University of Zurich[UZH]

Trends in European Climate Change Perception: Where the Effects of Climate Change go unnoticed

GEO 511 Master's Thesis

Author
Philipp Graf
16-711-988

Supervised by
Prof. Dr. Ross Purves

Faculty representative
Prof. Dr. Ross Purves

30.01.2022
Department of Geography, University of Zurich

Universität Zürich[UZH]

Master's Thesis • Department of Geography, University of Zurich

Trends in European Climate Change Perception:
Where the Effects of Climate Change go unnoticed

Philipp Graf

16-711-988

Supervisor and Faculty Member:

Prof. Dr. Ross Purves

January 2022

Acknowledgements

This master’s thesis marks the end of my studies at the University of Zurich, a time of discovery and learning made possible the many professors and fellow students accompanying me on this journey.

I would like to express my special thanks to my supervisor *Prof. Dr. Ross Purves* for his excellent guidance in the entire planning and writing process of this thesis. The many pleasant meetings we had were always fruitful for my progress and the expert inputs and feedback provided helped me tremendously.

To my *parents* and *my beloved partner*, I have infinite gratitude for always believing in what I do and being there for any support I needed. Whether I needed to take my mind off this thesis or required encouragement, they were always there for me to fall back on.

Abstract

Climate change threatens global impacts in a variety of domains that must be limited by adaptation and mitigation measures. The successful implementation of such policies can strongly benefit from the general public's cooperation motivated by their own risk perceptions. Public participation can be promoted by tailoring policies to the populations they affect, which in turn results in the need for a deeper understanding of how different communities interact with the issue of climate change. Social media platforms such as the microblogging service Twitter have opened unprecedented opportunities for research on public perception in recent years, offering a continuous stream of user-generated data. Simultaneously, they represent a crucial discursive space in which members of the public develop and discuss their opinions and concerns about climate change.

Subsequently, this thesis gains insight into the characteristics of public reactions to individual climate change effects and processes by investing corresponding corpora of tweets spanning a decade. For seven western European countries, the spatial, temporal, and thematic reaction patterns are determined with a further assessment of the drivers behind each finding. Tweets are collected, classified, georeferenced, and clustered using a selection of Geographic Information Retrieval as well as Natural Language Processing methods before being analysed regarding thematic trends in their content, spatial distributions and influences of environmental factors, as well temporal distributions and impacts of real-world events.

The findings illustrate diverse climate change perceptions that vary across spatial, temporal, and thematic dimensions. Communities tend to focus more on issues relevant to their local or national environment, leading populations to develop a certain degree of specialisation for these aspects of climate change. This typically coincides with a substantially more domestic discourse on the subject and a decrease in interest for corresponding international events. In a similar sense, the tangibility of an event drives the magnitude of reactions. However, while more tangible events are more frequently recognised and discussed, less tangible events tend to be more frequently attributed to climate change as the public shifts their focus from immediate impacts on the personal scale to impacts on the global scale. Additionally, traditional news media are shown to retain a high level of control over science communication and the climate change discourse on Twitter, likely influencing the public's perspective on global warming. Individual real-world events such as major climate conferences and scientific releases only occasionally elicit strong public reactions when they are topically related to an event type, whereas global protests can lead to significant discussion across various event types. Inversely, global crises such as the COVID-19 pandemic significantly reduce public concern about climate change processes.

Table of Contents

Table of Figures

Table of Tables

Acronyms

ADM1	First-order administrative division of a country
ADM2	Second-order administrative division of a country
API	Application Programming Interface
COP	United Nations Climate Change conference (Conference of the Parties)
ECV	Essential Climate Variables
FFF	Fridays for Future
FUA	Functional Urban Area
GCOS	Global Climate Observing System
GDP	Gross Domestic Product
GHG	Greenhouse gas(es)
GIR	Geographic Information Retrieval
GIS	Geographic Information System(s)
HDI	Human Development Index
HHI	Heat Health Index
IPCC	Intergovernmental Panel on Climate Change
IR	Information Retrieval
LDA	Latent Dirichlet Allocation
LTET	Long-term event type
MAUP	Modifiable Area Unit Problem
MRSD	Mean regional standard deviation
NER	Named Entity Recognition
NGO	Non-governmental organisation
NLP	Natural Language Processing
OECD	Organisation for Economic Co-operation and Development
PERMOS	Permafrost Monitoring Service
RCP	Representative Concentration Pathway
RSD	Regional standard deviation
SLR	Sea level rise
STET	Short-term event type
UGC	User-generated content
VGI	Volunteered Geographic Information
WGMS	World Glacier Monitoring Service
WMO	World Meteorological Organisation

Trends in European Climate Change Perception:

Where the Effects of Climate Change go unnoticed

PHILIPP GRAF

Department of Geography, University of Zurich, Switzerland

1. Introduction

1.1 Motivation and research gap

Human-made climate change caused by greenhouse gases and further anthropogenic forcings (IPCC, 2014: 6) is arguably one of the most pressing issues of the 21st century, not only in geography, but across many disciplines (Haines et al., 2020: 1; Javadinejad et al., 2019: 10; Kerr, 2007: 1231). Its effects are observable in a variety of natural processes including physical systems (e.g. floods, droughts, erosion, glacier shrinkage), biological systems (e.g. wildfires) and also human systems (e.g. food production, health) (IPCC, 2014: 49–52). In order to contain the "*severe, pervasive and irreversible impacts for people and ecosystems*" resulting from uncontrolled climate change (IPCC, 2014: 8), adaptation and mitigation strategies are proposed to regulate and minimize the risks of climate change. The adoption of climate policies is heavily influenced by the risk perceptions of not only governments and organizations, but also individuals, which can present a constraining factor on the implementation of policies and measures (IPCC, 2014: 17, 27, 95; O'Connor et al., 1999: 469–470).

Public recognition, cooperation and initiative are thus key factors when attempting to implement climate policies in order to mitigate the effects of climate change (Hansen et al., 2012: E2415; IPCC, 2014: 27; Lorenzoni et al., 2007: 446; Semenza et al., 2008: 479). Crucially, however, it has been shown that public interest on the topic has generally decreased in many parts of the world recently (Ratter et al., 2012: 6). As such, it is essential to investigate how the public experiences and reacts to climate change to gauge the effectiveness of certain action plans (Howe et al., 2012: 1).

This public climate change perception is, however, not entirely uncomplicated as it is affected by an overwhelming variety of influences and contextual factors including the fact that information regarding the subject is often assimilated by individuals to fit their worldviews and political values. The decision-making process surrounding climate policies can thus strongly benefit from the consideration of societal values, circumstances, and interests expressed differently by local population groups (IPCC, 2014: 19, 26; Ruiz et al., 2020: 112). Any type of climate policy from local to national must be tailored to these characteristics of the target population in order to maximize citizen engagement (Lee et al., 2015: 1019; Lorenzoni et al., 2007: 454; Whitmarsh, 2011: 698).

Beyond the general characteristics of how the public reacts to climate change, it is crucial to identify *why* these trends materialise (Ruiz et al., 2020: 112). The current academic understanding of such climate change perception drivers remains 'modest' according to Ruiz et al. (2020: 112), with the literature being rather scattered across methodologies and spatiotemporal areas of interest. A comprehensive review of theses drivers on an extensive spatial scale is considered desirable, especially for decision-makers.

The manifestations of *how* and *why* people react to climate change can be captured through a variety of media and methodologies (e.g., Lorenzoni et al., 2007; Ratter et al., 2012; Semenza et al., 2008; Whitmarsh, 2011). With the emergence of microblogging services such as Twitter, there is now an unparalleled availability of user-generated content representing individuals' opinions and perspectives, offering unprecedented opportunities for detecting environmental events (e.g., de Bruijn et al., 2019; Ostermann & Spinsanti, 2012; Sakaki et al., 2010) and – importantly for this thesis – the research on the public perception of such events (Atefeh & Khreich, 2015: 132–133, 138; Hu, 2018: 1–2). Individuals across the globe are already correctly identifying recent changes in their surrounding natural systems such as local temperature anomalies and weather patterns (Lee et al., 2015: 1017).

The considerable diversity and uncertainty in such public discussion of climate change on social media specifically can threaten to impair the support for mitigation and adaptation strategies achieved through political action. This highlights the ever-increasing importance of understanding the online climate change discourse (Veltri & Atanasova, 2015: 721; Williams et al., 2015: 126).

Up until a few years ago, literature regarding the climate change discourse on Twitter was still rather thin, even though opinion making on social media is expected to have a large impact on public climate change perception, not only through direct influence but also through the perpetuation of the discourse into the 'offline' sphere (Brossard & Scheufele, 2013: 41; Kirilenko & Stepchenkova, 2014: 171; Williams et al., 2015: 127, 136–137). Especially research taking into account a multitude of climate change indicators is needed in the context of climate change perceptions, as studies using either climate change in its entirety or singular effects (e.g., local temperature variations) are insufficient to fully grasp perception-climate relationships due to their one-dimensional nature (Hamilton & Keim, 2009: 2351; Taylor et al., 2019: 158). Studies examining the long-term influence of climate indicators on the media and subsequently the public are furthermore rare (Schäfer et al., 2014: 153).

1.2 Research aims

The main objective of this thesis is to contribute findings to the climate change perception discourse by determining the different characteristics of the public's perception during the past decade for a variety of climate-change-related environmental processes in western Europe. These insights should help inform more closely tailored climate policy approaches in the future.

With the above-mentioned importance of public awareness, it is imperative to gain more knowledge about the ways in which the public interacts not only with climate change as a whole, but also how this interaction varies spatially, temporally and in regard to different natural processes. This will allow for the tailoring of programs aiming to improve climate change awareness to these specific characteristics (Lee et al., 2015: 1019).

Using social-media-based Volunteered Geographic Information (VGI) related to various environmental processes representing the effects of climate change (e.g., temperature anomalies and glacier shrinkage), this thesis aims to investigate trends of public perception and interaction regarding these individual effects and thus climate change as a whole. The goal is to detect *what* (type of), *where,* and *when* public reactions regarding climate change events are elicited and conversely in which instances the effects of climate change go unnoticed or hardly discussed by the public. This should identify both aspects of climate change where the public already shows concern and might be susceptible to climate policies and where knowledge and/or concern could be improved or incited. This could give indication to policy-makers and others seeking to push behavioural change in society regarding climate change where policies might already be fruitful and where potential for increased awareness and education lies. Furthering public understanding on climate issues is argued to be a crucial part of successfully developing climate action (Lee et al., 2015: 1014).

As a result of the research gaps and aims described above, my thesis will investigate the following research questions to determine *how* and *why* the public reacts to climate change:

> *RQ1: What are the spatial, temporal, and thematic patterns eliciting public reactions to climate change effects and processes on Twitter?*

> *RQ2: What are the driving forces behind these public reaction patterns?*

The posed research questions imply two hypotheses that will be investigated in this thesis. The first hypothesis *H1* insinuated by *RQ1* states:

> *H1: Spatial, temporal, and thematic variability in the European public's reactions to climate change-related events is observable.*

Literature suggests that public reactions on Twitter related to climate change should indeed vary spatially, temporally, and thematically across different types of climate change processes (e.g., Abbar et al., 2016: 9; Frondel et al., 2017: 174; Kirilenko & Stepchenkova, 2014: 174–178; Ratter et al., 2012: 6; Stier et al., 2018: 1918; Taylor et al., 2019: 158; Veltri & Atanasova, 2015: 13). A second hypothesis *H2* relates to the reasons *why* some of the observations from above might occur:

> *H2: The effect of different climate change perception drivers on the observed trends varies.*

Aspects such as education, politics, media, and local weather should be expected to be among the most dominant drivers (Borick & Rabe, 2010; Lee et al., 2015; Ruiz et al., 2020; Whitmarsh, 2011; more in section 2.2.3), but their impacts are seldom quantified for different spatial, temporal, and thematic extents. To summarise, this thesis analyses *how* and *why* people on Twitter react to various climate change-related events across space and time.

1.3 Outline

In *Chapter 2*, the state of the art regarding academic literature on the fundamental concepts is provided to introduce background information of important climate change processes, the general public's perception thereof, and how these reactions can be captured and spatially contextualised using Twitter.

Chapter 3 focusses on the methodology used in this thesis, first explaining the choice in study parameters, and presenting the major data sources before giving a more detailed insight into the separate approaches used to process and analyse the data.

Chapter 4 displays the results of the various analyses in the thematic, spatial, and temporal dimensions detailing how the public reacts to climate change processes on these different levels.

Chapter 5 summarises and discusses the results presented in the previous chapter regarding existing literature and potential weaknesses or biases in the methodology used.

Chapter 6 reflects on the contributions of this thesis and offers paths for future research to build on the findings of this study.

2. Background

Elaborating on the current state of the art, this chapter introduces the basic concepts the thesis is built upon by summarising the most essential knowledge found in the corresponding literature. This includes a short overview of the climate change effects investigated in this thesis, background on public climate change perception, and the fundamentals of conducting geographic research with Twitter data.

2.1 Climate change effects and processes

With the extensive range of wide-spread consequences threatened by human-made climate change (IPCC, 2014: 49–52), it is common-place to select a representative sample of effects embodying the overall process (e.g., Bojinski et al., 2014; Trewin et al., 2021). Choosing such environmental variables has been an important step in monitoring climate change (Karl et al., 1999: 6). The monitoring of processes is reported on frequently by various teams of scientist such as the Intergovernmental Panel on Climate Change (IPCC) and the World Meteorological Organisation (e.g., IPCC, 2014; 2019; World Meteorological Organization, 2019).

When selecting indicators representing climate change for this thesis, there are multiple sources to draw from. The climate variables used by the Global Climate Observing System (GCOS) consist of seven headline indicators, chosen in an attempt to capture as much of the climate systems as possible while keeping in mind data availability and timeliness (Trewin et al., 2021: E22). The headline indicators are based on the Essential Climate Variables (ECVs) – a set of indicators useful in determining the development of earth's climate and planning mitigation as well as adaptation measures (Bojinski et al., 2014: 1432). ECVs can be summarised in four categories: (1) Temperature and energy (*global surface temperature, global ocean heat content*), (2) ocean and water (*ocean acidification, sea level change*), (3) cryosphere (g*lacier mass balance, Arctic and Antarctic sea ice extent*) and (4) atmosphere (a*tmospheric* CO_2).

A selection of indicators – discussed in more detail in the following sub-sections – has been made to cover a variety of the factors proposed by the different sources. The indicator selection process itself is detailed in the methods (section 3.1.1). The following indicators presented in this section will form the basis for the public perception analysis carried out in this thesis.

2.1.1 Temperature (Heatwaves)

The average global temperature rise calculated between 1880 and 2012 amounts to approximately 0.85°C (IPCC, 2014: 2). Additionally, the increased occurrence of extreme weather events such as temperature anomalies is already a reality (Fig. 1): Hansen et al. (2012: E2417–E2418) found that the area of the planet covered by summers 3σ warmer than the base period average (1951-1980) had risen from as low as 0.1% during the base period to 13% in 2010, warranting a new category of "extremely hot summers" which was practically non-existent a few decades ago. Meanwhile, global area covered by "hot" summers ($> 0.43\sigma$) has increased by over 40% from 33% to 75% over the same time period. In Europe specifically, Zhang et al. (2020: 1) report an "*exceptional number of pronounced heatwave events*" since the beginning of the 21st century. Just recently, the heatwaves in June and July of 2019 broke temperature records all over Europe (Xu, Wang, et al., 2020: 3).

The trend will continue with the further development of climate change and the extremely hot summers ($+3\sigma$ anomaly) will become the norm with additional occurrences of $+5\sigma$ temperature anomalies, signifying a dramatic change in seasonal-mean temperatures. These findings also contest claims suggesting that the new extreme temperature events in recent times are solely due to phenomena such as blocking and La Niñas, which have – unlike the found anomalies – always been common (Hansen et al., 2012:

E2420). This conclusion is coherent with other research on the increased occurrence of extreme temperatures (IPCC, 2014: 7–8; Zhang et al., 2020: 7).

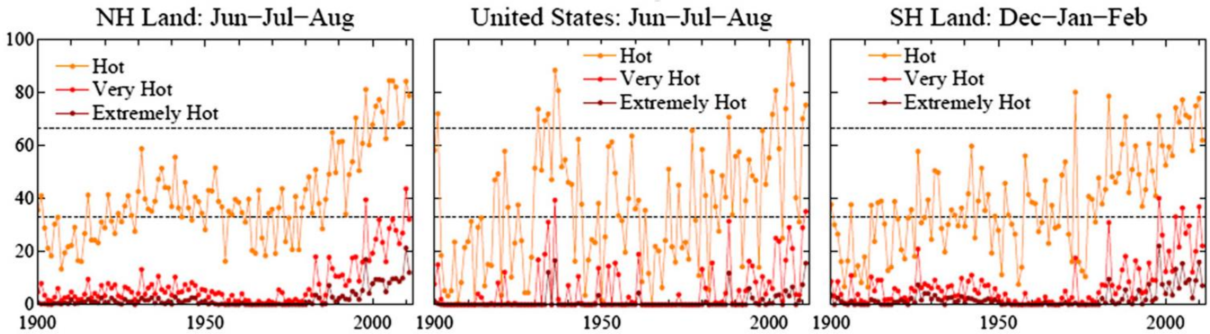

Fig. 1: Hot area percentages in summer for the northern hemisphere, the United States, and the southern hemisphere with categories defined as hot ($> 0.43\sigma$), very hot (2σ), and extremely hot (3σ) (Hansen et al., 2012: E2420). 1σ signifies a temperature anomaly 1 standard deviation above the historic mean (1950-1980).

Exposure to heatwaves can be measured using the health heat index (HHI), setting the level of dangerous heat at 40.6°C (105°F). Under a 2.0°C global warming scenario – as targeted by the Paris agreement – the area affected by dangerous heat is already 15.6% larger compared to a warming of 1.5°C. In the worst-case RCP 8.5 scenario, almost 80% of global land area could experience dangerous heat levels by 2100 (Sun et al., 2019: 128–129). These are important distinctions regarding the health risks associated with heatwaves, as such temperature anomalies are the leading cause of mortality due to weather (Xu, Wang, et al., 2020: 1). The 2003 European heatwave alone caused an excess mortality of around 50000 (Zacharias et al., 2015: 101). In this sense, even small increases in global warming can affect a large number of humans, as more and more regions become subject to dangerous heat levels.

