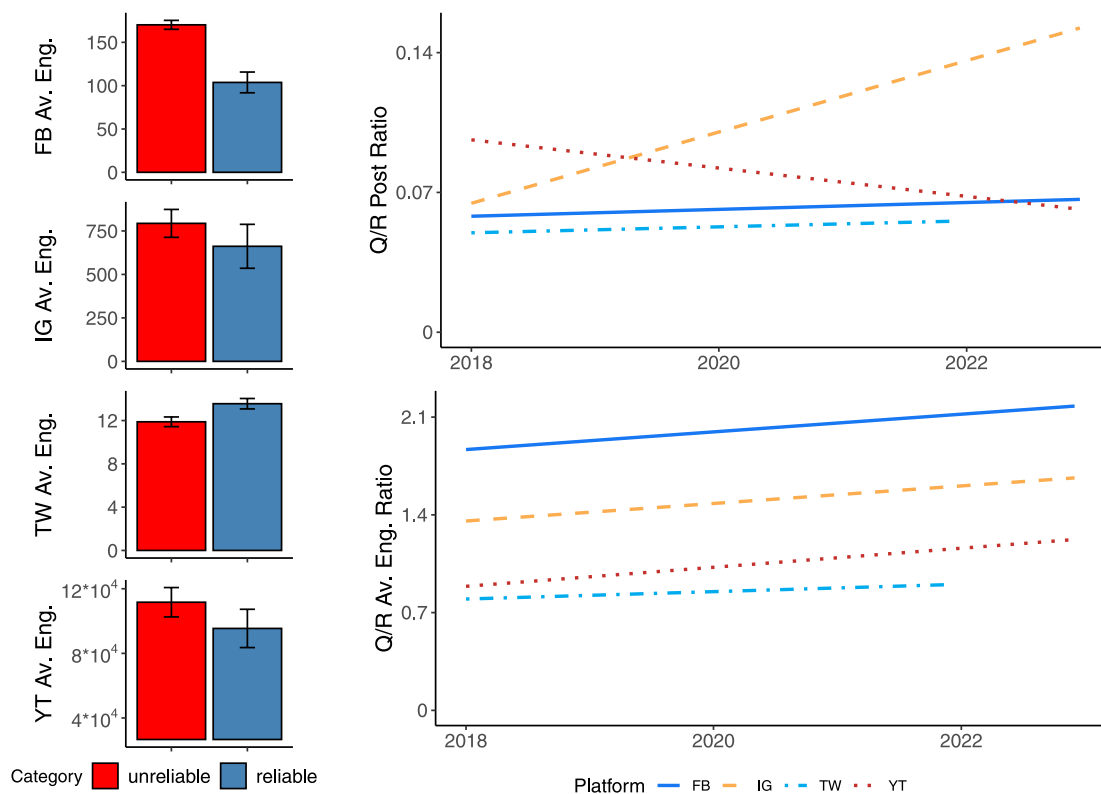


Fig. 4. Comparison of average engagement between posts from unreliable and reliable sources. The four bar charts in the left column show the mean values of bootstrap samples from posts linking to news outlets belonging to two categories: unreliable and reliable. The right column contains regression lines for the four platforms, obtained from time series data of the ratio of posts categorised as unreliable to reliable (upper panel), and from the ratio of average interactions with unreliable to reliable (bottom panel).

of the full dataset, these examples offer insight into the types of narratives that tend to drive user interaction. On Facebook, unreliable content includes claims like “We’ve been lied to for a long time about climate change. Everyone in America needs to see this!”, whereas reliable posts focus on factual summaries such as “scientists say the melting ice in Antarctica is responsible for about one-third of all sea-level rise around the world.” On Instagram, unreliable narratives refer to geoengineering as “not a conspiracy theory but very real indeed”, while reliable content emphasizes scientific data, such as the environmental impact of animal agriculture. On YouTube, unreliable content often takes a confrontational tone—for example, calling Greta Thunberg a “TERRIBLE Role Model”—in contrast to reliable posts that highlight scientific innovations like “artificial photosynthesis to fight climate change”. On Twitter, unreliable tweets may simply state “Climate change hoax” whereas reliable ones stress scientific consensus, such as declarations of climate emergency signed by thousands of scientists. These examples suggest that posts conveying sensationalist or polarizing narratives—particularly those questioning mainstream perspectives or invoking controversy—tend to resonate strongly with users. However, engagement is not solely driven by tone or topic. Prior literature highlights the role of structural dynamics such as opinion polarization, echo chambers, and filter bubbles in reinforcing users’ preexisting beliefs and preferences^{10,57,58}. These mechanisms likely contribute to the amplification of certain types of content regardless of their factual accuracy, shaping how information is consumed and shared on social media.

The right column of Fig. 4 displays the results of two linear regression analyses. The top right panel shows the monthly ratios of the number of posts linking to unreliable sources relative to those linking to reliable sources, while the bottom right panel illustrates the monthly ratios of average engagement with posts linking to unreliable sources compared to reliable ones. In both cases, time (measured in months) is the independent variable, and the ratio is the dependent variable. The regression lines were estimated using the Ordinary Least Squares (OLS) method. The applied regression model is $Y_t = \beta_0 + \beta_1 t + \varepsilon_t$, where Y_t represents the monthly ratio, β_0 is the intercept, β_1 the slope indicating the temporal trend, and ε_t the error term. The resulting trends highlight a number of relevant dynamics. Overall, we note that the share of posts from unreliable sources with respect to those from reliable ones is approximately 7% thus highlighting a significantly lower production of such kind of content. In more detail, the intercepts range from ~ 0.05 in the case of Twitter to ~ 0.1 in the case of YouTube while the trends show different slopes ranging from positive to slightly negative values. Conversely, the bottom right panel reveals a notable disparity in engagement levels, with unreliable content receiving, on average, 1.2 times the engagement of its reliable counterparts. The higher average engagement with unreliable content varies across platforms and generally shows a slight increasing trend over time. Detailed regression coefficients are provided in SI (Tables 3 and 4). Note also that, in the case of YouTube, we obtained consistent results excluding views in the count of engagement as shown in SI. Despite the lower volume of unreliable content compared to reliable sources, its disproportionate impact on user engagement is noteworthy. This counterintuitive dynamic highlights a critical aspect of social media ecosystems: even though unreliable content is less prevalent in terms of volume, it often obtains significantly more user engagement.

Hashtags used by sources spreading unreliable and reliable information

After outlining the general dynamics of the climate change discussion, we now focus on how specific concepts are treated within unreliable and reliable posts in order to highlight differences and similarities in their respective discourses. According to previous literature^{24,59–61}, we operate under the premise that hashtags significantly encapsulate posts’ key aspects and therefore provide a compact way to analyse topics⁶². We first identify the most frequently used hashtags in posts labelled as either unreliable or reliable and compute their relative difference i.e., the difference between the relative frequency in each of the two categories in order to understand the importance with respect to the two sides. For an general overview of hashtag usage on the four platforms please refer to SI (Overview of hashtag usage section).

Figure 5 shows the difference between the relative frequencies normalised over the set made up of the twelve most used hashtags appearing in posts sourced by either unreliable or reliable accounts, aggregated over time. This analysis focuses on how these hashtags are utilised in posts from each source category, highlighting the distinctions in thematic engagement. In some cases, the top twelve hashtags were common between unreliable and reliable posts, as shown in magenta. The analysis indicates that accounts posting information from unreliable sources tend to focus on hashtags that align with narratives often seen as controversial or denying mainstream scientific consensus⁶³. For example, hashtags such as #climatehoax are prominent in content shared by unreliable sources, reflecting a tendency to challenge the scientific consensus on climate change. Conversely, reliable sources frequently use hashtags related to scientific discussions and global events such as #cop26, #cop27, and #actionclimate, emphasising their alignment with established scientific perspectives and international efforts to address climate change. In this context, Instagram shows a relatively different behaviour having hashtags appearing in posts linking to unreliable sources mostly related to the vegan community, and a group of hashtags appearing in posts linking to reliable sources having some geographic resolution. Exception made for Instagram, similar discussion themes have been observed in previous research on climate communication online⁴. The difference in hashtag use suggests that accounts that are more likely to engage with unreliable sources, employ a lexicon, in terms of hashtags, that may reinforce sceptical views on climate change. Conversely, accounts that employ reliable sources, present content that supports and promotes scientific and global consensus on the issue. This divergence in hashtags usage underscores the broader polarisation in how climate change is discussed and perceived across different segments of social media users.

After showing which hashtags are the most representative for each kind of content we attempt at estimating their separation in the public debate using network analysis. In more detail, we build up four undirected co-occurrence networks (one per each social media platform) using hashtags as nodes and co-occurrence of