Heatwaves are often defined as periods of unusually hot weather exceeding the 90th to 95th percentile of the local climatology derived from historical temperature data at the corresponding time of the year. Furthermore, to be considered a heatwave, the event should last a minimum of 3 to 4 consecutive days (Nairn & Fawcett, 2014: 228; Pezza et al., 2012: 211; Stefanon et al., 2012: 2).

2.1.2 Wildfires

The term wildfire refers to a wide array of vegetation fires including forest fires, bush fires, crop fires, and more (Bowman et al., 2020: 500). Climate variability can largely affect the occurrence of wildfires, which has been manifested in recent years by the unprecedented scale and duration observable in events covering the Amazon rainforest (2019/20), the Australian bush (2019/20) and Western North America (British Columbia 2017/18, Western United States 2018/20), indicating an increased risk in devastating wildfires (IPCC, 2014: 8; Xu, Yu, et al., 2020: 2173). In Europe, areas of Portugal, Spain, Greece, and Sweden suffered from large wildfires in recent years (Bowman et al., 2020: 501; Eriksen, 2020: 2). Projections show that the frequency as well the length of wildfires will further increase during the 21st century, with the extent of global warming having a major influence on the severity of this development (Fig. 2; Sun et al., 2019: 129–130).

Climate change not only increases the chances of wildfires by helping create highly combustible conditions and elongating the fire season (not least through increased occurrence of heatwaves), but is itself also likely accelerated due to greenhouse gas emissions and forest loss resulting from wildfires, possibly leading to a positive feedback loop (Sun et al., 2019: 129; Xu, Yu, et al., 2020: 2173). The permanent conversion of tropical forest to open lands for agriculture and real estate further contributes to climate change (Bowman et al., 2020: 501). Additionally, wildfires in permafrost regions threaten to accelerate its thaw, resulting in yet another increase in greenhouse gases (see also section 2.1.4; Bowman et al., 2020: 503; Gibson et al., 2018: 6).

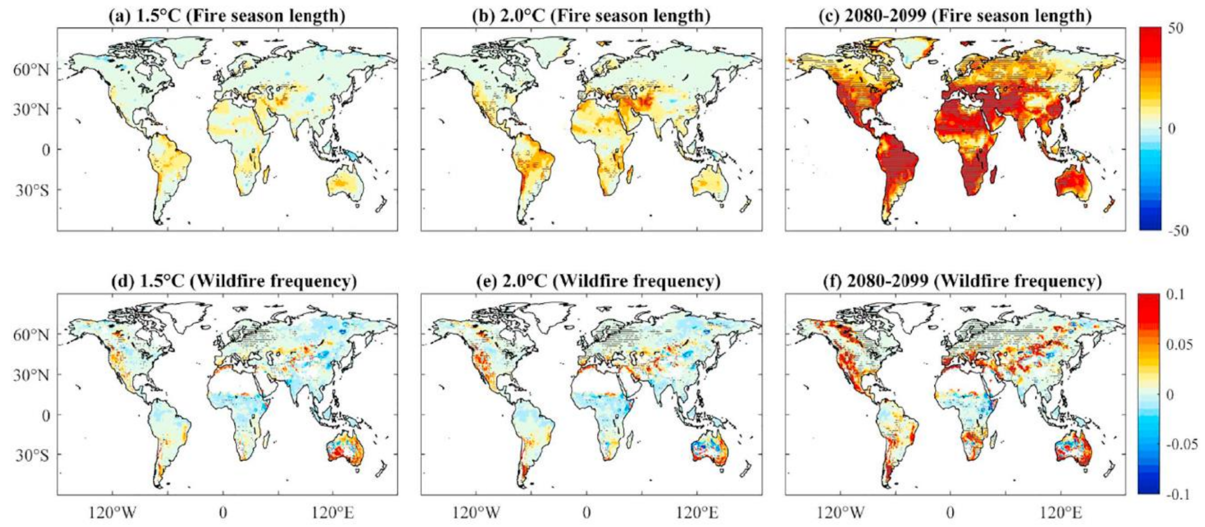

***Fig. 2:** Wildfire season length and frequency developments under different global warming models. The first column shows 1.5°C warming, the second column 2.0°C warming, and the third column the difference between the baseline (1981-2000) and the end of the 21st century (Sun et al., 2019: 130).*

Besides the environmental and often enormous economic impact (Bowman et al., 2020: 501), wildfires are a serious threat to human health. Increases in fire occurrence due to climate change are expected to lead to a rise of excess mortality and morbidity as a result of smoke inhalation, burns, and effects to mental health – with wildfire smoke becoming a more significant source of air pollution exposure for humans (Black et al., 2017: 192; Finlay et al., 2012; Reid et al., 2016: 1340). Air pollution generated by wildfires is argued to be more toxic than fossil fuel emissions (i.e. ambient air pollution; Dong et al., 2017: 272), highlighting the importance of wildfires independent from other emission sources. Furthermore, increased heat stress leading to more wildfires threatens global food security as the vulnerability of crops reaches its maximum during their reproductive period (Sun et al., 2019: 130).

It is worth noting that a study by Doerr and Santín (2016: 2) found increased reporting on wildfires both in popular media and scientific literature to not always align with increases in wildfire activity on longer timescales as reports base the assumption of change on rather short timescales. The recent fires mentioned at the beginning of this section have, however, occurred on unprecedented scales which might challenge the relevance of these findings in the current context (Bowman et al., 2020: 501; Xu, Yu, et al., 2020: 2173).

2.1.3 Glaciers and ice sheets

Glacier and ice sheet shrinkage due to climate change is an almost globally observable process with mass loss rates accelerating (Fig. 3; IPCC, 2014: 4; Zemp et al., 2015: 753–755). It is estimated that melting glaciers as well as the Greenlandic and Antarctic ice sheets contribute approximately 30% to the rising sea levels each, resulting in a combined contribution of around 60% (Gardner et al., 2013: 857). Glacier melt and sea level rise are therefore closely linked, which will further reveal itself in the analysis of public reactions. Increased glacial runoff as a result of global warming further has significant local impacts on the water cycle and hazard situations. Water availability is strongly governed by glacier contributions in drier climates, meaning that this dependency is expected to lead to significant water shortages in some regions once glaciers have retreated (Kaser et al., 2010: 20226). Environmental hazards related to glaciers – which are intensified by climate change – can arise from events such as glacier lake outbursts, mudflows, ice avalanches and rock fall from destabilised mountainsides (Kääb et al., 2003: 171).

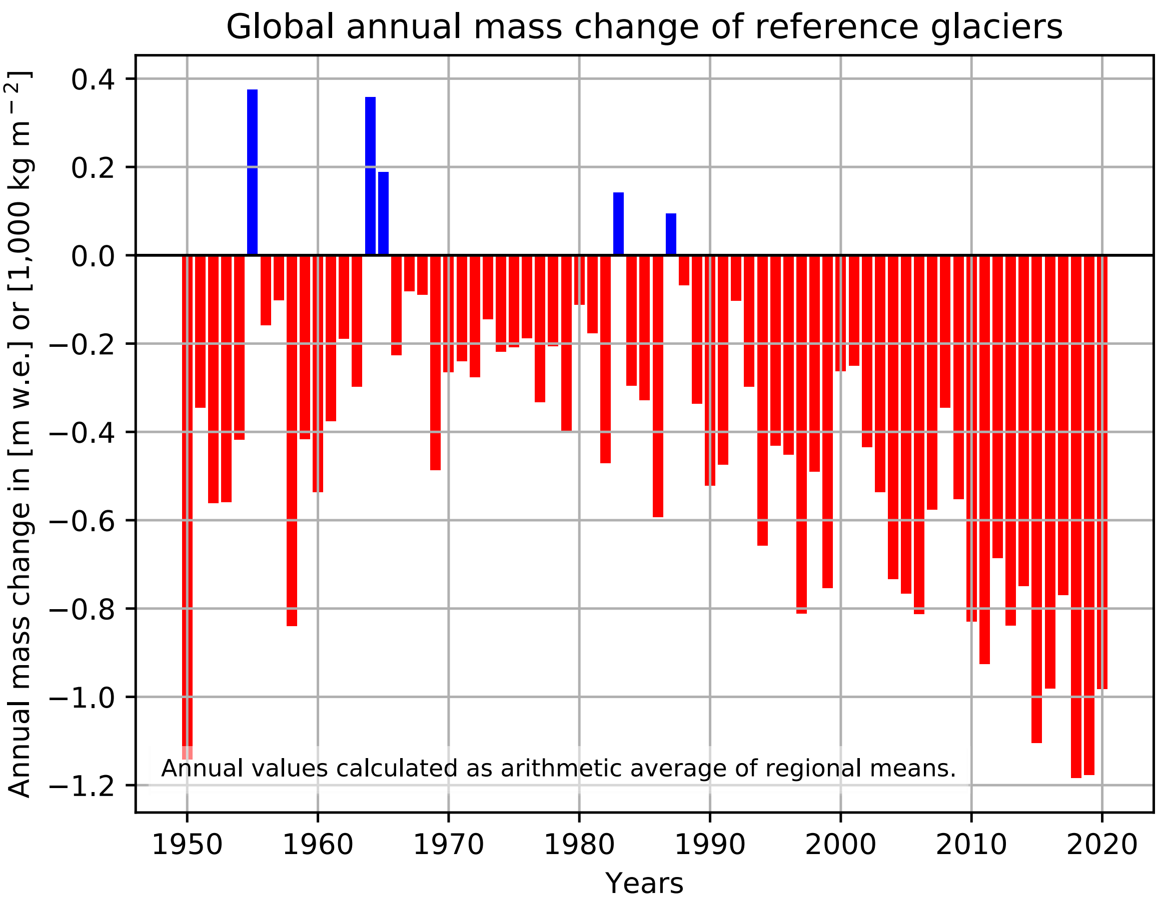

Fig. 3: Annual mass balance change in meters water equivalent averaged across reference glaciers with long-term observations (WGMS, 2021b).

Borick & Rabe (2010: 785–786) found that especially melting glaciers and polar ice had the largest impact on people's belief that climate change is real in the United States and suggest employing visualisations of these processes to confront climate change sceptics with particularly convincing evidence. In the same study, major climatic events such as hurricanes and droughts were found to have a similarly strong impact when an individual directly experienced it. Somewhat related, media articles claiming that there is insufficient evidence for global glacier retreat (or even suggesting that they are advancing) may be attempting to soften this impactful effect glaciers have on the public, as they support their conclusions with biased and globally insignificant datasets (Zemp et al., 2015: 754).

2.1.4 Permafrost

The continuously frozen permafrost ground covers large swaths of the polar regions and high mountain areas, importantly storing large masses of organic carbon. At the beginning of the 20th century, permafrost occupied approximately a quarter of global land area and although not densely populated, its development under climate change could have severe consequences affecting ecosystems, surface hydrology, vegetation as well as wildlife habitats, and the global carbon cycle (Jorgenson et al., 2010: 1220; Nelson et al., 2002: 206; Schuur et al., 2015: 171–172). While the extent of these consequence is still rather uncertain, increases in permafrost temperatures as well as thawing in conjunction with rising surface temperatures can already be observed, especially since temperatures in high-latitude areas have risen twice as fast as the global average (IPCC, 2014: 4; Schuur et al., 2015: 171).

The influence of climate change on permafrost is affected by both negative and positive feedbacks, somewhat complicating the relationship especially since they can outweigh the effect of global warming in some areas and lead to developments not explainable by changes in air temperature. Climate change is, however, expected to increase the overall occurrence of thawing (Jorgenson et al., 2010: 1230–1231; Schuur et al., 2015: 175). A recent study by Farquharson et al. (2019: 6687–6688) based on in-situ data from the Canadian Arctic found that permafrost thaw depths are already exceeding values expected in 2090 with the IPCC RCP 4.5 scenario, suggesting high-latitude permafrost degradation is highly susceptible to climate change. Wildfires, as mentioned before, are a further factor majorly impacting the development of permafrost by transforming the surrounding ecosystem (Jorgenson et al., 2010: 1231; Kirdyanov et al., 2020: 9).

Permafrost degradation adds heat gain to the affected areas due to changes in the surface albedo of the landscape and further emits CH_4 (methane) and CO_2 (carbon dioxide) to the atmosphere. While emerging vegetation could offset these emissions in the short term, it is expected that the long-term thaw of permafrost will turn it into a carbon source. During this century, the emissions released from permafrost will likely equal a tenth of fossil-fuel emissions, thus not having a dramatic impact on climate change development in the short term. Over much larger timescales, however, a positive feedback loop between greenhouse gas (GHG) emissions from permafrost thaw and climate change is possible (Jorgenson et al., 2010: 1233; Schuur et al., 2015: 176–178). Methane emissions are of special interest as the gas is far more potent in a climate change context compared to CO_2 and could start being released in significant quantities soon as some of the permafrost containing it is showing rapid warming. There are, however, large uncertainties surrounding the extent of this process and its effects on climate change are still poorly quantified (Anisimov, 2007: 6; van Huissteden, 2020: 433; O'Connor et al., 2010: 24).

2.1.5 Sea level

Hydrological systems will be subject to many developments in the context of climate change. In general, the increased occurrence of heavy precipitation events has become more likely with climate change, as have other hydrologic climate extremes such as droughts, floods, and cyclones. This trend is, however, not expected to develop uniformly across the globe (Hansen et al., 2012: E2422; IPCC, 2014: 8, 11). Regarding the ocean, population centres in coastal regions such the port cities of North-Western Europe are especially threatened by an increase in storms, hurricanes, storm surges and – on a larger time scale – sea level rise (SLR) threatening to result in coastal flooding (Hamilton & Keim, 2009: 2350; Kirezci et al., 2020: 4; Ratter et al., 2012: 4). Similar to other hydrologic systems, sea level rise will also not be uniformly impactful as relative sea levels continue to fall in areas undergoing post-glacial rebound (e.g., Baltic, Hudson Bay) and rise on subsiding coasts (e.g., India, China) (Nicholls & Cazenave, 2010: 1518).

Rising sea levels can on one hand be attributed to thermal expansion caused by ocean warming as well as ocean salinity changes and on the other hand an increased influx from inland water mass due to land ice melt (Milne et al., 2009: 472; Nicholls & Cazenave, 2010: 1517). Sea level rise is one of the major challenges for many nations presented by climate change and will require extensive mitigation as well as adaptation, as around 200 to 600 million people worldwide live in coastal flood zones depending on the definition (McGranahan et al., 2007: 22; Stern & Stern, 2007: 129). While a majority of this population resides in Asia – outside the study area – are the North Sea and Atlantic coasts of Europe among some of the areas most threatened by coastal flooding (Kirezci et al., 2020: 4).

While the focus of the general public often lies on the increase in global mean sea level rise (and its magnitude as well as accelerated development), the actual erosion and coastal flooding threats presented to many of the endangered areas develop from specific constellations of storm surges and wave setups temporarily increasing the water level beyond the results of sea level rise. Therefore, sea level rise can impact population centres on the time scale of mere hours (individual storm) up to multiple centuries (erosion and continued sea level rise). Climate change-driven developments in these aspects are projected to lead to a significantly increased frequency of episodic coastal flooding by 2100, turning current 1-in-100-year events potentially into 1-in-10-year events (Kirezci et al., 2020: 1, 7; Marcos et al., 2019: 4361). These impacts of climate change-induced sea level rise are, however, clouded in uncertainty as the adaption measures taken by means of building coastal defences will likely have a strong influence (Nicholls et al., 2008: 96). This influence is frequently omitted in sea level rise models, resulting in absolute worst-case scenarios being calculated (Nicholls & Cazenave, 2010: 3).

2.2 Climate change perception

Adaptation and mitigation measures necessary to combat climate change are not only subject to the risk assessments made by governments and organisations, but also heavily rely on the risk perceptions of the general public. As such, the implementation of these measures can greatly benefit from a population that recognises the dangers of climate change and is willing to cooperate (Hansen et al., 2012: E2415; IPCC, 2014: 27; Lorenzoni et al., 2007: 446; O'Connor et al., 1999: 469–470; Semenza et al., 2008: 479).

The public's perspective on the matter, however, is a tangled web of diverging interests, concerns and opinions, providing policymakers with difficulty when attempting to tailor policies to populations (Ruiz et al., 2020: 112; Whitmarsh, 2011: 691). Research on the public's stance towards the subject aims to untangle the mesh and offer a clearer picture to work with. Consequently, the following section will give an overview over public climate change perception and its key drivers and consequences.

2.2.1 On perceptions and reactions

Climate change perception can be defined as "*a state of opinion and awareness of anthropogenic climate change*" (Ruiz et al., 2020: 113), determining the level of recognition and concern individuals or communities display towards the subject. When investigating tweets, one might also consider the distinction between reaction and perception, as the term reaction is more closely tied with the reporting on events and perception more describing an attitude towards the event (Dunkel et al., 2019: 780–782). With tangible events such as heatwaves, reactions could be seen as an expression of perception, whilst events such as permafrost thaw might only be experienced through scientific or media reports and corresponding tweets thus rather qualify solely as reaction.

Additionally, some thematic aspects related to tweet contents (e.g., mentions of risk and danger when talking about an event) will be addressed in the analysis, meaning that this thesis sometimes walks the line between perception and reaction. I will, however, not employ any sentiment analysis during this process – which might be a common assumption of perception research in geography, especially when involving Twitter (e.g., Dahal et al., 2019; Lansley & Longley, 2016; Mitchell et al., 2013).

2.2.2 Current global trends

A wide range of studies has investigated the public perception of climate change on local and global scales. The two seemingly most commonly addressed statistics in the discourse are awareness and concern (or similarly perceptions of risks) related to climate change (e.g., Lee et al., 2015; Semenza et al., 2008). Most frequently, the data collection process involves telephone (e.g., Ratter et al., 2012; Semenza et al., 2008) or postal (e.g., Lorenzoni et al., 2007; Whitmarsh, 2011) surveys and interviews (e.g., De Longueville et al., 2010; Lorenzoni et al., 2007).