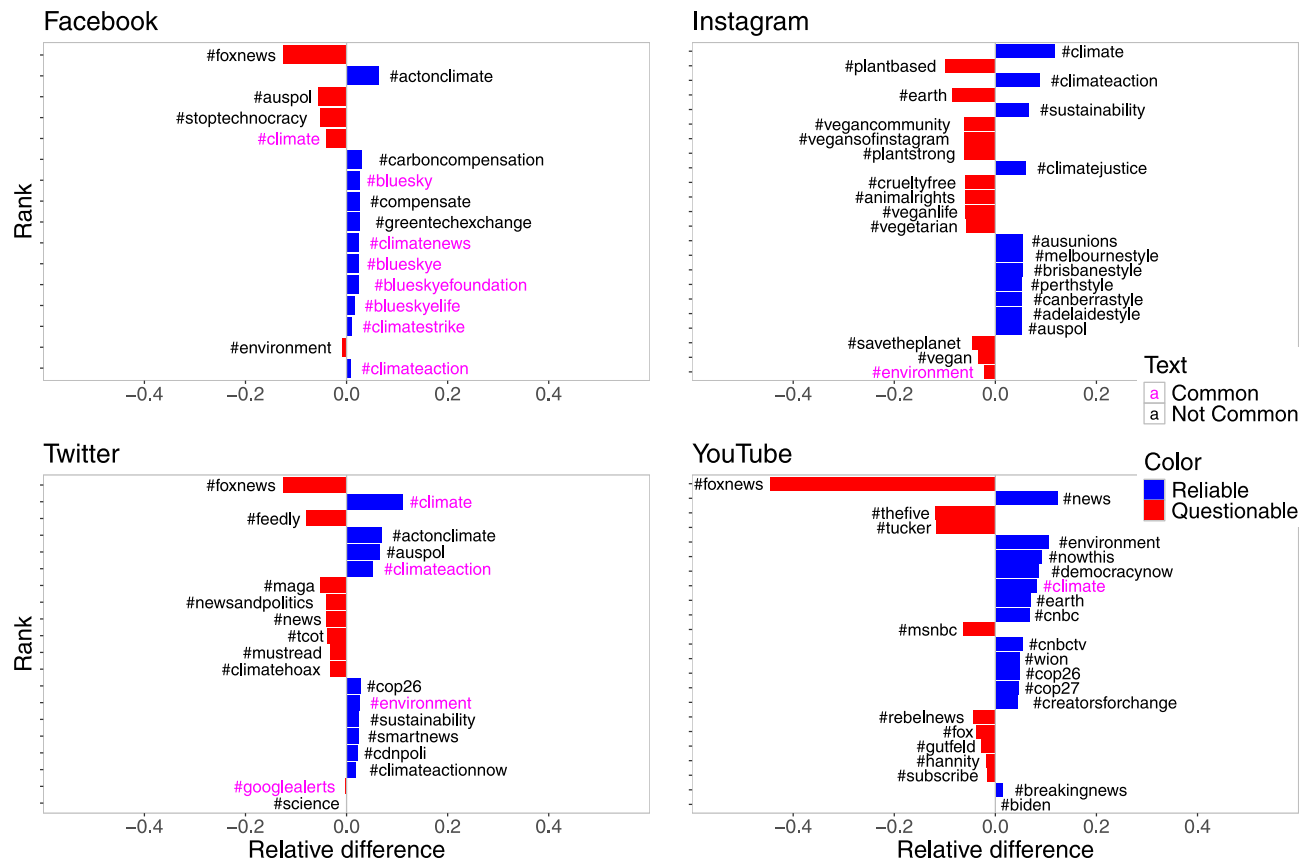


Fig. 5. The panel displays the top 12 hashtags with the highest relative frequencies for content referencing both reliable and unreliable sources on FB, IG, TW, and YT respectively. The bars represent the difference between the relative frequencies calculated on the top 12 hashtags used in reliable and unreliable content. In red and blue, we have those hashtags mentioned above in unreliable and reliable content, respectively. The magenta labels underline those hashtags that are present in the top 12 of both reliable and unreliable categories. Rank highlights the position in term of relative frequency difference among reliable and unreliable hashtags.

hashtags in posts as links. In these networks, the weight of a link corresponds to the number of times two hashtags co-occurred in posts. The networks are sparse (that is the number of links is much lower with respect to its maximum) and they are made up of {680,280; 569,962; 449,182; 19,431} nodes and {10,766,970; 13,703,602; 3,399,321; 517,121} links for Facebook, Instagram, Twitter and YouTube respectively. The resulting networks are not made up of a single connected component and therefore we extract their largest connected components which are made up of about {93%, 98%, 91%, 91%} original nodes respectively. Operating on the largest connected components, we compute the shortest paths among labelled hashtags in order to quantify their separation in the network and therefore the tendency to be used in different contexts. Hashtags were labelled as either unreliable or reliable based on the difference of their relative frequency in posts by unreliable and reliable sources. Specifically, after standardising the differences between relative frequencies, we labelled as unreliable hashtags having a z-score of standardised relative differences lower than -3 (indicating significant occurrence in unreliable posts), and reliable those with a z-score greater than +3. A more formal explanation is provided in Methods (see Network construction and hashtag labelling Section).

Table 4 provides values of the average shortest path distances between hashtags classified as “reliable” and “unreliable” on the four different social media platforms: Facebook, Instagram, Twitter, and YouTube. These values represent the average separation between hashtags of these categories in the co-occurrence networks constructed for each platform.

On Facebook, reliable hashtags tend to be very close to each other, indicating that they are often used together. In contrast, unreliable hashtags show a greater average distance (5.44 steps), suggesting less coherence and density in their usage. The distance between reliable and unreliable hashtags is intermediate (3.78 steps), indicating some interaction but with significant separation, very similar to the overall distance between the nodes of the network. The high standard deviation for l_{qq} and l_{rq} highlights considerable variability in the joint usage of these hashtags. On Instagram distances between labelled hashtags are, on average, consistent across classes (~ 2 steps) and lower with respect to the average distances, thus indicating their proximity in the network. The uniformity of such results could be due to a relatively poor identification of hashtags belonging to such categories, as discussed with respect to Fig. 5. On Twitter, reliable hashtags are also very close to each other (1.95 steps). Unreliable hashtags are relatively more distant (2.38 steps), but less fragmented than in the Facebook

	Facebook	Instagram	Twitter	YouTube
l_{rr}	2.05±0.39	1.90±0.30	1.95±0.31	1.93±0.46
l_{qq}	5.44±6.62	1.99±0.24	2.38±1.87	7.24±4.45
l_{rq}	3.78±4.77	1.93±0.33	2.14±1.35	4.71±3.40
l	3.94±1.54	3.89±10.88	4.08±2.07	3.96±2.25

Table 4. Average shortest path and standard deviation for reliable-reliable (l_{rr}), unreliable-unreliable (l_{qq}), reliable-unreliable (l_{rq}), and overall (l) hashtag pairs on Facebook, Instagram, Twitter, and YouTube.

case (also considering the average length of shortest paths in the network). The distance between reliable and unreliable hashtags is the smallest among the three platforms (2.14 steps), suggesting greater interaction between the two groups on Twitter compared to Facebook and YouTube. The standard deviation is lower than on Facebook, indicating more uniform distances. On YouTube, reliable hashtags have the shortest average distance (1.93 steps), similarly to Facebook and Twitter while unreliable hashtags show the highest average distance (7.24 steps), indicating significant dispersion. The distance between reliable and unreliable hashtags is also quite high (4.71 steps), suggesting a clear separation between the two groups. The large variability in distances (l_{qq} and l_{rq}) indicates a very heterogeneous use of hashtags on YouTube. These results suggest that reliable hashtags tend to form cohesive clusters, possibly reflecting a more organized online community built around shared and verified information sources. This behaviour is consistent with prior work that highlights the formation of echo chambers in which reliable information circulates within homogeneously connected users¹⁰. Conversely, the dispersion of unreliable hashtags may be indicative of more fragmented communication around misinformation-related narratives. Consequently, the use of questionable hashtags can create small and relatively isolated clusters, each presenting its own perspective regarding climate change and related events, which aligns with recent studies highlighting the heterogeneous and uncoordinated nature of misinformation dynamics⁶⁴. In summary, we note that reliable hashtags tend to cluster closely together across all platforms, while unreliable hashtags are more dispersed and variable indicating a higher fragmentation in their use.

Conclusions

Our analysis of the climate change discourse on social media platforms reveals intricate dynamics of user engagement, thematic developments through hashtags, and the challenging interplay between information spread by reliable unreliable sources. With the aim to explicitly focus on climate change misinformation to inform evidence-based policy interventions and improve climate communication, we find that engagement spikes significantly during major global events (Climate Action Week, Fridays for Future, and COP summits), underscoring the role of social media as a pivotal platform for climate activism and discussion. This enhanced engagement may promote extensive information dissemination and drive public participation in climate-related initiatives.

Our research highlights a consistent rise in climate change-related content across four different mainstream platforms that are likely to contain the majority of online climate discussions, although the nature and engagement trends vary markedly. This variability reflects the distinct roles that these platforms play in facilitating discussions on climate issues. For instance, we observed differences in engagement, likely influenced by platform-specific characteristics, such as user demographics or content formats. The decline in engagement on Facebook, for example, may be linked to its longevity and older user base⁶⁵, while each platform's audience and content format influence the patterns we observe. For instance, younger audiences on YouTube may engage more with sensationalist content, while Twitter's retweets and trending hashtags create a different form of viral content amplification. These differences are not solely demographic. Platform-specific affordances-such as Twitter's character limit and retweet functionality, YouTube's algorithmic promotion of engaging videos, or Instagram's emphasis on visual storytelling-may also shape the type of content that circulates and how users interact with it. For instance, polarizing posts may gain traction more easily on Twitter due to its trend-based structure and amplification mechanisms, while sensationalist videos may thrive on YouTube^{7,10}. Moreover, the different role of hashtags across platforms-more integrated and frequent on Twitter, less so on Facebook- may further influence the visibility and engagement with specific narratives. Identifying these different trends across platforms is important given that existing research has been largely platform specific, preventing a comparative analysis².

A critical point of our study is the significant engagement with content from unreliable sources and its increasing trends compared to content from reliable ones, particularly on platforms like Facebook. This issue is also substantiated by a clear separation in the topics treated within posts referring to unreliable versus reliable information sources. This analysis is important given that most existing research on climate communication has not focused specifically on climate misinformation⁴¹, and most studies of social media misinformation in general focus on political misinformation or Covid-19 misinformation^{4,30,66}.

Beyond the presence of potential confounding factors, which we tried to mitigate through robustness checks and extra statistical analysis, our study has limitations due to data accessibility⁶⁷. This is particularly true for Twitter, where restricted data availability, due to the academic API shutdown, limited the time frame of our analysis. Relatedly, in selecting platforms, we chose mainstream social media sites like Facebook, Instagram, Twitter, and YouTube due to their widespread use and established presence in the literature. We excluded platforms such as TikTok and Reddit, either due to data unavailability at the time of collection or the lack of substantial data volume on climate change. Other niche or messaging platforms (e.g., Telegram, WhatsApp) were

Criterion	Linear model	Quadratic model
AIC	2020.861	1998.047
BIC	2035.521	2014.802
ANOVA (F, p)	–	F = 27.15, p < 0.001

Table 5. Comparison between the linear and quadratic models. The quadratic model shows lower AIC and BIC values, and the ANOVA test confirms a significantly better fit ($p < 0.001$), indicating that the inclusion of a quadratic term improves model performance.

similarly excluded due to their specific functionalities or user bases, which would require a separate analysis. Also platform differences and keywords selection could represent a potential limitation in terms of representativeness of diverse groups of users⁶⁸. Similarly, the use of a source-based approach and the combination of two different sources^{10,53,69} for identifying unreliable and reliable accounts could hinder the study from providing a complete picture regarding engagement with misinformation. We strongly encourage future work to look at a broader set of social media platforms where possible, including emerging platforms which cater to different demographic groups of specific political communities, but we note that researchers should not lose sight of the important interdependencies between established and emerging social media platforms⁷⁰.

That being acknowledged, our study underscores the profound impact of social media in monitoring perceptions and discussions about climate change, highlighting the need of both fostering informed discourse and countering misinformation. Future works should focus on entering the debate in greater detail by tackling specific themes, targeting specific pages, and possibly comparing mainstream social media platforms with alternative ones. A broader analysis across diverse social media ecosystems and topics, incorporating both traditional and emerging platforms, would be therefore valuable for future studies. In this context, developing mechanisms and policies to strengthen the reliability of information and engage diverse audiences effectively will be crucial in leveraging social media’s capacity to support global climate initiatives in this digitally interconnected era.

Methodology
Data collection

The data collection process is tailored according to the characteristics of the four different platforms. In general, we performed keyword search to identify relevant social media content. For Facebook, Instagram, and YouTube, the keywords used to capture discussions related to climate change are: {*climate change, climate crisis, climate emergency, global warming*}. Due to unavailability of the Twitter API for academics after the Elon Musk acquisition we had to resort to an available dataset retrieved using only the keyword: {*climate change*}. We collected the Facebook and Instagram datasets using CrowdTangle, a Meta owned archive which grants access to posts (and their engagement) from 7M+ pages, groups, and verified profiles in the case of Facebook, and 2M+ public accounts in the case of Instagram. In more detail, the archive includes: all public Facebook pages with more than 50K likes, all public Facebook groups with 95k+ members, all US-based public groups with 2k+ members, and all verified profiles; all public Instagram accounts with more than 50K followers, as well as all verified accounts. More details are provided at <https://help.crowdtangle.com/en/articles/4201940-about-us>. For YouTube, we collected a large sample of video videos and their engagement data using the YouTube Data API v3, specifically searching for videos that match our keywords and crawling the relevant related videos. In the case of Twitter, we gathered tweets and user information using the official Twitter API in 2022, prior to Elon Musk’s takeover of the platform. The API is no longer available to researchers free of charge. It is important to note that the search query on Twitter was more limited compared to the other platforms, as previously mentioned. Another distinction between Facebook, Instagram, YouTube, and Twitter is the observation period. Data from the first three social media platforms range from January 2018 to December 2022. The time window for Twitter lasts four-years ending in December 2021. Table 1 provides a data breakdown.