On the global scale, Lee et al. (2015: 1014) state that "*levels of climate change awareness, knowledge, perceived risk, and support for mitigation or adaptation vary greatly across the world*". Awareness about climate change is essentially universal in developed countries (Lee et al., 2015: 1015), as for example 92% of respondents to a survey in the United States had heard of the term before (Semenza et al., 2008: 481). In developing countries, awareness can be as low as 35% (Lee et al., 2015: 1015). Ratter et al. (2012: 4–6) found that concern about climate change *as a serious threat* had decreased by 9% over the study period in the city of Hamburg, Germany, whilst the opposite opinion of a small to non-existent threat had gained in popularity (+ 3%). In their review of concern in other localities, they found very similar trends in Great Britain, the European Union, Canada, the United States, Australia, and New Zealand, leading to the conclusion that the observed decline in concern is persistent across many parts of the (western) world. It is worth noting that although awareness is widespread, a variety of studies suggest a consensus on the fact that concern about climate change often ranks secondary to other issues (Ruiz et al., 2020: 114; Whitmarsh, 2011: 691).

2.2.3 Drivers of climate change perception

As stated above, general awareness and recognition of climate change by the public is imperative. Furthermore, it is also essential to gain knowledge about the various distinct ways in which individuals and communities perceive climate change and its many facets in order to enable a smoother implementation of climate policies (Ruiz et al., 2020: 112). A variety of factors (e.g., national, cultural and geographic; Lee et al. 2015: 1015) has been suggested to explain fluctuations in global climate change perception. For the purpose of this short overview, I will use the categorization of climate change perception drivers developed by Ruiz et al. (2020), who reviewed and analysed a wide variety of related studies. They distinguish seven different classes, which will be loosely followed in this section.

Media exposure

News, in general, are considered to have a strong influence on our perception of the world (Segev & Hills, 2014: 67). Traditional media is a common source of climate change information for the general public, meaning their interpretations of the topic are frequently accepted as truth. Meanwhile, online media present more effective ways of communicating climate science in a more direct and unbiased manner (Brossard & Scheufele, 2013: 40–41; Ruiz et al., 2020: 114). Twitter, amongst other social media, is increasingly used as a news source and trending topics on the platform are often aligned with news tweets (Shariff et al., 2017: 785). A study by Shariff et al. (2017: 794) found that readers generally assign rather high credibility ratings to news tweets and often fail to identify rumour tweets attempting to spread misinformation. Combined with the fact that international reporting characteristics differ from country to country (Segev & Hills, 2014: 68), this can lead to divergent climate change perceptions based on media and news coverage.

Media attention regarding climate change – including the separate climate change processes described above – has shown a strong increase during the 21st century, however can varying degrees of coverage be observed between different countries (Doerr & Santín, 2016: 2; Schäfer et al., 2014: 153, 165). Information fatigue associated with attention cycles is suggested to be a reason for general fluctuation in climate change perception over the course of a few years (Ratter et al., 2012: 6–7). This could be influenced by the media and the characteristics of its reporting on the topic at a given time, as well as political events and views (Ratter et al., 2012: 6). For example, media reporting on the topic of climate change has increasingly been perceived as ‘exaggerated’ and ‘alarmist’ by the general public (Borick & Rabe, 2010: 791; Lorenzoni et al., 2007: 452; Whitmarsh, 2011: 697).

Politics (Ethnography)

Investigating newspaper coverage of climate change, Schäfer et al. (2014: 169) found that politics were the most influential driver on the focus of newspapers on climate change, whilst weather and climate only had some, and the publication of scientific articles on the topic barely any influence. This would suggest that political actors play a key role in determining the climate change discourse, at least in traditional media. Politics – amongst other ethnographic aspects such as natural and cultural environments – also play a major role in an individual’s perception of climate change, with political affiliation having been found to be a key predictor of climate change awareness and concern (Borick & Rabe, 2010: 783; Ruiz et al., 2020: 114; Whitmarsh, 2011: 697). Especially the separate reasons for believing in climate change were divided by partisan affiliation and “*partisan differences regarding agreement with the statements about the validity of global warming*” were even found to “*transcend individual belief in this phenomenon*” (Borick & Rabe, 2010: 793). Furthermore, polarization in the media about climate change leads people to turn towards their political affiliation to form an opinion (Ruiz et al., 2020: 114).

Further possible variations due to political perceptions can stem from the very structure of various democratic systems: Climate change as a political discourse is influenced by the institutional design, leading to significant differences between climate change perceptions in ‘consensus’ (multi-party arrangement) and ‘majoritarian’ (winner-takes-all system) democracies. Findings suggest that the implications of this difference for climate change politics and perception are threefold: Majoritarian system generally lead to (1) more polarized positions with (2) a more diverse set of perspectives – largely due to higher levels of scepticism – and (3) a larger number of non-political actors prominently involved in the climate change discourse. Generally speaking, climate change is used less as a political tool in consensus democracies, whilst scepticism is employed in majoritarian democracies to create polarization and draw a clear line between party values (Häussler et al., 2016: 95, 98–100). This could very well explain some national differences in climate change perception.

Education

Although education is commonly found to be the top predictor of climate change awareness and concern (Lee et al., 2015: 1017; Ruiz et al., 2020: 113), evidence exists contradicting the assumption that public climate change perception is primarily and uniformly determined by the type and content of information individuals are confronted with (e.g., through education). Instead, ideology, political stance, and personal worldview are suggested to influence risk perceptions and the processing of such information more fundamentally. Diverging views on the issue are therefore not necessarily an indicator of varying cognitive abilities but can represent personal and societal values (Ruiz et al., 2020: 114; Whitmarsh, 2011: 691). However, the comparatively high importance of environmental values in younger generations could be attributed to more elaborate environmental education (Whitmarsh, 2011: 698).

Personal experience of local weather and climate

Recent and local weather has been shown to have a significant influence on public climate change perceptions (Borick & Rabe, 2010: 785; Lee et al., 2015: 1016). Local temperature trends were found to only be surpassed by politics as the top predictor for perceiving climate change after adjusting for demographic, socioeconomic and religious factors (Hamilton & Keim, 2009: 2351). This aligns with findings suggesting that individuals tend to focus on local, tangible environmental issues. However is the connection between local weather events and climate change not always made by the general public, even flood victims (Boudet et al., 2020: 72; Lorenzoni et al., 2007: 452).

Whilst a general rise in average temperature associated with climate change may be difficult to perceive due to its long-term development, a correlation can be found between characteristics of public perceptions and magnitudes of observed temperature changes (Howe et al., 2012: 1). As a result, there is also an observable difference in climate change perceptions depending on the season (cold or warm) a survey is conducted in (Howe et al., 2012: 3). Besides the difficult recognition of long-term trends, individuals’ successful perception of local climate change might be further hindered by the fact that heavy snowfall is commonly recognized as harsh winter conditions by the public, even when no extremely low temperatures are present (Hansen et al., 2012: E2419). This misconception can thus lead to the opposite conclusion during the cold season.

Illustrating the link between local weather (and its potential environmental threat) and climate change is therefore instrumental in raising concern among the public as these issues are not only more immediate to the individual, but also offer visible opportunities for climate action and its necessity (Lorenzoni et al., 2007: 452; Ruiz et al., 2020: 114). Provoking the realization that climate change can be personally experienced in the local area and thus reducing the perceived abstract nature of its risks and consequences is argued to significantly motivate concern about the subject (Howe et al., 2012: 1).

Demographics

General demographic such as age, gender and race are commonly found to be an insignificant driver on climate change perception (Ruiz et al., 2020: 114). Geographical variability in climate change perception within a country or region can be partially explained by differing perspectives of urban and rural areas, with the latter demonstrating more scepticism towards the topic (Whitmarsh, 2011: 694–696). The author's further findings however suggest that the correlation of rural location and a certain political affiliation simply makes the geographical variability a proxy for political variability in perception. Furthermore, individuals in rural areas are more likely to be reliant on high-carbon lifestyles, especially in the aspect of transportation, and thus show more reluctance towards lifestyle changes. This coincides with their general conservative-leaning tendency (Whitmarsh, 2011: 696, 698).

There are, however, still differing perspectives from region to region, as a study on climate change perception in the rural United States shows (Hamilton & Keim, 2009: 2348). Interestingly, a correlation was found between the perception in these rural areas and the ongoing winter temperature trend (Fig. 4), demonstrating higher concern about climate change risks in states that experience the most winter warming. The outliers on the other end of the spectrum can be explained by occurrences of extreme weather events such as floods as well as hurricanes in Mississippi and a large proportion of conservatives in Kansas, resulting in a higher and lower concern, respectively, than can be explained by the winter temperature trend (Hamilton & Keim, 2009: 2350). This reiterates the above stated assumption that local climate (and weather events) can influence an individual's perception of climate change and confirms that such perceptions can be strongly regionally dependent.

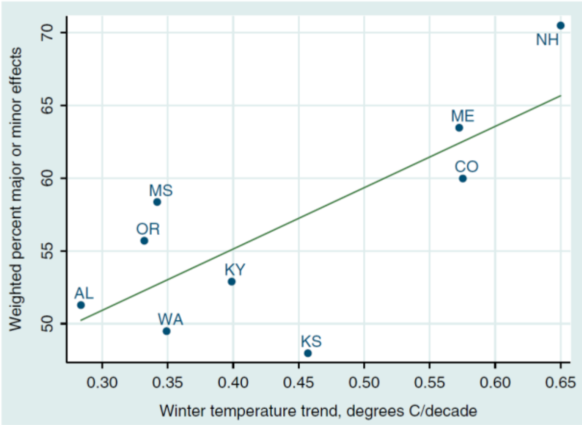

Fig. 4: Climate change risk perception vs. winter temperature trend in rural areas across 9 states of the US (Hamilton & Keim, 2009: 2350).

Findings by Semenza et al. (2008: 481) further describe a geographical variation in concern, with 90% of respondents in Portland, OR, United States versus 82% of respondents in Houston, TX, United States being very or somewhat concerned. Spatial dependence in perception is a generally observable trend (Hu, 2018: 7).

Wealth

Higher concern was also tied to lower income with the suggested explanation that higher-income individuals and communities have increased capacities to deal with the impacts of climate change through mitigation and adaptation (Ruiz et al., 2020: 114; Semenza et al., 2008: 481–482). This suggestion corresponds with findings that the detrimental consequences of climate change will have a greater impact on disadvantaged people and communities, for example eroding food security and increasing health issues besides displacement and financial suffering induced by natural disasters (IPCC, 2014: 13–16). Commonly used indices to assess sustainability on a national level such as the Human Development Index (HDI) and the Gross Domestic Product (GDP) have, however, been shown to be insufficient in predicting the global characteristics of climate change awareness and concern, highlighting the need for more specific and relevant social and cultural indicators tailored to the subject (Lee et al., 2015: 1019).

Interestingly, belief in climate change decreases during an economic recession as priorities are shifted (Ruiz et al., 2020: 114). The COVID-19 pandemic is one such shock to the global economy, but it also could impact the way people, governments and decision makers approach climate change policies in the future due to lessons learnt from unprecedented lifestyle and policy making situations (Hepburn et al., 2020: S374).

Influence of corporations

Corporations will frequently lobby for or against climate action depending on their own interests, aiming to influence both the media coverage and the public's opinion on the topic. Very often, such strategies are used by conservative groups as well as the energy and oil sectors to spread scepticism about the scientific consensus on human-made climate change. On the other end of the spectrum, groups such as non-governmental organisations (NGOs) make use of this approach to campaign in favour of climate action (Ruiz et al., 2020: 114–115). The effectiveness of the influence of corporations is significantly undermined by an individual's firm belief in the existence of anthropogenic climate change (Ruiz et al., 2020: 116).

2.2.4 Behaviour change and related barriers

Behaviour and lifestyle change is seen as a key climate change mitigation measure (IPCC, 2014: 28). The levels of awareness and concern about climate change strongly impact an individual's willingness and ability to change and adapt (Semenza et al., 2008: 480). Although general awareness and concern are widespread, behaviour change related to climate change is much rarer. This implies that awareness itself is insufficient to inspire public engagement and an understanding of the climate change risks is key in order to develop the concern necessary for the successful implementation of climate policies (Lorenzoni et al., 2007: 446–447).

Solely communicating the scientific evidence supporting the verity of climate change – although an essential part in raising general awareness – is not always sufficient to significantly change climate change perceptions (Ruiz et al., 2020: 114). Climate change is commonly perceived as a distant issue – both in space and time – in much of the developed world, which is further promoted by the fact that physical experiences of climate change effects are much rarer, due to for example widespread use of air conditioning (Howe et al., 2012: 1; Lorenzoni et al., 2007: 452; Ruiz et al., 2020: 114, 116). As a result, people's perceptions of the issue are much more easily influenced by the media and other drivers, as media for example has a much stronger impact on shaping opinions when topics are seen as distant and personally unrelated rather than local and immediate (Segev & Hills, 2014: 70). This highlights once more the importance of connecting people's perception of climate change with their local environmental processes.

Studies show that the most common barrier to behaviour change in terms of climate change is not knowing how or what to change in one's lifestyle, along with the assumption that one's individual actions will not make any significant differences (Lorenzoni et al., 2007: 451; Semenza et al., 2008: 483). Still, Semenza et al. (2008: 486) argue that "*findings from this and other studies suggest that the majority of consumers desire to be part of the solution to climate change*". This willingness to participate in adaptation and mitigation measures should lay the foundation for active citizen engagement once policies tailored to their interests and concerns are provided (Lee et al., 2015: 1019; Lorenzoni et al., 2007: 454; Whitmarsh, 2011: 698).

2.3 Using Twitter to investigate public reactions to climate change

Over the last decade, Twitter [1] has become a popular microblogging service, making it the "*microphone of the masses*" (Atefeh & Khreich, 2015: 132; Murthy, 2011: 779). This – especially in traditional media – unparalleled availability of user-generated content representing individuals' opinions and perspectives offers unprecedented opportunities for research on public perception (Atefeh & Khreich, 2015: 133, 138; Hu, 2018: 1–2). For instance, the extensive user base generates a large number of tweets when environmental events occur in their vicinity and the recognition of such perception bursts can be used to detect a large majority of certain event types that are observed by scientific or commercial methods, with an increased chance for larger-scale events (de Bruijn et al., 2018: 10–12; 2019: 8–9; Kirilenko &

[1] https://twitter.com

Stepchenkova, 2014: 172; Ostermann & Spinsanti, 2012: 32; Sakaki et al., 2010: 851). Especially of interest is the detection of events which go unreported by traditional methods, turning social media into a complementary source of information for the detection of natural disasters and the like (de Bruijn et al., 2019: 1, 9). On the flipside, the very tweets making up these detected events can be investigated to discover behaviour patterns surrounding such occurrences (e.g., Abbar et al., 2016; Holmberg & Hellsten, 2015; Kirilenko & Stepchenkova, 2014; Pearce et al., 2014).

2.3.1 User-generated content and Volunteered Geographic Information

User-generated content (UGC) comes in various shapes and covers a wide range of topics, which is – in contrast to VGI – neither necessarily geographic or explicitly volunteered (Schade et al., 2012: 809–810). Goodchild (2007: 212) coined the term *Volunteered Geographic Information* (VGI) to more precisely describe this geographic subcategory of user-generated content. Through a variety of platforms and media on the internet, large numbers of mostly untrained private citizens continuously contribute geographic information voluntarily, resulting in data with a wide variety in quality. Not only do Geographic Information Systems (GIS) profit from this timely, detailed, and freely available data, but the general public itself can greatly benefit from the resulting applications and services (Goodchild, 2007: 213–214; Goodchild & Li, 2012: 110). The amount of available VGI is expected to increase significantly in the near future (Ostermann & Spinsanti, 2012: 17). VGI is in many cases already of better quality than the often out-of-date traditional authoritative data due to the use of more accurate technologies (Goodchild & Li, 2012: 112).

VGI shines in the representation of local variations of geographic phenomena by capitalizing on the local knowledge and experience of the population. Especially the real-time reporting of current events and conditions is of great interest in disciplines such as disaster response. Tweets can offer a range of valuable information about the ongoing event and contribute to its detection in the first place (de Bruijn et al., 2018: 1; Goodchild, 2007: 218–220). Public recognition of recent changes in their surrounding natural systems such as local temperature anomalies and weather patterns is already observable. Increasing these types of first-hand experiences is expected to raise awareness about climate change, especially in developing countries (Lee et al., 2015: 1017).

VGI describing environmental events can be rather vague in the temporal, thematic and spatial dimensions, as it often only describes a small snapshot of a larger process and can seldom assign a very concise geographic location (De Longueville et al., 2010: 1). In fact, many geospatial concepts in general are inherently vague (Goodchild & Li, 2012: 111). Coupled with the large amounts of meaningless and irrelevant posts of social media as well as the rather unreliable and unverified nature of VGI items (especially in situations with a time constraint such as disaster response), decision-makers have shown a general reluctance to use VGI in the past (Atefeh & Khreich, 2015: 139; Ostermann & Spinsanti, 2012: 19). One might, however, apply the ‘wisdom of the crowds’ principle, suggesting that a large number of individuals eventually converge on the correct solution. This is especially relevant for event detection, where a larger number of independent reports proposes a larger likelihood for the veracity of the reports than a single mention (Goodchild & Li, 2012: 112).

In the context of climate change, VGI offers a promising research option as the environmental changes perceived by ‘citizen sensors’ (Goodchild, 2007) or ‘social sensors’ (Sakaki et al., 2010) can be related with the impacts caused by climate change (De Longueville et al., 2010: 2). Whilst these types of sensors might commonly have a negative connotation of subjectivity and bias, it is worth noting that traditional physical sensors are far from objective and human-independent (Schade et al., 2012: 808). Furthermore, VGI sensors have the advantage over traditional geographic information practices of being more up-to-date and richer in pre-processed content. However, a lack in meta-data and poor structure make quality control a challenging task (Schade et al., 2012: 808). A further advantage of using VGI is the non-intrusive nature of the data collection process, negating the influence of the researcher on the views expressed by the subjects (Kirilenko & Stepchenkova, 2014: 181).

2.3.2 Event detection from tweets

To investigate the conversation around a certain topic or event on Twitter, algorithms are developed to detect instances of tweets mentioning said aspects. Generally, events are detected on Twitter by a sudden burst in keyword frequencies, either in the incoming stream of data or the distribution of a pre-existing corpus (Atefeh & Khreich, 2015: 142; de Bruijn et al., 2019: 4).

Adopting Twitter users as social sensors results in much noisier data than a conventional physical sensor would produce, making detection based on keywords complex. Some keywords may have multiple meanings in normal language or may be used in metaphors and sayings (e.g., 'wildfire' and 'spread like a wildfire'). Natural language processing (NLP) algorithms can be used to filter these types of issues and classify tweets into relevant and irrelevant groups according to the intended use (Atefeh & Khreich, 2015: 139; de Bruijn et al., 2019: 3–4; Sakaki et al., 2010: 852). Some research carrying out event detection on Twitter relies on the assumption that only a single instance of an event can occur at any given time in the area of interest (Sakaki et al., 2010: 860), which will not apply for the analyses in this thesis. The detection of a variety of simultaneous events of the same type (e.g., wildfires) in different locations presents an additional challenge. For instance, localisation cannot be performed after detection, since simultaneous events of the same type need to be kept apart geographically (de Bruijn et al., 2019: 2).

The fact that Twitter users are not entirely independent sensors must further be considered, as the option to follow other people on the platform and interact with their content can lead to information diffusion. In the case of real-time environmental events, however, it was found that this effect was rather negligible and had no significant impact on the independence of the 'sensors' (Sakaki et al., 2010: 856). This may not be the case for event types that predominantly see peaks in online discussion upon the publishing of related news articles as the primary source shifts from a person's own personal experience to another's writing.

2.3.3 Geoparsing

To spatially contextualise tweets, they each need to be assigned a specific real-world location. Twitter users have the option of attaching a location to their tweets, however is this feature only used in approximately 1% of posts, therefore suggesting that this set of tweets is not representative of the entire Twitter population with it likely being biased towards more technologically advanced users (Kirilenko & Stepchenkova, 2014: 173; Lee et al., 2013: 499; Ostermann & Spinsanti, 2012: 21). Alternatively, the user's registered location ('hometown') on their profile can be accessed (if specified) and used as an approximation of the location of their tweets, as it is probable that the tweet was made in the general area of the respective hometown. This feature is, however, neither entirely reliable nor necessarily precise or unambiguous (de Bruijn et al., 2018: 6; Sakaki et al., 2010: 853). In a similar sense, one must further consider the difference between tweets *from* and *about* places and the related implications of using profile locations (MacEachren et al., 2011: 186). Furthermore, spatial references in text can refer to multiple locations or only parts of a geographical feature (Hu, 2018: 14). Thus, direct mentions of *toponyms* (placenames) in the plain text of a tweet are frequently used in spatial Twitter studies in a two-step process involving (1) toponym recognition and (2) toponym resolution (e.g., de Bruijn et al., 2018; Ostermann & Spinsanti, 2012). This geoparsing approach plays a central role in this thesis as it connects any insights gained from the tweets to real-world locations, allowing for a spatial analysis of trends.

Toponym recognition describes the identification of natural language text describing a geographic location or feature (also '*geo-text*') and allows for the linking of human experience and perception with locations (Hu, 2018: 14). *Toponym resolution* describes the act of linking this type of geographically related information to a uniquely identifiable feature in space by assigning it the corresponding coordinates or footprint (Goodchild, 2007: 215). The procedure plays a central role in a variety of applications aiming to derive information through social media geolocation, amongst other uses (Acheson et al., 2017: 311).

Gazetteers – datasets storing information about places such as name, type, location and more – are crucial during this process as they provide the necessary link between text and space. Since no top-down, authoritative global database of toponyms exists (Acheson et al., 2017: 311), research on multinational, continental, or global scales generally uses partly crowdsourced gazetteers such as GeoNames[2], which is arguably the most popular resource for placenames in current academic works (Acheson et al., 2017: 312; Kirilenko & Stepchenkova, 2014: 173). The non-uniform nature of such gazetteers, however, leads to a variety of issues related to coverage and balance. On one hand, a significant difference in coverage can be observed across country borders, whilst even very common natural features such as streams and hills are heavily underrepresented when compared to populated places, skewing the balance strongly in the favour of urban areas (Acheson et al., 2017: 319). It must be acknowledged that these issues can lead to side-effects such as geographical distributions observed in the analysis reflecting gazetteer properties as opposed to any meaningful spatial variation. Furthermore, when tweets contain toponyms that also appear in regular speech such as ‘Turkey’, all mentions of this word can be assigned to the location of the corresponding toponym (de Bruijn et al., 2018: 12).

One approach to circumvent some of the described issues is the toponym-based algorithm for grouped geoparsing of social media (TAGGS) proposed by de Bruijn et al. (2018: 2), which clusters tweets seemingly reporting on the same event and allows for spatial information to be shared between these related tweets instead of geoparsing each item individually as done traditionally. The TAGGS algorithm disambiguates the true location by assigning cumulative scores to the various candidates through a voting process (de Bruijn et al., 2018: 4, 6). In a similar sense, language can provide some indication for the approximate location of a tweet – especially in linguistically diverse regions – as a study on forest fires found that most of VGI on Twitter and Flickr was produced in the country’s main language (Ostermann & Spinsanti, 2012: 25). Therefore, locations matching the language of a tweet can be prioritised in the toponym resolution.

[2] http://www.geonames.org/

3. Methods

The methodology used in this thesis builds on a rich literature in Geographic Information Retrieval (GIR), Natural Language Processing (NLP), and a variety of spatial and temporal analyses. In this chapter, the extents of the thesis are first outlined, and a short description of the utilised data sources is given. The following sections then elaborate on the exact methods used in the implementation of this thesis to collect, process, and analyse the data.

3.1 Thesis extent

Before data can be collected and analysed, the extent of the thesis must first be defined in all dimensions. This includes both the study area and period, but also extends into the delineation of the thematic boundaries with the selection of the climate change effects to be investigated in the analysis. This chapter will shortly discuss the decisions made in the selection of all essential dimensions with a final section summarising the utilised analyses to define the methodological extent of the thesis.

3.1.1 Climate change indicators

For my final selection of indicators representing climate change in this thesis, a compromise between the IPCC's assessment of affected system, ECVs, and the GCOS headline indicators mentioned in section 2.1 was made to cover a variety of climate change processes, while maintaining the ability of humans to directly detect or picture said processes as a core value.

Taking the GCOS variables as a starting point, surface temperature, glacier mass balance and sea level are chosen to represent three of the four categories. Atmospheric CO_2 is discarded due to its rather abstract nature and instead permafrost is recruited as it will potentially play an important role in the atmospheric aspect of climate change – specifically on larger time scales – through its emission of both carbon dioxide and methane, the latter of which being a considerably more potent greenhouse gas compared to former (Anisimov, 2007: 2; O'Connor et al., 2010: 24). Permafrost is arguably more tangible than the CO_2 in the atmosphere and the consequences of its melting directly impact inhabitants' livelihoods in the affected regions (Nelson et al., 2002: 218).

While only two out of the three affected systems in the IPCC model directly relate to the natural environment, only the physical aspect has been covered by the four indicators so far. As a result, wildfires have been chosen to represent the biological systems and their faring under climate change.

To enable an analysis of reactions by the general public on Twitter to these environmental variables, the processes associated with them – and contributing to climate change – were selected. Therefore, the final set of climate change indicators used in this thesis is as follows:

- *Heatwaves* (representing extreme events related to surface temperature)
- *Wildfires*
- *Glacier and ice sheet shrinkage* (representing glacier and ice sheet mass balance)
- *Permafrost thaw/degradation* (representing atmospheric implications of climate change)
- *Sea level rise*

Some of the indicators are interrelated, with especially increased temperature – the centre point of global warming – affecting all other variables. Heatwaves specifically, though, are strongly linked to the occurrence of wildfires (Sun et al., 2019: 129) and glacier melt goes hand in hand with sea level rise (Gardner et al., 2013: 857). Additionally, wildfires in boreal forests cause increased permafrost thaw (Bowman et al., 2020: 503; Gibson et al., 2018: 6). These connections will be subject to investigation during the analysis.

3.1.2 Study area

The study area is largely determined by the extent of my language knowledge as many steps of the analysis require methods such as annotating tweets, therefore leading to the necessity of being familiar with the corresponding language. This results in a final study area covering large parts of Western Europe, including Great Britain and Ireland (English), Germany and Austria (German), Spain (Spanish), France (French) and Switzerland (German, French). Switzerland was selected despite Italian-language tweets not being included as a large majority of the country is represented through the available languages. For statistical purposes, the population of Italian-speaking regions will be excluded as their tweets are also not represented.

The exclusion of the Americas despite the lack of language barrier for a large majority of the region is designed to keep the scope of thesis focused and realistic, enabling a deeper investigation into the individual areas as a result of the smaller spatial extent. Lastly, small European countries such as Liechtenstein which are also covered by the four languages are further excluded due to an expected data scarcity in said areas.

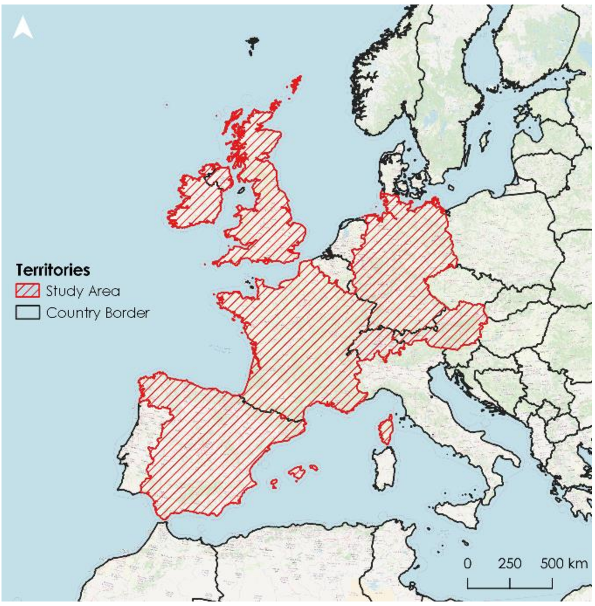

Fig. 5: Study area of the thesis.

Therefore, the study area is generally defined by the countries *Austria*, *France*, *Germany*, *Ireland*, *Spain*, *Switzerland* and the *United Kingdom* (Fig. 5). A further bounding box excludes any overseas territories of these states on or closer to a different continent than the main territory to retain the European focus. As some of the results are expected to be influenced spatially by environmental factors, this step is taken to increase homogeneity between the countries' populations' tangible experiences thereof, which should lead to more meaningful results for the European context. For mapping purposes, Annoni et al. (2001: 10) recommend to adopt the Lambert Azimuthal Equal Area (ETRS-LAEA) coordinate reference system for the statistical analysis and visualisation of pan-Europe data.

3.1.3 Study period

The study period is strongly tied to the availability of Twitter data, which dates back to its launch in 2006 followed by a rapid increase in users in the subsequent years (Sakaki et al., 2010: 851). As it took some years for Twitter to gain global popularity (reaching 500 million users in 2012 (Kirilenko & Stepchenkova, 2014: 171)) the opportunity presented itself to investigate the second decade of the 21st century spanning from 2011 to 2020. The year of 2021 was excluded as many analyses will be based on annual progressions requiring tweets for the full 12 months and data availability for such recent contextual data would be sparse as well.

3.1.4 Overview of analyses

As introduced with the research questions, this thesis investigates *how* and *why* people react to climate change-related processes in the thematic, spatial, and temporal dimensions. Tab. 1 offers an overview of the specific analyses conducted in this thesis to answer these questions. In a first step, I will determine *what* elicits a reaction from the public (thematic), *where* these reactions are elicited (spatial) and *when* they are elicited (temporal). In a second step, follow-up questions are asked to examine *why* the results from the first step materialised in the manner that they did. Further sections (3.4 – 3.6) will describe the methodology used for the individual analyses in more detail.

Tab. 1: Analyses carried out in this thesis. Chapters indicate where the corresponding methods and results are described.

Dimension	Name	Questing	Chapters	
Thematic	Themes	How do themes (e.g., ‘climate change’) influence the magnitude of the public’s reaction to an event?	M: R:	3.4.1 4.1.1-4.1.6
	Conversation characteristics	Which events and themes lead to the most conversation between users and which are the most controversial?	M: R:	3.4.3 4.1.7
Spatial	Spatial distribution	Where are reactions by the public elicited?	M: R:	3.5.1 4.2.1
	Spatiotemporal patterns	How do these geographic distributions change over time?	M: R:	3.5.2 4.2.2
	Transnationality	Events in which parts of the world elicit the strongest reactions?	M: R:	3.5.3 4.2.3
	Environmental context	How does the local, national, and global environment relevant to each event influence the public’s reactions?	M: R:	3.5.4 4.2.4
Temporal	Periodicity and seasonality	During what times and seasons is a reaction elicited?	M: R:	3.6 4.3.1
	User population changes	How does the population of users reacting to events change over time?	M: R:	3.6.1 4.3.2
	Event timelines	Which events elicited the strongest reactions during the study period?	M: R:	3.6.2 4.3.3
	Impact of key events	How did other events such as conferences, protests and the COVID-19 pandemic influence the reactions of the public?	M: R:	3.6.3 4.3.4-4.3.7
	M = Methods R = Results			

3.2 Data and software

Two groups of data were required for the analyses in this thesis. Tweets represent the primary data source as all public reaction results will be derived from them. There is, however, a need for an array of further data to contextualise and evaluate these results. Therefore, sections 3.2.2 and 3.2.3 will briefly summarise the collection and processing of additional data used in the analysis of the results data. This includes spatial data on one hand as well as background information such as demographics on the other hand.

3.2.1 Tweets

Tweets are character-limited (140 or 280 depending on year) posts made by users of the microblogging service Twitter. Essentially representing individual items of UGC or VGI in the presence of a spatial reference, tweets serve as the main data source for this thesis. They are retrieved from the Twitter database using an Application Programming Interface (API), which returns a requested number of tweets matching the specified keyword query. As the description of tweet retrieval methodology (chapter 3.3.1) will outline in more detail, tweet data retrieved from the API contains a wide array of associated information including public user data, therefore allowing for the analyses of various aspects such as spatial components (user’s hometown), information sharing behaviour (included URLs), content analysis (tweet text), and more (e.g., de Bruijn et al., 2019; Kirilenko & Stepchenkova, 2014; Moernaut et al., 2020; Pearce et al., 2014).

3.2.2 Supplementary spatial data

This thesis makes use of two different sets of spatial data: (1) General administrative boundaries and (2) contextual environmental data that materialises across space. The former will be discussed in the following, whereas the latter has a dedicated section in the spatial analysis chapter (3.5.4).

Beyond simple country borders, this thesis deals with data on a subnational level. Such first- and second-order administrative divisions (ADM1 and ADM2) are rather inconsistent across the study area. For all countries except the United Kingdom, ADM1 regions were selected to represent the subnational divisions. However, the proportions of region size and population to those of the country differ heavily between as well as within countries. In the UK, where ADM1 consists of England, Northern Ireland, Scotland and Wales, a higher resolution division was desired. A mixed approach based on the different commonly used ways to define the administrative divisions of the UK (OfficeforNationalStatistics, 2021) was taken to create similar types of divisions for each of these regions resulting in the following rules:

- *England:* Counties are used, whereby metropolitan districts are grouped with their corresponding county. Furthermore, the London boroughs are grouped into one Greater London area.
- *Northern Ireland:* Local government districts are used.
- *Scotland:* Council areas are used.
- *Wales:* Unitary authority divisions are grouped into counties.

Overall, this attempts to divide the study area as evenly as possible using existing administrative divisions of the countries. This should minimise the impact of the Modifiable Area Unit Problem (MAUP) by reducing the variability in the size of the aggregated units to achieve the most consistent results possible (Wong, 2004: 573–574). Nonetheless, its effects will have to be taken into account during the discussion of the results. The attributes such as population and administrative level of any spatial region in the thesis was either taken directly or derived from the GeoNames gazetteer (GeoNames, 2021).

Additionally, the concept of Functional Urban Areas (FUAs) is implemented when focussing on metropolitan areas. FUAs are a regions developed by the Organisation for Economic Co-operation and Development (OECD), which can be summarised as "*urban area[s] composed of densely inhabited urban core(s) and hinterland*" (OECD, 2012: 14). The dataset of region delineations (OECD, 2021a) is available alongside supplementary regional statistics on aspects such as population, economy, environment and more (OECD, 2021b).

3.2.3 Supplementary demographics data

To put the retrieved Twitter data into context, information about the number of Twitter users in each of the study area's regions must be approximated as a baseline (e.g., Hawelka et al., 2014: 4). To enable comparisons between regions, normalising values by population is essential. Whilst this data is not directly available for each year and country in the study, a series of datasets can be used to estimate the values.

Tab. 2: Datasets used to approximate yearly Twitter users for the countries in the study area.

Dataset	Description	Reference
Population	Country populations (yearly)	World Bank, 2021a
Internet users	Percentage of population using internet (yearly)	World Bank, 2021b
Twitter users (language)	Twitter users active in each language (yearly)	Alshaabi et al., 2021
Twitter users (country)	Percentage of internet users active on Twitter in each country (some years)	DataReportal, 2021

The existing survey-based data on Twitter user percentages per country (DataReportal, 2021) are not available for the entire span of the study period and changes in the survey procedure in 2017 break the data series in terms of continuous comparability. For each country, the 2019 value was chosen as a representative and then combined with further data for an approximation of each year's value.

The closest data to a continuous country-by-country, year-by-year Twitter user statistics is offered on a per language basis by Alshaabi et al. (2021: 11), which contains temporally high-resolution data based on roughly 10% of all tweets posted since September 2008. By defining the proportion contributed to these language totals by the individual countries, an estimation of the yearly number of Twitter users in each country is possible.

Thus, the ratio of Twitter users per country in 2019 (from DataReportal (2021)) versus Twitter users tweeting in the country's official language(s) (from Alshaabi et al. (2021)) in the same year was calculated. The assumption was made that changes in this proportion over the study period were limited. In a second step, this ratio could then be applied as a factor to the language data of all the other years in the study period to estimate the number of Twitter users in each country based on the number of Twitter users in each language that year. The additional country-specific data available for 2018 and 2020 was then further used to evaluate the performance of the factor and to add additional calibration to achieve close results.