Interrupted time series

The interrupted time series (ITS) approach is a widely used method for evaluating the impact of an event or intervention on a time-dependent variable. This method models the data as a time series before and after the event, allowing for the identification of structural changes in the trend following the event. Unlike simple pre-post comparisons, ITS accounts for underlying temporal dynamics, distinguishing between immediate effects and longer-term changes. ITS is commonly applied in various fields, including public health, economics, policy evaluation, and environmental science^{71–73}.

We explored different specifications for modelling engagement dynamics as show in Table 5. Model comparisons using ANOVA⁷⁴, Akaike Information Criterion (AIC)⁷⁵, and Bayesian Information Criterion (BIC)⁷⁶ confirmed that a quadratic time trend significantly improved model fit, capturing non-linear engagement trajectories⁷⁷.

The final ITS model is specified as follows:

$$Y_t = \beta_0 + \beta_1 T + \beta_2 T^2 + \beta_3 D_1 + \beta_4 D_2 + \beta_5 (T \times D_1) + \beta_6 (T \times D_2) + \epsilon_t \tag{1}$$

where Y_t represents engagement at time t , T is the time variable, and T^2 is its quadratic term. The dummy variables D_1 and D_2 indicate the post-event periods for the first and second events, respectively. The interaction

terms $T \times D_1$ and $T \times D_2$ capture long-term changes in engagement after each event. Finally, ϵ_t represents the error term.

Hodrick-Prescott filter

The Hodrick-Prescott filter is a statistical method commonly employed in economics and finance to analyse economic time series, especially when they display non-stationary behaviour and to identify cyclical variation^{78–80}. It is used to decompose the time series into two main components: the long-term trend and the short-term cyclical component.

The mathematical equation for the Hodrick-Prescott filter is :

$$y_t = \tau_t + c_t, \quad (2)$$

while the function to be optimised is:

$$\min_{\tau_t} \sum_{t=1}^T (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - 2\tau_t + \tau_{t-1})^2]. \quad (3)$$

The equation represents the minimisation problem to obtain the decomposition of the time series y_t into two components: the trend τ_t and the cyclical component c_t . The first term $\sum_{t=1}^T (y_t - \tau_t)^2$ measures the adherence of the trend component to the original data. The second term $\lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - 2\tau_t + \tau_{t-1})^2]$ penalises excessive variations in the trend over time, where λ is the smoothing parameter controlling the relative importance of the two terms in the objective function.

Matching sources

In order to assign reliability labels to content we combine the classification of web pages provided by two independent agencies that assess the credibility of news sources, namely NewsGuard and Media Bias/Fact Check (MBFC).

NewsGuard employs a team of highly experienced journalists and editors to analyse sites that publish news and information, providing ratings and reliability scores based on nine journalistic criteria. These criteria pertain to the basic standards of credibility and transparency in journalism. Each criterion is assigned a certain score, and the total, ranging from 0 to 100, represents the site's overall reliability. This score is also reported as a percentage. It indicates the extent to which the site meets the various criteria. Each site is accompanied by an information sheet, accessible by clicking on the score, which details the reasons for NewsGuard's assessment. Generally, a score above 60 indicates trustworthiness.

MBFC assigns a qualitative (non-numerical) rating for each news and commentary site, placing them on the political spectrum from left to right. The scale includes values such as 'extreme left', 'leaning left', 'center-left', 'poorly biased', 'leaning center-right', 'leaning right' and 'extreme right'. Additionally, a second rating on the accuracy and impartial objectivity of the published content is assigned, with values like 'very high', 'high', 'mostly factual', 'mixed', 'low', and 'very low'. Furthermore, MBFC provides extra labels for certain source including "unreliable" and "Conspiracy/Pseudo-Science". According to the MBFC's website: "A unreliable source exhibits one or more of the following: extreme bias, consistent promotion of propaganda/conspiracies, poor or no sourcing to credible information, a complete lack of transparency, and/or is fake news. [...] Sources listed in the unreliable Category may be very untrustworthy and should be fact-checked on a per-article basis". While a "Conspiracy/Pseudo-Science": "This category is for sources that disseminate unverified information related to known conspiracies or pseudoscientific claims. For instance, sources denying human-influenced climate change or promoting anti-vaccination stances are labelled as pseudoscience. According to the Stanford Encyclopedia of Philosophy on Science and Pseudoscience, there's a consensus among knowledge disciplines that certain topics, like creationism and climate change denial, are pseudosciences. To be included in this category, a source's primary focus must be on conspiracies or pseudoscience."⁸¹.

Using the classifications from both organisations, we assigned a dichotomous label (either reliable or unreliable) to each source employing both the score assigned by NewsGuard (if < 60 the source was considered unreliable otherwise reliable) and the label from MBFC (if equals to one on the following factuality labels {'mixed', 'low', and 'very low'} the source was considered unreliable otherwise reliable). This procedure is consistent with the guidelines of the two agencies where in the case of NewsGuard the official website states: "The language that introduces and summarizes each website's score varies, depending on the score, as follows: 100 – High Credibility: This website adheres to all nine standards of credibility and transparency. 75–99 – Generally Credible: This website mostly adheres to basic standards of credibility and transparency. 60–74 – Credible with Exceptions: This website generally maintains basic standards of credibility and transparency-with significant exceptions. 40–59 – Proceed with Caution: This website is unreliable because it fails to adhere to several basic journalistic standards. 0–39 Proceed with Maximum Caution: This website is unreliable because it severely violates basic journalistic standards"⁸². We used NewsGuard scores as the primary source for categorizing outlets, and for those domains not listed in the NewsGuard database, we incorporated ratings from MBFC. According to our procedure, an outlet was labelled as "unreliable" if it received a score below 60 from NewsGuard. If an outlet was not included in the NewsGuard dataset, we used the MBFC rating and, consistently with previous work¹⁰, we classified it as unreliable if it received the "unreliable" label by MBFC itself or if it was marked as either "low" or "very low" in terms of accuracy and impartial objectivity or if it was labelled as "Conspiracy/Pseudo-Science" and had a

label lower than or equal to “mixed”. More details, regarding the categorized domains and the matching between sources are available in SI.