Kirilenko & Stepchenkova (2014: 175) estimated the respective user bases in different countries by collecting random, keyword-free samples in regular time intervals from the Twitter API to generate a baseline dataset. This approach was, however, not suitable for my thesis as I was already making use of the entire 10 million tweet rate limit most months and thus could not afford to retrieve additional tweets.

3.2.4 Software

A variety of software was used when conducting this thesis. Tab. 3 gives a short overview of the different programs and their roles in the analysis. Large parts of the analysis were carried out using the Python programming language for algorithms and a password-protected PostgreSQL database for data storage. Spatial data was handled using the QGIS GIS application and the R programming language was occasionally adopted for specific statistical tasks.

Tab. 3: List of software used in the thesis and their corresponding tasks.

Tasks	Software
Data retrieval, processing and analysis, general algorithms	Python 3.9[3]
Spatial analysis, mapping	QGIS 3.20[4]
Data storage	PostgreSQL 13[5]
Additional statistics	R 3.6[6]

[3] https://www.python.org/
[4] https://www.qgis.org/
[5] https://www.postgresql.org/
[6] https://www.r-project.org/

3.3 Tweet collection and processing

The collection and processing of tweets represents a core part of this thesis as the entirety of the primary dataset is generated in this process. Thus, every step needed to be carefully considered and copious amounts of testing finetuned the final approaches. In essence, the process consists of tweet retrieval, filtering, and geoparsing (Fig. 6), which will all be discussed in the following sections. Lastly, the detection and clustering of events based on the collected tweets finishes this chapter.

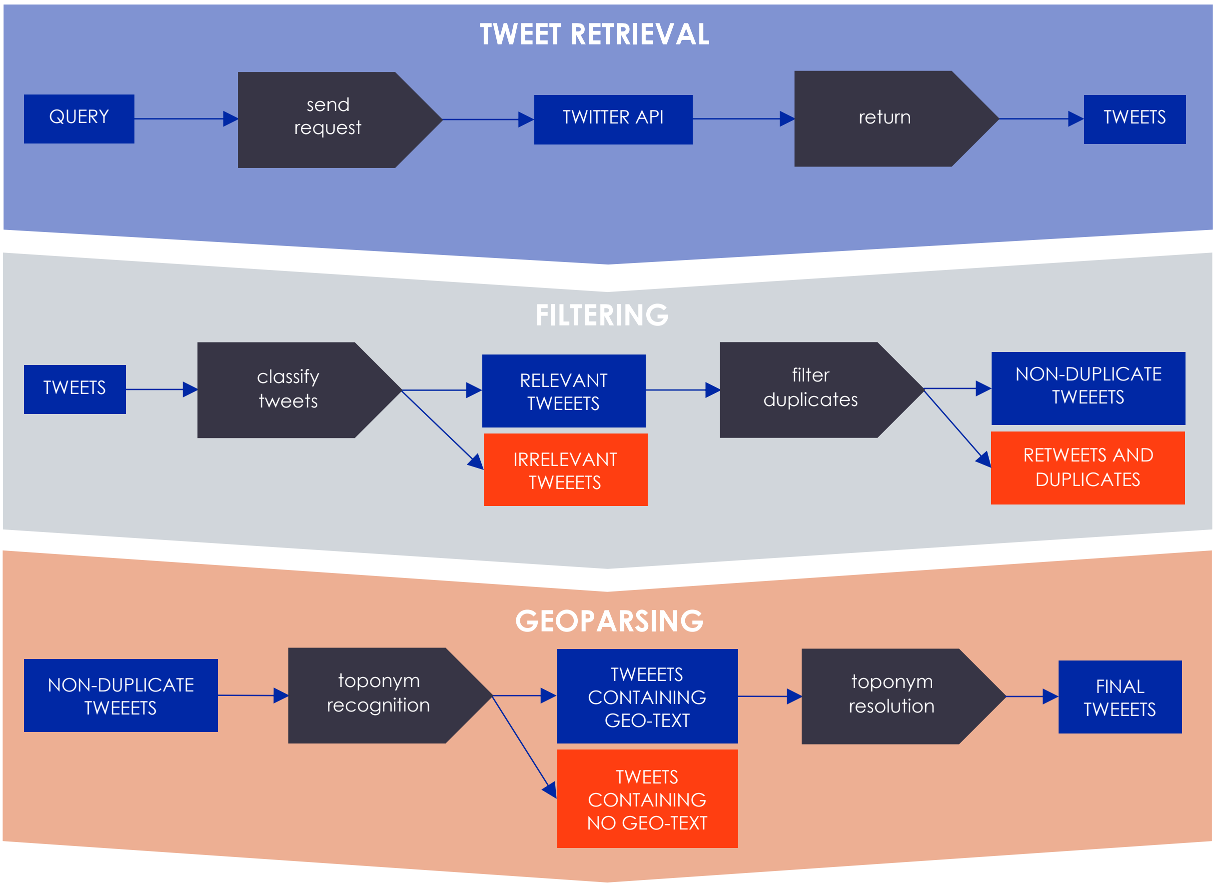

Fig. 6: Summary of the main steps and processes involved in retrieving, filtering and geoparsing the tweets ahead of analysis. Tweets belonging to the red boxes are discarded.

3.3.1 Tweet retrieval using Twitter API

The basic goal of an information retrieval (IR) workflow in a text-based analysis such as collecting tweets is to retrieve a set of documents from a large collection which match the given query (also referred to as information need) (Manning et al., 2008: 1). As such, queries form the foundation of the data collection in this context and are to be carefully constructed. In this thesis, the information need is determined by the events I am investigating with the goal of maximising the retrieval of tweets (*documents*) referring to said events whilst keeping the number of irrelevant tweets to a minimum. Finding this balance between precision and recall[7] is a multi-step process and begins with the query.

A query should therefore capture as many aspects of an event as possible (e.g., different types of wildfires or multiple terms describing the same process), whilst avoiding being too general and collecting

[7] Precision and recall determine the percentage of relevant documents in a corpus as well as the ratio of retrieved relevant documents compared to the total available number of relevant documents, respectively (Davis & Goadrich, 2006: 235).

tweets not related to the event in question. Queries can further be limited in their use of keywords to match the scale of the project (Parsons et al., 2015: 1222).

Tab. 4: Investigated climate change effects and processes along with English keywords used when querying the Twitter database. Alternate spellings, grammatical variations, and excluded terms are not included in this table to increase readability.

Event type	Keywords (English)
Heatwaves	*heat, heatwave, hot spell, hot day, hot weather, extreme temperature*
Sea level rise	*sea level, coastal flooding, permanent inundation*
Glacier shrinkage	*glacier, ice sheet [and further glaciologic terms describing land ice], ice melt, ice retreat, ice loss, glaciology*
Permafrost melt	*permafrost*
Wildfires	*wildfire, forest fire, bushfire, fire aircraft*

Tab. 5: Query parameters specified when connecting to the Twitter API. The meta fields contain values specified ahead of retrieval, whilst tweet and user fields solely form lists of information to be retrieved from the Twitter API about every tweet and the corresponding user (Twitter, 2021).

Query Parameter	Description
Meta fields specifying query	
`query`	Single-string query used to retrieve tweets matching the specified keywords. Has a maximum length of 1024 characters and can include operators such as AND, OR etc.
`max_results`	Number of tweets returned per request in the range between 10 and 500. For this analysis, the maximum of 500 was chosen as the default.
`start_time`, `end_time`	Start and end times of the query window in the form of UTC timestamps.
Tweet fields to be returned with results	
`id`	Unique ID of the tweet.
`text`	The written content of the tweet.
`created_at`	Exact timestamp of the date and time the tweet was posted by the user.
`lang`	Language of the tweet, as identified by Twitter.
`geo`	Location of the tweet if tagged by the user.
`referenced_tweets`	List of tweets referenced through a reply, retweet or quote by the current tweet. Returns both ID of referenced tweet(s) and type(s) of referencing.
`source`	Service or app used by the user to post the tweet.
`entities.urls`	URLs included in the tweet, both in shortened and expanded form
`public_metrics`	Engagement metrics (number of retweets, likes, replies and quotes) of the tweet at the time of retrieval.
User fields to be returned with results	
`author_id`	Unique ID of the user.
`username`	The user's unique screen name (e.g., "@twitter").
`name`	The user's name as specified by them. No unique constraints. Is often, but not necessarily, used to display the person's name.
`description`	The content of the user's profile description (or "bio").
`location`	If provided, the location specified by the user in freeform. Does not necessarily represent a real-world location.

To retrieve tweets from the entire Twitter archive, a connection to the `GET /2/tweets/search/all` endpoint of the API must be established (Twitter, 2021), which allows the user to specify a variety of parameters. These include the base query needed for keyword matching as well as a specification of attributes to be returned along with the matched tweets (Tab. 5). In the response to the API call, a maximum number of 500 tweets are returned per request, matching at least one of the queried words or phrases and not containing any excluded terms (specified with a "-" in the query). The academic research product track of the Twitter API used for this thesis allows 300 requests per 15 minutes and a total of 10 million tweets pulled from the full archive per month (Twitter, 2021).

Concerning these limitations and the arguments provided above, the queries in this thesis were designed to be stricter to enable the retrieval of tweets throughout the entire study period and across the study area whilst staying within the overall rate limit. Retrieving the entire population of tweets on a topic of interest is seen as favourable compared to a sample, especially to represent geographic as well as thematic minorities correctly (Kirilenko & Stepchenkova, 2014: 180), which is why the query was restricted instead of sampling tweets in intervals. In Tab. 4, a simplified list of all English keywords used in the analysis of the various climate change effects can be found.

The queries are rather diverse and differ not only between event types but also between languages, as some terms are composed and defined differently. All queries were tested and refined until a compromise was found between precision and recall, whilst also maintaining a high similarity between languages (i.e., not restricting queries in popular languages more than in others) and keeping track of the overall rate limit.

The data returned by the API in response to these queries can be divided into two domains: (1) The main data for this study relates to the tweets themselves, whilst (2) user data is used as supplementary information to determine attributes such as their general home area (Kirilenko & Stepchenkova, 2014: 172). Tab. 5 offers an insight in the attributes used in the retrieval ('meta fields') and the two return domains of (1) 'tweet fields' and (2) 'user fields'.

3.3.2 Tweet classification

Although the queries are designed to maximise precision, some compromises must be taken to maintain a decent level of recall. Therefore, the analysis relies on the classification and filtering of the retrieved tweets to ensure a satisfactory level of precision. The purpose of classification is to evaluate whether a tweet relates to the currently investigated event or not. As the size of the dataset far exceeds a number reasonable for the manual method of annotating each item individually, an automated approach has to be taken.

First, a rule-based algorithm scoring tweets based on keyword frequency and importance was implemented, which allowed for fine-grained adjustments of included keywords and their score values but was heavily disadvantaged by its rigidity. As a result, a strict version (Tab. 6, a) of this approach could achieve excellent precision but score very poorly in terms of recall, whilst a less strict version (Tab. 6, b) displayed mediocre precision with increased recall.

Second, a machine learning algorithm using the Random Forest Classifier from Scikit-learn for Python (Pedregosa et al., 2011) was implemented, which describes itself as a "*meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting*" (Scikit-learn, n.d.), thus using multiple machine learning algorithms to reach a classification. In the context of Twitter data, a Random Forest can be used to classify tweets into categories with the help of a wide array of features that include both the tweet itself as well as meta-information to provide the algorithm with more context (Schnebly & Sengupta, 2019: 507–508).

Training sets are annotated to teach the algorithm the differences between relevant and non-relevant tweets. This is no simple task as some tweets are hard to classify, even for humans (Sakaki et al., 2010: 853, 856). To improve the performance of the classifier, tweets are stripped of symbols and artefacts (e.g., "?", "RT", "@username") as well as common stop words are removed. The remaining words are tokenised (Lee et al., 2013: 502; Verma et al., 2011: 390). Applying this method while using the same sample from the rule-based algorithm as training data, a significant improvement in the F-score could be achieved with a much better balance between precision and recall (Tab. 6, c). Testing revealed a sample size of 500 tweets covering at least the extent of a year to be the minimum requirements to achieve good accuracy levels.

Tab. 6: Tested approaches of tweet classification. Note that the Random Forest values demonstrate the mean over ten different random seeds to account for the random nature of the classifier. All values listed are from classifications dealing with German-language heatwave tweets.

Approach	Precision	Recall	F-score
(a) Keywords, rule-based (strict)			
- Irrelevant	0.51	0.95	0.66
- Relevant	0.89	0.33	0.48
(b) Keywords, rule-based (less strict)			
- Irrelevant	0.60	0.37	0.45
- Relevant	0.64	0.82	0.72
(c) Random Forest Classifier (including date)			
- Irrelevant	0.86	0.69	0.77
- Relevant	0.85	0.94	0.89
(d) Random Forest Classifier (excluding date)			
- Irrelevant	0.61	0.68	0.64
- Relevant	0.81	0.75	0.78

The most important features during this training of the Random Forest were the content of the tweet itself and the time of the year a tweet was posted, making up 85% of the weights in the *heatwave – German* combination the Random Forest was first trained on (Tab. 7). Whilst the date therefore significantly contributes to the accuracy of the model, it established itself as highly problematic during further testing as all predictions made by the model had a strong temporal bias based on the dates of relevant tweets in the training data. Fig. 7 shows how the annual distributions of tweets roughly follow the trends of the 2019 distribution used as training data regardless of the fact that heatwaves occur at different points throughout the summer in these years.

Tab. 7: Features involved in the training of the Random Forest Classifier. The importance column displays values from the heatwave – German classification to give an idea of feature importance (before removing the date feature).

Feature	Description	Importance*
tweet_text	*tweet content (tokenised)*	0.526
date	*time of year (month & day)*	0.324
tweet_length	*number of characters in tweet*	0.043
avg_word_length	*average length of tweet content tokens*	0.033
n_words	*number of tweet content tokens*	0.025
source	*"Twitter for Android" etc.*	0.025
time	*time of day (hours & minutes)*	0.017
tweet_type	*regular, reply or quote*	0.004
text_links	*number of links in text*	0.002

As the environmental events investigated in this thesis are all temporally variable from year to year, the date feature had to be excluded, resulting in a small decrease in precision for relevant tweets and a larger decrease in recall (Tab. 6, d). Precision has priority over recall for this specific process as recall was already maximised during the tweet retrieval at the cost of precision. Therefore, the slight loss in precision was accepted in favour of removing the temporal bias.

It is worth noting that classifications for heatwave and wildfire tweets were the least accurate as users often use words related to heat and fire in context with events that do not constitute a heatwave or wildfire (e.g., a house fire), or in an entirely different context altogether (e.g., heat can be used in conjunction with cooking and baking). For the other event types, the associated terms are generally much less ambiguous and classification F-scores subsequently often reach values of 0.9 and above.

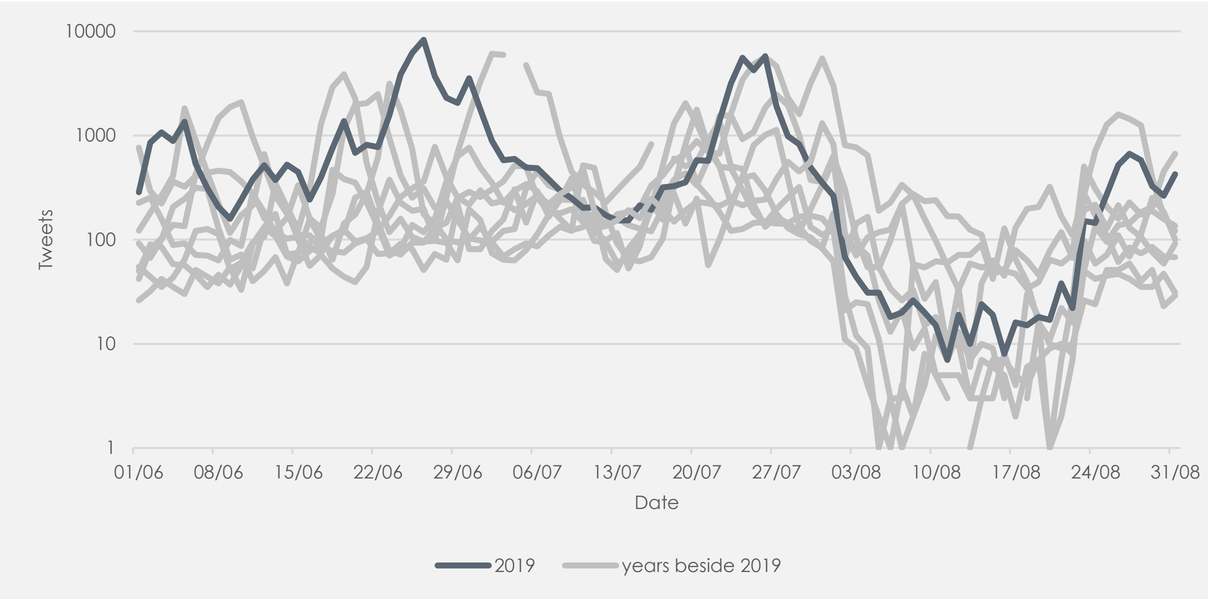

Fig. 7: Tweet frequency distribution for German-language heatwave tweets 2012-2020 demonstrating date bias incurred from classification. The y-axis has a logarithmic scale due to large differences in tweet frequencies between the beginning and end of the study period.

Machine learning models used to classify tweets about environmental events have the advantage that they can often be re-used with high accuracy for events of the same type without having to create training data for each individual event (Verma et al., 2011: 391) . This allows me to classify tweets in the entire study period based on one initial training set (per event type and language).

3.3.3 Tweet filtering

The main purpose of classifying tweets is of course to remove any items irrelevant to the analysis. In total, there are three criteria used to exclude tweets from the dataset:

1. ***Retweets and near-duplicate tweets***

Using tweets to detect events in the environment heavily relies upon the 'wisdom of the crowds' principle, suggesting that a large number of individuals eventually converge on the correct solution. More specifically, higher amounts of independent reports propose a larger likelihood for the veracity of the reports than a single mention (Goodchild & Li, 2012: 112). Therefore, any duplicates or near-duplicates of an original tweet such as retweets or other copies of the tweet's content add little reliability to the detection of an event as these reports are not independent (de Bruijn et al., 2019: 4). As a result, retweets were discarded from the dataset.

A further step was taken to scan the data for prolific users providing more than three relevant tweets in a single day. The contributions to user-generated content are not necessarily even across the population and prolific users providing a large amount of relevant information can thus skew the data analysis towards their characteristics and introduce bias (Hollenstein & Purves, 2010: 31–32). To limit the influence of any single user, any group of duplicate or near-duplicate tweets from an individual account was reduced to a single tweet of that group on a daily basis. This retains contributions relevant to the analysis (as a user might make multiple different observations about an event), whilst getting rid of (near-)duplicates and limiting the influence of prolific users.

2. ***Tweets containing no geo-text***

As the spatial dimension is an essential component of every analysis carried out in this study, tweets containing no recognised geo-text in the content of the tweet or the specified hometown are discarded.

3. ***Tweets classified as irrelevant***

   Tweets deemed irrelevant to the investigated event in the random forest classification process are discarded as they cannot be seen as a reliable source of information for the intended purpose.

3.3.4 Geoparsing

As introduced above (section 2.3.3), it is necessary to derive the location of a tweet from its content and associated user profile in the case of a spatial analysis as the number of tweets with a georeferenced location tag is extremely low (Lee et al., 2013: 499; Ostermann & Spinsanti, 2012: 21). In this thesis, toponyms identified in the tweet's content as well as the user's hometown location will be used for the analysis and thus enable me to broadly differentiate between the implied *from* (hometown) and *about* (tweet content) of a toponym (MacEachren et al., 2011: 186).

While using toponyms detected in the tweet's content to determine the location a user is tweeting *about* is common practice amongst studies using tweets for event detection (de Bruijn et al., 2019; Ostermann & Spinsanti, 2012; Sakaki et al., 2010), the usage of hometowns is more diverse. Especially in the context of event detection, authors are far more concerned with the *about* as it refers to the location of the event and thus might use the hometown to further approximate the *about* rather than differentiating a *from*. Different approaches use the hometown as a contextual clue during the toponym resolution (de Bruijn et al., 2018: 6) or make the assumption that a tweet likely refers to an event within close proximity of the specified hometown (Sakaki et al., 2010: 853). Others choose not to use hometowns at all due to their problematic nature (Ostermann & Spinsanti, 2012: 22). For this thesis, however, it is crucial to obtain the knowledge of a tweet's (broad) spatial origin (i.e., the *from*) as it is the goal to determine the different characteristics of various geographical regions.

The process of geoparsing is divided into two steps, first identifying toponyms within the text, and then attempting to match these toponyms with real-world locations. The following describes how this procedure was carried out in this thesis.

1. ***Toponym recognition***

   As both tweet content and hometowns are pieces of natural language, the NLP library *spaCy* for Python (SpaCy, 2021) is used to analyse them. SpaCy was chosen due to its multi-lingual Named Entity Recognition (NER) model which is not included in other popular NLP libraries such as NLTK, whose tagger performing NER is English-only. SpaCy's NER model uses annotated Wikipedia data from Nothman et al. (2013) to recognise named entities (locations, organisations, etc.) in a piece of natural language and covers all languages included in this thesis. T*oponym recognition* is essentially a part of NER, as locations represent one form of named entities. As such, the NER is used to identify place names, whilst any other tagged objects are disregarded.

2. ***Toponym resolution***

   Any toponym candidates identified in the recognition phase are then checked with the *GeoNames* gazetteer (GeoNames, 2021) to either find the geographic location of the placename or discard it as a non-geographic entity. During this process of *toponym resolution*, *toponym ambiguity* plays a major role and must be considered (Purves et al., 2018: 214). To deal with such uncertainties, an approach akin to the toponym-based algorithm for grouped geoparsing of social media (TAGGS) proposed by de Bruijn et al., (2018: 2) can be considered, which clusters tweets seemingly reporting on the same event and allows for spatial information to be shared between these related tweets instead of geoparsing each item individually as is done traditionally. This grouping approach has been shown to achieve significantly better results when detecting floods worldwide using tweets (de Bruijn et al., 2018).

Geoparsing in this thesis is always carried out on datasets containing tweets from a (full) single day, single language and single event. In a first step, the potential placenames identified in during the recognition process are checked against all country and ADM1-level names in the GeoNames database. Toponym ambiguity increases significantly with a higher spatial resolution as ambiguous toponyms tend to be spatially autocorrelated and found in close proximity of each other (Brunner & Purves, 2008: 26). This, in turn, means that there is far less ambiguity to be expected at these top levels and toponym resolution can be carried out in this rather straightforward manner.

Once all country and ADM1-level toponyms in the dataset are resolved, a similar approach to the TAGGS algorithm is used. Therefore, the same assumption mentioned above is made, grouping all remaining identical toponyms (e.g., “Dresden”) together and later assigning them the same real-world location. During this process, any other toponyms co-occurring with the toponym in the same tweet or hometown (e.g., “Germany” or “Saxony”) are used for context. Out of these co-occurrences, only the previously resolved country and ADM1-level toponyms are selected as they would suggest a containment relationship.

Tab. 8: Location candidates for the toponym “Dresden”, when searching for exact matches or matches in the ‘alternatenames’ column of the GeoNames dataset. Filtered for populated places and regions.

GeoNames ID	Name	Latitude	Longitude	Country	ADM1	Population
3305799	Direktionsbezirk Dresden	51.16667	14.08333	DE	Saxony	1631486
2935020	Kreisfreie Stadt Dresden	51.0833	13.7666	DE	Saxony	554649
6551127	Dresden	51.05	13.75	DE	Saxony	554649
2935022	Dresden	51.05089	13.73832	DE	Saxony	486854
5087183	Town of Hanover	43.71556	-72.1913	US	NH	11401
7171634	Town of Dresden	36.27844	-88.6939	US	TN	3005
4619013	Dresden	36.29145	-88.7081	US	TN	2898
5152291	Dresden	40.12146	-82.0107	US	OH	1706
4962833	Town of Dresden	44.07944	-69.7394	US	ME	1639
7317524	Village of Dresden	40.12145	-82.0113	US	OH	1529

Finally, a list of possible candidates matching the potential toponym (“Dresden”) are retrieved from GeoNames (Tab. 8) and then each assigned a score depending on the number of times their corresponding country and ADM1-region are present in the list of co-occurrences. In the end, the candidate with the highest score is assigned as the real-world location of the group and in the case of a tie, the candidate with the largest population is selected. Therefore, if the co-occurrences are dominated by mentions of “Germany” and “Saxony”, *ID 3305799* will be chosen to represent “Dresden”. If more co-occurrences included “United States”, a result would be chosen depending on the most frequently mentioned state.

Tab. 9: Review of geoparsing precision and recall. A sample of 500 German heatwave tweets was used.

Process	Precision	Recall	F-Score
Toponym recognition (solo)	0.73	0.84	0.78
Toponym resolution (solo)	0.86	0.95	0.90
Complete Geoparsing process	0.86	0.79	0.82

This combined geoparsing process has led to satisfactory results, especially on the lower levels of resolution which are predominantly used for the analysis. Tab. 9 displays the performance analysis of the geoparsing process described above, demonstrating good values for the recognition and better values for the resolution. The overall evaluation shows that many shortcomings of the recognition can be compensated for with good resolution.

3.3.5 Event detection/clustering

The five event types studied in this thesis materialise over different time frames in nature. Heatwaves and wildfires can be considered short-term event types (STETs), as the events themselves only have durations spanning a few days to several weeks (Pezza et al., 2012: 211; Stefanon et al., 2012: 2; Sun et al., 2019: 131). Meanwhile, glacier shrinkage, permafrost melt, and sea level rise are long-term event types (LTETs) taking place over the span of decades and centuries and exhibiting much slower changes (Church & White, 2011: 585; Schuur et al., 2015: 177–178; Zemp et al., 2015: 749–750).

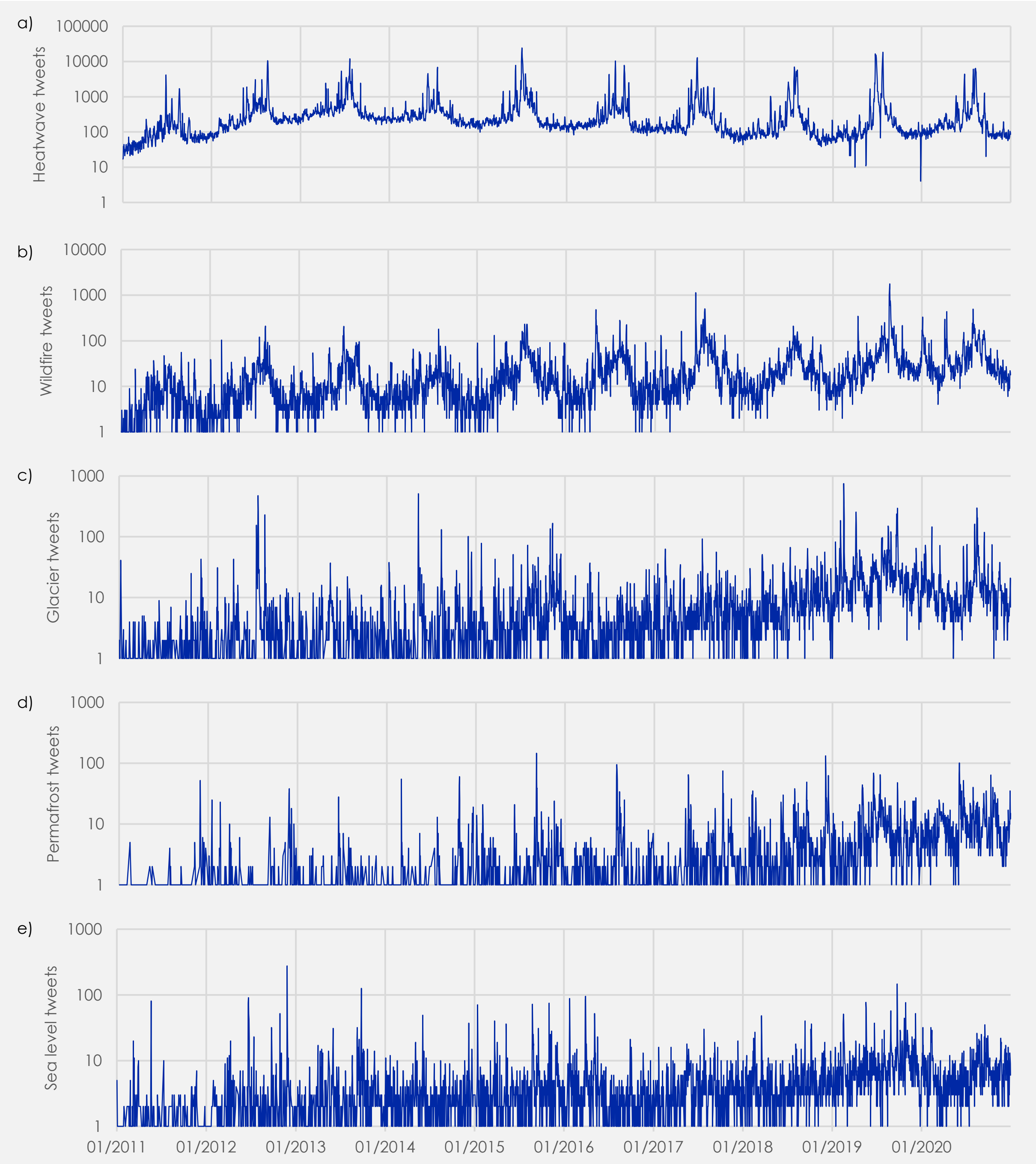

Fig. 8: Tweet counts for French language tweets. Y-scale is logarithmic.

Fig. 8 shows that especially in the case of heatwaves, the short-term events are easily distinguishable from the noise when looking at tweets regarding the issue. Similarly for wildfires, events can be made out simply from inspecting the timeline as they exhibit clear peaks. Interestingly, the long-term event types are not only expressed as random noise but can in some cases display event-like characteristics as well. This is expected to be the result of notable occasions such as publications of media stories and scientific reports, proxy-events such as ice shelf collapses, the viral spreading of videos and images on the internet, political debates and policy decisions, as well as climate change conferences and others (Kirilenko & Stepchenkova, 2014: 176) inciting public discourse on the issue without necessarily demonstrating the immediate occurrence of an extreme event – which is in turn commonly the case with heatwaves and wildfires (e.g., Wang et al., 2016: 529).

As a result, detecting ‘events’ for these long-term climate change effects would mostly yield meta-events – occasions where public debate is sparked by an intermediary source about a process, rather than the direct observation of the process itself. Thus, selecting only tweets from these ‘events’ would bias the dataset towards the most popular secondary reports and disregard less prominent, yet still valid/relevant reports. In contrast, the short-term nature of heatwave and wildfires generally results in a quasi-co-occurrence of real-world events, first-hand reports and meta-discourse in the close period around the peak of the event (e.g., Wang et al., 2016: 529). Thus, delimitating the short-term dataset into event tweets and noise based on a threshold value does not set the border between more and less popular meta-information, but along the real-world extent of the event[8]. The following description will therefore focus on the detection of short-term events by threshold, whilst long-term events are determined based on regular time intervals (annually).

Jumps in a frequency distribution can be determined visually or mathematically, with the former usually requiring considerable experience (Riley, 2008: 154). As apparent from Fig. 8, however, peaks in the heatwave and wildfire tweet distributions strongly deviate from the noise of non-event times, therefore allowing for a threshold based on visual inspection. In practise, events for these two event types are determined for each year separately (to account for changes in user population size), with the threshold being defined as a certain number of standard deviations from the annual mean. The annual data is grouped by country and encompasses all languages. The threshold is set empirically based on inspection of the tweet frequencies and remains the same for each year (in terms of number of standard deviations, actual values differ accordingly). With this approach, the strong spatial and temporal variations present in the dataset do not impact the detection of events beyond the scale of a country and a year.

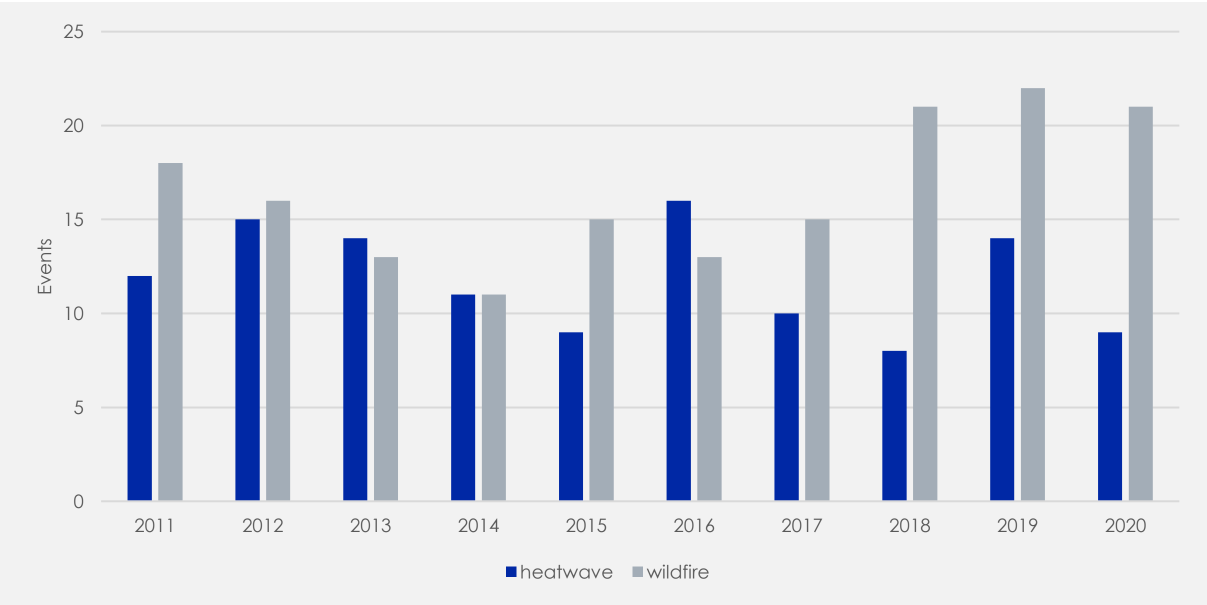

Fig. 9: Total annual event counts for heatwave and wildfire. Counts summarise all regions of the study area.

To provide an example, the French heatwave tweets from Fig. 8 can be divided into annual diagrams with visual indications of various degrees of deviation from the corresponding annual mean.

[8] LTET ‘events’ or peaks are nonetheless useful in some analyses where these meta influences are investigated and will therefore be utilised in these specific instances only.

Subsequently, a single deviation threshold is empirically determined (in this case $+2\sigma$) that can distinguish events from noise across all years. Applied back to the actual daily tweet numbers of each annual distribution, this results in the same relative but different absolute threshold depending on each annual mean. Event detection has the added benefit of raising the slightly poorer classification precision of STETs (section 3.3.2) by discarding tweets outside of major relevant events.

As Fig. 9 shows, the distribution of events is somewhat random over time and does not exhibit any clear trend throughout the study period. The LTETs glacier, permafrost, and sea level all have seven events per year as every country's tweets are separated on an annual basis.

3.4 Thematic Analysis

The remaining chapters of this methods section will now describe how the different analyses were carried out. In the *thematic* (or *content*) *analysis*, the main focus lies on the contents of the tweets. Specifically, the intention is to represent content in units "*as large as is meaningful [and] as small as is feasible*" (Krippendorff, 2018: 106). Tweets lend themselves for such an analysis as they themselves are already highly condensed statements due to the character limit for each message (Kirilenko & Stepchenkova, 2014: 181).

This qualitative approach extracts themes from the conversation surrounding an event and determines concerns and interests Twitter users connect with their occurrence, identifying spatiotemporal patterns derived from these tweet contents – or the lack thereof. Such thematic trends include for example how strongly twitter users link the occurrence of the different events with climate change or how strongly news media is represented compared to scientific works.