This approach allowed us to encompass a broader range of sources compared to individual companies, thereby enhancing the accuracy and reliability of our classification. For labelling Facebook posts, we focus on content posted by Pages, Groups, and Verified Profiles that include URLs linking to external websites that, when decompressed, point to a domain among those labelled by either NewsGuard or MBFC. We followed a different procedure for Instagram, Twitter, and YouTube. We extracted URLs referring to external websites from the text, image, or video descriptions, depending on the platform. After collecting all the URL links present in the text, we expanded the compressed ones. This enabled us to cross-reference the domains of the extracted links with our dataset containing reliability and unreliability classifications of the sources.

Network construction and hashtag labelling

We start with a graph $G = (V, E)$, where V represents the set of nodes n and E represents the set of m edges connecting pairs of nodes. Specifically, nodes represent hashtags, and connections are determined by the number of times two hashtags appear together in the same post. The links $w_{(i,j)}$ indicate the connection strength (i.e. the number of co-occurrences) between nodes i and j , making our network weighted. We denote our weighted network as $N = (V, W)$, where the matrix W is our adjacency matrix containing the weights of the links $w_{i,j}$.

We calculate the relative frequency of hashtags, specifically the frequency with which hashtags appear in reliable and unreliable posts respectively, normalised by the total number of hashtags in reliable or unreliable categories. The relative frequency $f_{h,c}$ of a hashtag h in category c (either reliable or unreliable) is calculated as:

$$f_{h,c} = \frac{\text{count of hashtag } h \text{ in category } c}{\text{total count of all hashtags in category } c}$$

To label hashtags as either unreliable or reliable, we calculate the z-score of each hashtag. The z-score z_h is computed as follows:

$$z_h = \frac{d_h - \mu_d}{\sigma_d}$$

where $d_h = f_{h,\text{reliable}} - f_{h,\text{unreliable}}$ is the difference in relative frequencies of the hashtag h between the reliable and unreliable categories, μ_d is the mean of these differences for all hashtags, and σ_d is the standard deviation of these differences.

By computing these metrics, we can effectively categorize hashtags and analyse their distribution within different content reliability categories.

Data availability

The datasets generated during and analysed during the current study are available from the corresponding author on reasonable request.

Received: 23 July 2024; Accepted: 19 May 2025

Published online: 28 May 2025

References

- Falkenberg, M. et al. Growing polarization around climate change on social media. *Nat. Clim. Chang.* **12**, 1114–1121 (2022).
- Pearce, W., Niederer, S., Özkula, S. M. & Sánchez Querubin, N. The social media life of climate change: Platforms, publics, and future imaginaries. *Wiley Interdiscip. Rev. Clim. Change* **10**, e569 (2019).
- Chen, T. H. Y., Salloum, A., Gronow, A., Ylä-Anttila, T. & Kivelä, M. Polarization of climate politics results from partisan sorting: Evidence from Finnish twittersphere. *Glob. Environ. Chang.* **71**, 102348 (2021).
- Del Vicario, M. et al. The spreading of misinformation online. *Proc. Natl. Acad. Sci.* **113**, 554–559 (2016).
- Del Vicario, M. et al. Echo chambers: Emotional contagion and group polarization on Facebook. *Sci. Rep.* **6**, 37825 (2016).
- Falkenberg, M., Zollo, F., Quattrociochi, W., Pfeffer, J. & Baronchelli, A. Affective and interactional polarization align across countries. *arXiv preprint, arXiv:2311.18535* (2023).
- Avalle, M. et al. Persistent interaction patterns across social media platforms and over time. *Nature* **628**, 582–589 (2024).
- Brady, W. J. & Van Bavel, J. J. The mad model of moral contagion: The role of motivation, attention, and design in the spread of moralized content online. *Perspect. Psychol. Sci.* **15**, 978–1010 (2020).
- Brady, W. J., Jackson, J. C., Lindström, B. & Crockett, M. Algorithm-mediated social learning in online social networks. *Trends Cognit. Sci.* (2023).
- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociochi, W. & Starnini, M. The echo chamber effect on social media. *Proc. Natl. Acad. Sci.* **118**, e2023301118 (2021).
- Valensise, C. M., Cinelli, M. & Quattrociochi, W. The drivers of online polarization: Fitting models to data. *Inf. Sci.* **642**, 119152 (2023).
- Salloum, A., Chen, T. H. Y. & Kivelä, M. Anatomy of elite and mass polarization in social networks (2024). [arXiv:2406.12525](https://arxiv.org/abs/2406.12525).
- Roozenbeek, J. & Zollo, F. Democratize social-media research-with access and funding. *Nature* **612**, 404 (2022).
- Wagner, M. W. Independence by permission. *Science* **381**, 388–391 (2023).
- Masson-Delmotte, V. et al. Climate change 2021: the physical science basis. *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* **2**, 2391 (2021).
- Auer, M. R., Zhang, Y. & Lee, P. The potential of microblogs for the study of public perceptions of climate change. *Wiley Interdiscip. Rev. Clim. Change* **5**, 291–296 (2014).
- Vicario, M. D., Quattrociochi, W., Scala, A. & Zollo, F. Polarization and fake news: Early warning of potential misinformation targets. *ACM Trans. Web (TWEB)* **13**, 1–22 (2019).