A distinction is made between user-generated content and URLs (web links) when conducting a content analysis with tweets (Veltri & Atanasova, 2015: 725). The following sections will first focus on the former, while a later section will be dedicated to URLs specifically.

3.4.1 Themes in event tweets

Various approaches for the analysis of user-generated content were tested before I settled on a keyword-based method which seemed the most appropriate for my data and objective. In the following sections, the separate approaches will be discussed individually – first the unfit topic modelling and then the more suitable relative word frequency approach.

Topic modelling approach

Topic modelling, in its simplest terms, is a field of NLP which represents documents of a corpus as vectors in a high-dimensional space and then interprets the resulting structures in regard to topical relatedness and similarity. It is argued that a corpus of natural language texts can be summarized using a limited number of topics (Rehurek & Sojka, 2010: 46).

One popular topic modelling technique is the Latent Dirichlet Allocation (LDA) conceived by Blei et al. (2003), which is used in a variety of studies attempting to derive topics from Twitter and other social network posts (e.g., Bauer et al., 2012; Resch et al., 2018). LDA is an unsupervised machine learning approach based on the co-occurrences of words within documents as well as word-document counts, extracting topics and their corresponding tokens without the need for prior knoswledge about potential topics in the corpus (Bauer et al., 2012: 349; Rehurek & Sojka, 2010: 48; Resch et al., 2018: 363).

Literature shows that LDA can be successfully combined with spatiotemporal analyses for the investigation of tweets, for example giving insight into the local characteristics of earthquake events (Resch et al., 2018: 374) or highlighting trends in descriptions of local urban environments (Bauer et al., 2012: 356). Adapting the approaches taken by these studies in regards to model inputs and data handling, I used the *Gensim* library (Rehurek & Sojka, 2010) to carry out LDA for my data, treating tweets as documents and choosing varying time periods to delimit the corpus (event-based, year-based, entire dataset).

Tab. 10: Top ten topics of LDA output for approximately 60000 English-language heatwave tweets from 2011-2020 (Model parameters: α = 0.00001, β = 0.1, passes = 10, N topics = 25).

Topic	Tokens (descending relevance)
1	hot, weather, rain, today, day, get, go, cold, sun, check
2	love, hot_weather, year, find, feel, bad, mean, stop, get, train
3	hot_day, today, home, dog, call, office, ice_cream, go, sunny, enough
4	work, temperature, new, stay_cool, remember, tell, today, look_forward, hot_weather, advice
5	climate_change, take, want, long, tonight, walk, start, even, cause, ever
6	summer, sun, thunderstorm, last, country, britain, early, get, extreme_heat_keyword, arrive
7	thank, drink, hope, actually, happy, hot_weather, live, fun, deal, put
8	week, know, help, especially, many, hot_weather, warm, wait, experience, share
9	cool, see, right, place, hot_weather, care, warn, ireland, stand, way
10	london, cope, back, set, record, gon, show, hit, today, friday

However, even with extended testing of different variable combinations, model parameters and data pre-processing methods, the resulting outputs did not amount to any meaningful sets of topics, often either demonstrating large overlaps between topics and/or limited intra-topic coherence between the tokens (see Tab. 10). Thus, using these LDA-generated topics to infer any general trends within the data would have constituted rather subjective guesswork to arrive at any meaningful interpretation. This is a known issue with machine learning topic modelling and there is no academic consensus on appropriate interpretation methods (Resch et al., 2018: 372).

Relative word frequency approach

As the topic modelling approach seemed unfit for my analysis past its use as an initial exploration of the data, I considered a slightly more straight-forward method based on keywords. As such, the term *theme* will be used from here on instead of *topic* to signify the absence of any traditional *topic* modelling algorithms.

Parsons et al. (2015: 1222) suggest a three-tiered bottom-up approach for identifying themes in tweets related to disasters, separating "*Basic Themes*, *Organised Themes* and a *Global Theme*". I adapted this framework for my own analysis without deviating far from the methods detailed by the authors. In their first step, they identify their *global theme* to be "[tweets about] UK Storms and Floods 2012/14", which conceptually matches my data structure of tweets about separate event types in Western Europe during the period 2011-2020. This goes hand in hand with the assumption that a large majority of the tweets remaining at this stage of my analysis belong to this global theme due to the previous classification of tweets during the retrieval process, where irrelevant tweets were discarded.

In order to recognise *basic themes*, Parsons et al. (2015: 1223) annotate a sample of tweets with keywords describing the main message of each tweet and how it was expressed. I decided on a slightly different methodology, which generalised the sampling and annotating of keywords (however retaining the bottom-up approach) and then allowed me to apply the acquired knowledge on the entire dataset. This deviation is mainly possible since I only investigated the information conveyed with each tweet and not the emotional context in which it was expressed. My extraction of basic themes follows a three-step preparation process:

1. For each event type, the word frequencies for all event tweets in each of the four languages (German, English, French, Spanish) are counted and normalised, excluding a small set of stop words, twitter usernames ('*@user*') and hyperlinks. This yields a list of the most common words in each language per event type.
2. To identify the most meaningful themes for the analysis, I am however looking for the most important keywords relating to the event rather than just the most common ones. To find such

relevant keywords, word frequency lists for general tweets were taken from the *WorldLex* datasets collected by Gimenes & New (2015), which allow for a comparison of keywords between my datasets and general language[9], thus resulting in a relative frequency value describing this deviation from the norm for each word. Ranking the list by this figure seemed to yield good results upon visual inspection, bringing keywords from the original retrieval query and other terms closely related to the top. This would build a solid base for sourcing basic theme keywords.

3. On top of the relative frequencies, a Chi-Square test was performed to check the statistical significance of each deviation. The results suggest that a large majority of terms with a high relative frequency (large deviation) also exhibit a significant deviation ($p < 0.05$) with a gradual decrease in significance the lower the relative frequency drops. Thus, the top few hundred keywords were hardly affected by a lack of significance and only a handful of terms was excluded in this step.

To finally identify basic themes, the top 300 keywords ranked by relative importance for each language-event type combination were annotated following a set of rules. During the first iteration, the terms were categorised in more detail, whereas a second iteration sought to combine these fine-grained basic themes into more generalised *organised themes* in a similar approach to Parsons et al. (2015: 1223), who also achieved this final step by grouping the basic themes corresponding to a set of rules into *organised themes*. A similar approach was taken by Vieweg et al. (2010: 1082), who started by annotating tweets and then iteratively refined the annotations into categories, finally applying the annotation scheme to all tweets in the dataset. Furthermore, Moernaut et al. (2020: 6–7) also used a comparable method, dividing their dataset of climate-change-related tweets into a 'frame set' consisting of frames and subframes analogous to the organised and basic themes discussed above.

In my approach, each tweet in the datasets could be assigned scores for every organised theme based on the number of theme-related keyword mentions occurring in the tweet. This compromise still allows for full coverage of all tweets, as the manual annotation of millions of tweets would not have been possible. Organised themes stretching across all of the event types and of special interest for the analysis ('climate change', 'risk/danger', 'news', 'science') were further enhanced by a set of default keywords including the most basic terms connected to said themes (e.g., 'scientist' and 'research' for the science theme), which should result in more comparable values when cross-examining the same themes for different event types. Scores are given in a simple count (mentions/tweet) and there is no limit on the number of different themes a single tweet can include. Having assigned each tweet its themes, the associated spatiotemporal information can then be used to investigate developments of the distribution of themes across space and time for the different events. For example, Parsons et al., 2015 (1224–1225) discuss how patterns change during the lifecycle of an event (Fig. 10).

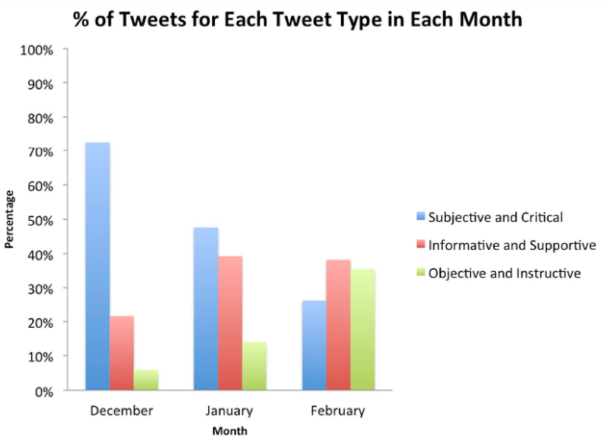

Fig. 10: Distribution of organised themes in disaster tweets by month (Parsons et al., 2015: 1225).

3.4.2 Special tweet content

Hashtags are still part of the UGC using the distinction made in the chapter's introduction above, however do they hold a special status within this content. Twitter users utilise hashtags to situate their tweet within a topic or event, thus automatically self-categorising or self-annotating their contribution (Williams et al., 2015: 127). This includes, for example, hashtags representing the impactful events mentioned above. Looking at climate change-related hashtags, it is also possible to make educated

[9] WorldLex includes a dataset with word frequencies for general language derived from tweets.

guesses about users' stances on the topic: Williams et al. (2015: 129) found that the hashtags #climate and #climatechange are dominated by activists, while #climaterealists is largely used by sceptics. Both #globalwarming and #agw contained a mixed population with both activists and sceptics contributing to the discussion. Similar to themes, the temporal distribution of hashtags can illustrate the rise and fall of important topics and events in the climate change discourse on Twitter (Kirilenko & Stepchenkova, 2014: 176). Hashtags are thus used in this thesis to help understand and further contextualise and augment findings made using the theme approach described above.

Web links or *URLs* (and the destination web pages) contained in tweets are only selected, but not created by the users themselves. They are argued to be an important point of analysis regarding climate change content on Twitter as they represent a major form of information sharing, with users taking on the role of curators rather than writers (Veltri & Atanasova, 2015: 722, 733). When discussing climate change, Twitter users often refer to external sources in their tweets using URLs, most commonly sharing legacy news articles, news aggregator content, blog posts and a variety of popular as well as political science pages. Websites frequently used by a variety of users can be described as authoritative information sources regarding the tweets in the dataset (Kirilenko & Stepchenkova, 2014: 177, 180; Veltri & Atanasova, 2015: 733).

As a result, web links influence my thematic analysis twofold: For one, the 50 most popular general web pages (e.g., www.theguardian.com) for each country across all event types are annotated based on categories found in Kirilenko and Stepchenkova (2014: 177–179) to allow for an evaluation of the importance of different source types in various contexts. Second, frequently shared links pointing to specific articles and similar are used to derive important individual sources that shape the discourse at a certain point in time.

3.4.3 Conversation characteristics

Twitter consists not only of original tweets, but also enables users to share the statements and opinions of others by retweeting as well as engaging in conversations with each other using public replies and since 2015 also quoted retweets, which allow users to comment on a post they are retweeting. Conversation on Twitter regarding climate change can be quite segregated between activists and sceptics, with especially follower and retweet networks demonstrating patterns of strong homophily, meaning few connections between the two groups exist. Looking at mention networks (users addressing each other directly in their tweets), however, a stronger heterophily can be observed, thus suggesting that this is the medium chosen by members of different groups to interact with each other (Fig. 11) (Williams et al., 2015: 130). Across their dataset of tweets, Williams et al. (2015: 131–132) found that 28% of users in mention networks are members of an 'open forum' (as opposed to an 'echo chamber'), compared to only 2% and 3% for follower and retweet networks, respectively. Therefore, if one intends to investigate a two-sided debate related to climate change on Twitter, mention tweets are the primary source.

Based on these findings about conversations on Twitter, three different metrics for conversation characteristics are investigated in this thesis:

1. ***Conversational Tweets***

   As Williams et al.'s (2015: 131–132) findings introduced above establish reply tweets as the medium of choice for debate between members of different perspectives on Twitter, the percentage of reply as well as quote[10] tweets in a corpus serves as an indicator for the level of conversation and discussion occurring. This should give an insight into which themes elicit the most debate.

[10] Not included in Williams et al. (2015) due to their launch after the completion of the study. Their similar functionality to reply tweets enabling commenting on another user's statements is, however, expected to result in comparable outcomes.

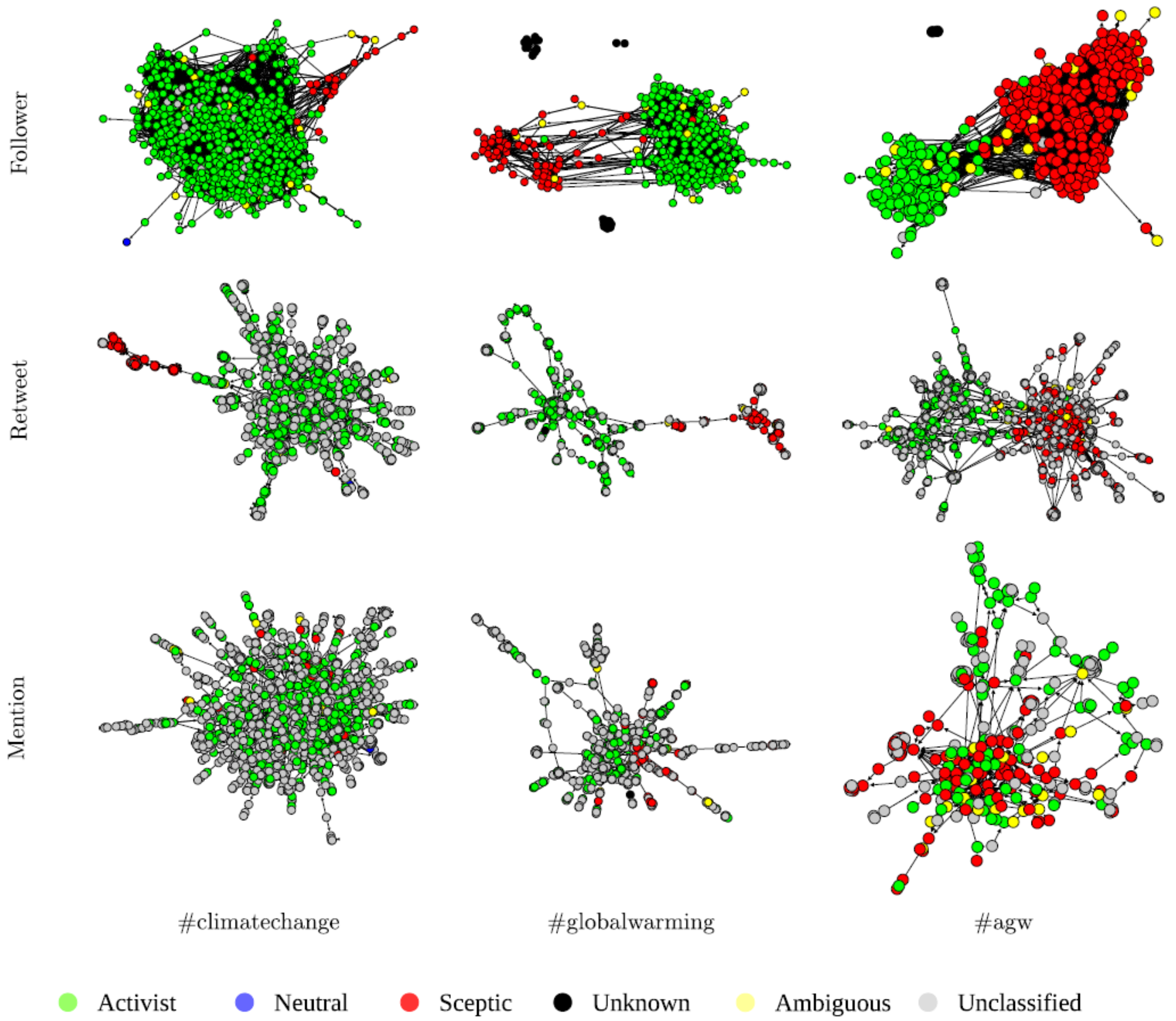

Fig. 11: Attitude towards climate change in different types of networks (nodes=users, edges=interactions) on Twitter for tweets using selected hashtags (Williams et al., 2015: 130).

2. *Controversial Tweets*

It can be approximated how controversial a tweet is by calculating the ratio of likes and retweets (generally indicating agreement) versus replies, which can indicate disagreement when the latter outweighs the former (Minot et al., 2021: 2). Using this method, I can gain a rough understanding on the extent to which the various discourses elicit controversy.

3. *Tweet sharing*

Users can assign importance to a tweet by retweeting it and therefore spreading it to their audience. Moernaut et al. (2020: 9–11) investigated this type of sharing behaviour for a sample of climate change-related heatwave tweets to ascertain retweeting patterns for different perspectives on the topic. In a similar fashion, I compare how frequently tweets belonging to different themes are shared (i.e., retweeted) to quantify how important other members of the public perceive contributions to the various sub-groups of tweets in my corpus.

3.5 Spatial Analysis

The spatial component is derived from toponyms mentioned both in the tweets and the hometowns specified by the corresponding users to assess where the tweet is from *and* about. To determine spatial variability, an analysis of this geocoded data makes statements about how the reactions to events are distributed across space between and within countries, which can additionally be augmented by temporal information. Further measures such as the degree of transnationality (Reber, 2020: 3, 7) of tweets can give additional insights into the networks of climate change reactions. Finally, an investigation into the influences of relevant environmental processes require the collection of various supplementary datasets.

3.5.1 Spatial distribution

To determine the spatial distribution of tweets, each tweet must first be assigned a single location representing it. As the user hometown and tweet text often generate multiple resolved toponyms as a result of the geoparsing described in section 3.3.4 (e.g., a city and a country in the hometown field), decisions must be made systematically to define the single most likely user location from the selection of toponyms associated with a tweet. The following set of rules was used to carry out this process:

1. If there are user hometown toponyms, tweet text toponyms are discarded.
2. From these hometown toponyms, it is likely that they are formatted in increasing granularity (e.g., ‘Zurich, Switzerland’), therefore the first hometown toponym (e.g., ‘Zurich’) is chosen.
3. If step 1 is not the case, the first toponym from the tweet’s text is chosen. Formatting in the text is much less predictable, however would any formatting likely follow the pattern described in step 2.

Users very commonly specify a hometown and the approximation using the tweet’s content is used much more rarely. It serves, though, to include the group of users not willing to specify a location in the analysis as well. With the locations determined, the point distributions are then aggregated to subnational polygons, where the point distribution is normalised with the corresponding user population size.