18. Cinelli, M. et al. The covid-19 social media infodemic. *Sci. Rep.* **10**, 1–10 (2020).
19. Moser, S. C. Communicating climate change: history, challenges, process and future directions. *Wiley Interdiscip. Rev. Clim. Change* **1**, 31–53 (2010).
20. McCombs, M. E. & Shaw, D. L. The agenda-setting function of mass media. *Public Opin. Q.* **36**, 176–187 (1972).
21. Forum, W. E. *Global risk report 2024* (Tech, Rep, 2024).
22. Farrell, J., McConnell, K. & Brulle, R. Evidence-based strategies to combat scientific misinformation. *Nat. Clim. Chang.* **9**, 191–195 (2019).
23. Gupta, A., Lamba, H., Kumaraguru, P. & Joshi, A. Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. In *Proceedings of the 22nd international conference on World Wide Web*, 729–736 (2013).
24. Torricelli, M. et al. How does extreme weather impact the climate change discourse? insights from the twitter discussion on hurricanes. *Plos Climate* **2**, e0000277 (2023).
25. Flynn, D. J., Nyhan, B. & Reifler, J. The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Polit. Psychol.* **38**, 127–150 (2017).
26. Lazer, D. M. et al. The science of fake news. *Science* **359**, 1094–1096 (2018).
27. Allcott, H. & Gentzkow, M. Social media and fake news in the 2016 election. *J. Econ. Perspect.* **31**, 211–236 (2017).
28. Loy, L. S., Hamann, K. R. & Reese, G. Navigating through the jungle of information. informational self-efficacy predicts climate change-related media exposure, knowledge, and behaviour. *Clim. Change* **163**, 2097–2116 (2020).
29. Hassan, I., Musa, R. M., Latiff Azmi, M. N., Razali Abdullah, M. & Yusoff, S. Z. Analysis of climate change disinformation across types, agents and media platforms. *Information Development* 02666669221148693 (2023).
30. Sharma, K., Seo, S., Meng, C., Rambhatla, S. & Liu, Y. Covid-19 on social media: Analyzing misinformation in twitter conversations. *arXiv preprint, arXiv:2003.12309* (2020).
31. Pearce, W., Holmberg, K., Hellsten, I. & Nerlich, B. Climate change on twitter: Topics, communities and conversations about the 2013 ipcc working group 1 report. *PLoS One* **9**, e94785 (2014).
32. O'Neill, S., Williams, H. T., Kurz, T., Wiersma, B. & Boykoff, M. Dominant frames in legacy and social media coverage of the ipcc fifth assessment report. *Nat. Clim. Chang.* **5**, 380–385 (2015).
33. Walter, S., Lörcher, I. & Brüggemann, M. Scientific networks on twitter: Analyzing scientists' interactions in the climate change debate. *Public Underst. Sci.* **28**, 696–712 (2019).
34. Cody, E. M., Reagan, A. J., Mitchell, L., Dodds, P. S. & Danforth, C. M. Climate change sentiment on twitter: An unsolicited public opinion poll. *PLoS One* **10**, e0136092 (2015).
35. Sanford, M., Painter, J., Yasseri, T. & Lorimer, J. Controversy around climate change reports: a case study of twitter responses to the 2019 ipcc report on land. *Clim. Change* **167**, 59 (2021).
36. Sanford, M. & Painter, J. Divergences between mainstream and social media discourses after cop26, and why they matter. *Oxford Open Climate Change* kgae006 (2024).
37. Müller-Hansen, F. et al. Attention, sentiments and emotions towards emerging climate technologies on twitter. *Glob. Environ. Chang.* **83**, 102765 (2023).
38. Debnath, R., Creutzig, F., Sovacool, B. K. & Shuckburgh, E. Harnessing human and machine intelligence for planetary-level climate action. *npj Clim. Action* **2**, 20 (2023).
39. Pera, A. & Aiello, L. M. Shifting climates: Climate change communication from youtube to tiktok. *arXiv preprint, arXiv:2312.04974* (2023).
40. Parth, A.-M., Weiss, J., Firat, R. & Eberhardt, M. "how dare you!"-the influence of fridays for future on the political attitudes of young adults. *Front. Political Sci.* **2**, 611139 (2020).
41. Spaiser, V., Nisbett, N. & Stefan, C. G. "how dare you?"-the normative challenge posed by fridays for future. *PLoS Clim.* **1**, e0000053 (2022).
42. Pera, A. & Aiello, L. M. Narratives of collective action in youtube's discourse on veganism. *arXiv preprint, arXiv:2401.09210* (2024).
43. Debnath, R., Bardhan, R., Reiner, D. M. & Miller, J. Political, economic, social, technological, legal and environmental dimensions of electric vehicle adoption in the united states: A social-media interaction analysis. *Renew. Sustain. Energy Rev.* **152**, 111707 (2021).
44. Debnath, R. et al. Conspiracy spillovers and geoengineering. *Iscience* **26** (2023).
45. Vu, H. T. et al. Social media and environmental activism: Framing climate change on facebook by global ngos. *Sci. Commun.* **43**, 91–115 (2021).
46. Shapiro, M. A. & Park, H. W. More than entertainment: Youtube and public responses to the science of global warming and climate change. *Soc. Sci. Inf.* **54**, 115–145 (2015).
47. Hautea, S., Parks, P., Takahashi, B. & Zeng, J. Showing they care (or don't): Affective publics and ambivalent climate activism on tiktok. *Social media+ society* **7**, 20563051211012344 (2021).
48. Hamid, S., Ijab, M. T., Sulaiman, H., Anwar, M. R. & Norman, A. A. Social media for environmental sustainability awareness in higher education. *Int. J. Sustain. Higher Edu.* **18**, 474–491 (2017).
49. Mavrodieva, A. V., Rachman, O. K., Harahap, V. B. & Shaw, R. Role of social media as a soft power tool in raising public awareness and engagement in addressing climate change. *Climate* **7**, 122 (2019).
50. Yin, J., Brooks, M., Wang, D. & Chi, G. Characterizing climate change sentiments in alaska on social media. *Digital Geograp. Soc.* **8**, 100110 (2025).
51. Khan, M. L. Social media engagement: What motivates user participation and consumption on youtube?. *Comput. Hum. Behav.* **66**, 236–247 (2017).
52. Auxier, B. et al. Social media use in 2021. *Pew Res. Center* **1**, 1–4 (2021).
53. Bovet, A. & Makse, H. A. Influence of fake news in twitter during the 2016 us presidential election. *Nat. Commun.* **10**, 7 (2019).
54. Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**, 1146–1151 (2018).
55. Allcott, H., Gentzkow, M. & Yu, C. Trends in the diffusion of misinformation on social media. *Res. Politics* **6**, 2053168019848554 (2019).
56. Germani, F. & Biller-Andorno, N. The anti-vaccination infodemic on social media: A behavioral analysis. *PLoS One* **16**, e0247642 (2021).
57. Bakshy, E., Messing, S. & Adamic, L. A. Exposure to ideologically diverse news and opinion on facebook. *Science* **348**, 1130–1132 (2015).
58. Nyhan, B. et al. Like-minded sources on facebook are prevalent but not polarizing. *Nature* **620**, 137–144 (2023).
59. Small, T. A. What the hashtag? a content analysis of canadian politics on twitter. *Inform. Commun. Soc.* **14**, 872–895 (2011).
60. Rogers, R. et al. Digital methods for cross-platform analysis. *The SAGE handbook of social media* 91–110 (2017).
61. Bovet, A., Morone, F. & Makse, H. A. Validation of twitter opinion trends with national polling aggregates: Hillary clinton vs donald trump. *Sci. Rep.* **8**, 8673 (2018).
62. Cardoso, F. M., Meloni, S., Santanchè, A. & Moreno, Y. Topical alignment in online social systems. *Front. Phys.* **7**, 58 (2019).
63. Coan, T. G., Boussalis, C., Cook, J. & Nanko, M. O. Computer-assisted classification of contrarian claims about climate change. *Sci. Rep.* **11**, 22320 (2021).
64. Zollo, S., Cinelli, M., Etta, G., Cerqueti, R. & Quattrociochi, W. Inference of social media opinion trends in 2022 italian elections. *Expert Syst. Appl.* **269**, 126377 (2025).
65. The Economist. Youngsters are avoiding facebook-but not the firm's other platforms (2019). Accessed: 2023-10-08.

66. Budak, C., Nyhan, B., Rothschild, D. M., Thorson, E. & Watts, D. J. Misunderstanding the harms of online misinformation. *Nature* **630**, 45–53 (2024).
67. Di Bona, G. et al. Sampled datasets risk substantial bias in the identification of political polarization on social media. *arXiv preprint, arXiv:2406.19867* (2024).
68. Allgaier, J. Science and environmental communication on youtube: Strategically distorted communications in online videos on climate change and climate engineering. *Front. Commun.* **4**, 36 (2019).
69. Yang, K.-C. et al. The covid-19 infodemic: Twitter versus facebook. *Big Data Soc.* **8**, 20539517211013860 (2021).
70. Mekacher, A., Falkenberg, M. & Baronchelli, A. The systemic impact of deplatforming on social media. *PNAS Nexus* **2**, pgad346, <https://doi.org/10.1093/pnasnexus/pgad346> (2023). <https://academic.oup.com/pnasnexus/article-pdf/2/11/pgad346/53180059/pgad346.pdf>.
71. Bernal, J. L., Cummins, S. & Gasparrini, A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int. J. Epidemiol.* **46**, 348–355 (2017).
72. Li, Y. et al. Impacts of forest loss on local climate across the conterminous united states: Evidence from satellite time-series observations. *Sci. Total Environ.* **802**, 149651 (2022).
73. Ferron, J. & Rendina-Gobioff, G. Interrupted time series design. *Encyclopedia of Statistics in Behavioral Science* (2005).
74. Fisher, R. A. *Statistical methods for research workers*. 5 (Oliver and Boyd, 1928).
75. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **19**, 716–723 (1974).
76. Schwarz, G. Estimating the dimension of a model. *The Annals of Statistics* 461–464 (1978).
77. Cruz, M., Bender, M. & Ombao, H. A robust interrupted time series model for analyzing complex health care intervention data. *Stat. Med.* **36**, 4660–4676 (2017).
78. Blanchard, O. J. & Quah, D. The dynamic effects of aggregate demand and supply disturbances. *Am. Econ. Rev.* **74**, 655–673 (1988).
79. Ravn, M. O. & Uhlig, H. On adjusting the hodrick-prescott filter for the frequency of observations. *Rev. Econ. Stat.* **84**, 371–376 (2002).
80. Canova, F. Detrending and business cycle facts. *J. Monet. Econ.* **41**, 475–512 (1998).
81. Media Bias/Fact Check. Fake news - media bias/fact check. <https://mediabiasfactcheck.com/fake-news/> (2024). Accessed: 2024-10-08.
82. NewsGuard. Rating process & criteria. <https://www.newsguardtech.com/ratings/rating-process-criteria/> (2024). Accessed: 2024-10-08.

Acknowledgements

The work is supported by IRIS Infodemic Coalition (UK government, grant no. SCH-00001-3391), SERICS (PE00000014) under the NRRP MUR program funded by the EU NextGenerationEU project CRESP from the Italian Ministry of Health under the program CCM 2022, PON project ‘Ricerca e Innovazione’ 2014-2020, PRIN Project MUSMA for Italian Ministry of University and Research (MUR) through the PRIN 2022CUP G53D23002930006 and EU Next-Generation EU, M4 C2 I1.1, and project SEED PNR n. SP122184858BEDB3. M.F. acknowledges funding from the EU through the MSCA Postdoctoral Fellowship “POLARCLIMATE”, grant number 101148808.

Author contributions

S.S., M.F. and M.C. downloaded the data; S.S. and M.C. analysed the data; S.S., M.F. and M.C. conceived the study. All authors participated in the discussions, wrote and revised the manuscript.

Declarations

Competing interest

Authors have no competing interest to declare.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-03082-9>.

Correspondence and requests for materials should be addressed to M.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025