3.5.2 Spatiotemporal patterns

The goal of the spatiotemporal analysis is to determine how the distribution of tweet activity changes within the study area over time, therefore requiring a measure defining these fluctuations for each event type. The chosen approach ranks each subnational area based on its tweet frequency (relative to the other areas) on an annual basis. As Fig. 12 shows can this rank fluctuate different amounts over the study period as the tweet frequency changes from year to year and some regions show more interest than others.

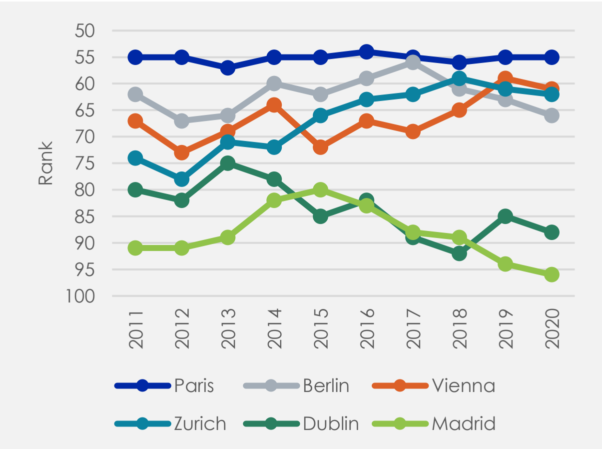

Fig. 12: Selected regions’ annual ranks for glacier tweet frequencies.

The standard deviation of each region’s (RSD) ten ranks (2011-2020) is then used as an indicator to identify the degree of temporal variability in a region’s tweet activity. Taking the mean of all regional standard deviations (MRSD) in the study area gives an indicator on the event type-scale as it summarises all areas.

To examine whether regions with higher tweet activity experience the same spatiotemporal variation as less active regions, a comparison between this overall MRSD and the MRSD of the top 20 most active regions is made. This relationship is further investigated by calculating the correlation (Pearson’s r) between each regions overall tweet activity and RSD, which gives insight into whether fluctuations in tweet activity are related to ranking (i.e., overall tweet activity).

3.5.3 Transnationality

Climate policies are implemented on local, national, and global scales, with the national arena retaining the most importance as it controls both general domestic policies and international engagement in climate action. As a result, an integrated public climate change discourse on the national level looking both inward on local efforts as well as outwards on transnational aspects is of utmost importance (Reber, 2020: 1). The concept of transnationality can be used to investigate patterns of domestic and international discourses by drawing connections between users' assumed home and places they mention in their tweets (Krackhardt & Stern, 1988: 127; Reber, 2020: 3, 7). This degree of transnationality can be measured by the *E-I index*, which determines the ratio of domestic links d to transnational links t in a network (Krackhardt & Stern, 1988: 127):

$$EI_i = \frac{t_i - d_i}{t_i + d_i}$$

Using Twitter data, such referential links can be extracted by investigating the relation between the specified hometown of a user and toponyms mentioned in the tweet. This will allow for the calculation of a measure determining how transnational people in different places perceive and discuss climate change, allowing for inferences about the importance of domestic and international aspects of each subject.

3.5.4 Environmental context

To get a better understanding of how patterns in the public's reception are formed, the influence of environmental variables is investigated. The goal is to determine how people react to fluctuations of these variables on the local, national, and global scale to identify the granularity of events that elicit the strongest reactions.

Tab. 11: Collection and processing of event-related supplementary data.

Event type	Dataset	Processing	Data reference
Heatwave	Daily temperature measurements at stations across Europe	Averaging of values across subnational and national scales	Klein Tank et al. (2002)
	Global daily temperature anomaly in a 1°x1° grid	Averaging of values across subnational, national, and global scales	Berkely Earth (2021)
Wildfire	All active wildfire areas detected by the MODIS sensor from 2011-2018	Aggregating of burnt areas to local, national and global sums for each day in the dataset	Artés Vivancos et al. (2019); Artés Vivancos & San-Miguel-Ayanz (2018)
Glacier shrinkage	Annual glacier mass balance measurements	Aggregating of values to subdivisions of study area as well as calculation of European and global changes	WGMS (2021a)
Permafrost melt	Annual permafrost extent for the Northern Hemisphere	Annual delta to previous year, on the global scale	Obu et al. (2021)
Sea level rise	Satellite altimetry data monitoring global and regional changes in sea level	Aggregating values to monthly and annual intervals	NOAA Laboratory for Satellite Altimetry (2020)

For each of the event types, additional data was gathered to put the tweets into context with observations of the corresponding real-world environmental variables. Tab. 11 summarises the utilised data and describes the processing done to adjust the data to fit into the spatial and temporal frameworks used in the analyses of tweets in this thesis. In general, values were aggregated over space to match subnational and national divisions of the study area and a global value was calculated to compare the dependence on environmental variables inside and outside the study area. On the temporal axis, time intervals of data points largely depend on the event type, with the short-term nature of heatwave and wildfire events favouring daily data, whereas the long-term and often season-dependent event types glacier, permafrost and sea level are more suited to annual data points. These intervals further match the scale of individual events as defined by the event detection in chapter 3.3.5.

With the data above processed, correlations (Pearson's r) were calculated matching each region's tweet activity with the environmental values at the different scales. This allows for the comparison between reactions to events at the different granularities. With these values attached to each day (and therefore each tweet) in the case of STETs, it is also possible to identify national thresholds at which users start to react to events. For example can temperature thresholds for different definitions of heatwaves across the study area be empirically assessed this way.

3.6 Temporal Analysis

For the temporal analysis, only a limited amount of data processing was necessary as large parts of the analysis are characterised by a more straightforward interpretation of rather basic data. The following section will thus give a very brief overview of the main algorithms used to generate this data. The analysis of periodicity will not be discussed, as it simply aggregates tweet timestamps to various measures of time.

3.6.1 User population over time

To determine the development of user populations over time and observe changes in discourse participants, the composition of users can be analysed in different time intervals using the Sørensen similarity (Williams et al., 2015: 128). As this thesis is based on an event-based system, user populations are compared between events and on a per-country basis.

The Sørensen similarity index (Sørensen, 1957) simply measures the overlap between two populations and normalises the value with the size of the same populations:

$$S = \frac{2\ (pop_1\ \cap\ pop_2)}{pop_1\ \cup\ pop_2}$$

This returns a value between 0 – denoting no overlap – and 1 for identical populations. In the case of Twitter user populations, it is further interesting to determine this overlap for the most prolific users to investigate if the voices dominating the discourse change at the same rate as users at the periphery of the discussion. To achieve this, the similarity is not only calculated for the entire userbase of each event, but also for the 10 and 100 users who published the largest number of tweets per event (Williams et al., 2015: 128).

3.6.2 Event timelines: finding causes for events

Temporary peaks in climate-change-related tweets tend to stem from either local weather or major news events (Kirilenko & Stepchenkova, 2014: 176). To gain further insight into the causes of peaks in tweet activity surrounding event types that are less directly related to local weather (glaciers, permafrost, and sea level), the tweet frequency distribution over the study period is analysed for each country and event type. The most notable peaks for each country are summarised in a timeline depicting the most important occurrences of the study period causing increased tweet activity.

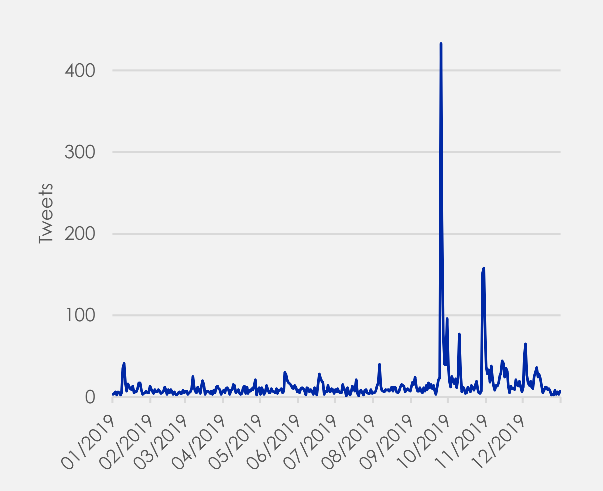

Fig. 13: Tweet frequency distribution for sea level tweets made by Spanish users in 2019.

To find notable increases in tweet activity, the largest single-day peak for each country-event type combination is investigated on an annual basis (e.g., the late September peak in Fig. 13 for Spanish sea level tweets). The peak is included in the timeline if it (1) is larger than ten tweets in size, (2) represents a considerable anomaly in tweet activity for the corresponding annual distribution, and (3) consists of less than half obviously corporate tweets (e.g., tweets by newspapers, organisations, etc.).

To help better contextualise the importance of the peak – and thus the impact of the event causing the peak – the deviation from the annual mean of daily tweets for the observed distribution (e.g., Fig. 13) is used. Specifically, the anomaly represented by the peak is measured in the number of standard deviations from this annual mean. The clearly discernible peak during late September in the example of Fig. 13 exhibits a high anomaly of +14.5 standard deviations above the annual mean.

A peak with a high anomaly therefore signifies the occurrence of an event that largely overshadowed any other events in terms of public interest during the year, whereas a lower anomaly either indicates a generally lower interest in the event (peak is closer to daily baseline/noise) or the co-existence of multiple high-interest peaks during the year (raising the annual mean and in turn decreasing the number of deviations). Testing has shown that even with three or four high-profile peaks in the same year, high anomalies can still be reached. The former is therefore generally a more likely interpretation.

3.6.3 Impact of key events

While the analysis above attempts to find occurrences in the real world by investigating tweet frequencies, the inverse analysis can be carried out with a set of real-world occurrences as a starting point. The goal is to determine whether certain events such as the occurrence of United Nations Climate Change Conferences (COPs), the publishing of IPCC reports, global climate strikes, and the COVID-19 pandemic have the same influence on the tweeting behaviour in this thesis as has been observed in previous research on climate change tweets (Kirilenko & Stepchenkova, 2014: 176).

For COPs, IPCC reports, and strikes, the quantitative analysis used to determine the level of public reaction during these occurrences is largely identical to chapter 3.6.2, with the two differences being that the timespan of the distribution is limited to the month before and after the event and that the anomalies are not calculated for the highest peak of the distribution, but of course for the day(s) of the event.

To identify the impact of COVID-19, the monthly tweeting activity is compared between 2019 and 2020 for each event type. As the Twitter userbase grew considerably between the two years (Statista, 2021), months in 2020 exhibiting very similar or even lower tweet numbers compared to the corresponding months in 2019 could potentially be impacted by the pandemic and users prioritising COVID-19 over climate change. This comparison is further augmented by a complete overview of tweet frequencies throughout the study period, which is calculated as an annual tweets per user metric.

4. Results

4.1 Thematic dimension

I will begin the presentation of the results with the thematic dimension before moving on to the spatial and temporal counterparts. During these descriptions, I will use the term ‘heatwave tweets’ – and analogously all other event types – for simplicity’s sake to refer to the tweets published during any detected event. As introduced in section 3.3.5, the two STET tweet corpora therefore only include tweets immediately surrounding wildfire and heatwave events, while the three LTET corpora include any relevant tweets published during the entire study period.

The thematic analysis of the tweets’ contents generated a wide array of results regarding themes and conversation characteristics. The first part of this chapter will analyse a selection of these themes with a special interest in climate change and the balance between news media and science-related content. A final section quantifies how strongly various aspects of the different discourses elicit conversation and controversy between users.

4.1.1 Climate change

This thesis investigates climate change effects, but the question remains to which degree climate change itself is actually discussed. The importance of climate change differs mainly between short-term and long-term events (Fig. 14). While roughly a fifth to a third of glacier, permafrost and sea level-related tweets contain a climate change keyword, less than 5% of wildfire and heatwave-related tweets contain the theme for a large part of the study period. The exception to this rule are wildfire tweets in the time span between 2018 and 2020, where the keyword frequency rapidly increased, reaching around 12% in the final year.

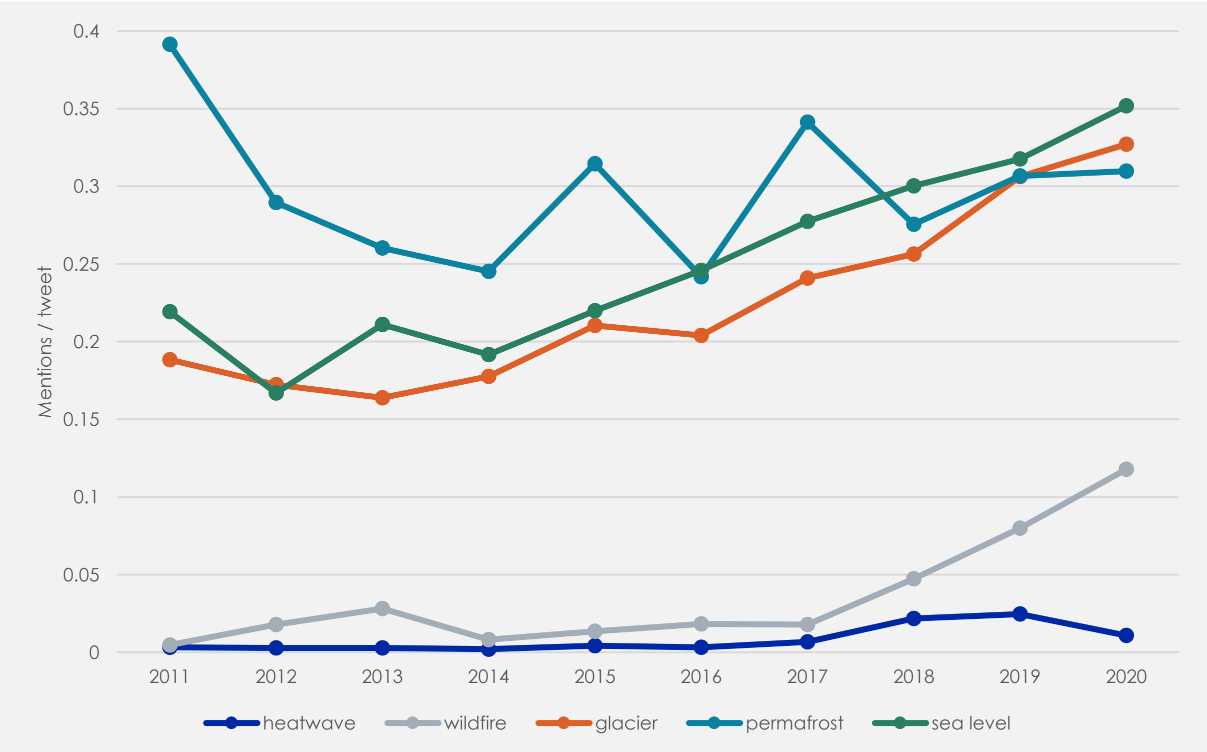

Fig. 14: Importance of climate change in event tweets.

Looking at other trends, the importance of climate change in sea level and glacier tweets was steadily increasing throughout the latter two thirds of the study period, both climbing from below 20% to above

30%. Their rather similar values and trends suggests that the events are often mentioned in combination with each other – at least when discussing climate change – which is likely due to the correlation between glacier melt and sea level rise. In fact, this connection can be confirmed when looking into the word frequency distributions for the tweets of each event type (more specifically the relative frequency, as introduced in chapter 3.4.1): For sea level tweets, references to glaciers and their melting can be found within the 50 most important keywords of every language. Similarly, direct mentions of sea level reside in the top 25 keywords for glacier tweets. This causal relationship will reveal itself in the results on multiple further occasions.

It is interesting to note the overall high percentage of glacier tweets containing a climate change keyword (33% in 2020) as imagery of melting glaciers and ice caps has been suggested to be one of the most fruitful methods of convincing people in the United States that climate change exists (Borick & Rabe, 2010: 785–786). The high number of glacier tweets mentioning climate change suggests that the phenomenon is also a frequently used for the same argument in Western Europe. In fact, of those of glacier tweets mentioning climate change, a further 18.7% also mention risk/danger, indicating that glaciers are not only used to reinforce the climate change argument, but also to do so with a tone of urgency to highlight the severity of potential outcomes.

Tab. 12: Overlap between tweets mentioning climate change and tweets mentioning risk/danger.

Event	Tweets	Climate change mentions	Thereof danger mentions	Overlap [%]	Overlap as % of all tweets
Heatwave	786827	8370	2458	29.4	0.31
Wildfire	281434	14828	5145	34.7	1.83
Glacier	255389	67036	12533	18.7	4.91
Permafrost	38992	12066	3573	29.6	9.16
Sea level	116797	33605	8617	25.6	7.38

This high overlap between climate change and risk/danger mentions can be universally observed across all event types with values ranging between 18.7 and 34.7% (Tab. 12), indicating that there certainly is considerable concern about the dangers of climate change when users discuss events related to global warming. It is worth keeping in mind, however, that the original proportion mentioning climate change in the first place differs strongly, especially between STETs and LTETs. Therefore, the actual percentage of all users concerned about the dangers of climate change is an order of magnitude lower in heatwave tweets compared to wildfire tweets. LTET tweets display an even more significant overall overlap between the two themes. Judging by these values, the risks associated with permafrost melt due to climate change are especially concerning to users and contribute majorly to the permafrost discourse.

Returning to Fig. 14, the importance of climate change for permafrost is highly dependent on whether it is a central point in the news stories and scientific articles causing users to tweet about the issue every year (see section 4.3.3), thus resulting in a curve without a clear trend that rises and falls each year accordingly.

Although still displaying a low percentage compared to other event types, heatwave tweets saw a significant increase in climate change mentions during the latter half of the study period with a peak during the record-setting heat year in 2019 (Xu, Wang, et al., 2020: 3) before decreasing again the following, less hot year. This suggests that users do associate heatwaves more with climate change the more extreme the events are, however only in more recent years as previous hot summers in the early years of the study period do not exhibit any significant peaks. While the summers in the latter half of the study period have indeed been hotter than the former half, the differences are not separated by the same factor observable in Fig. 15 (Xu, Wang, et al., 2020: 3). This would suggest a certain threshold being crossed between 2017 and 2018 where users suddenly get considerably more concerned about climate change related to temperature anomalies.

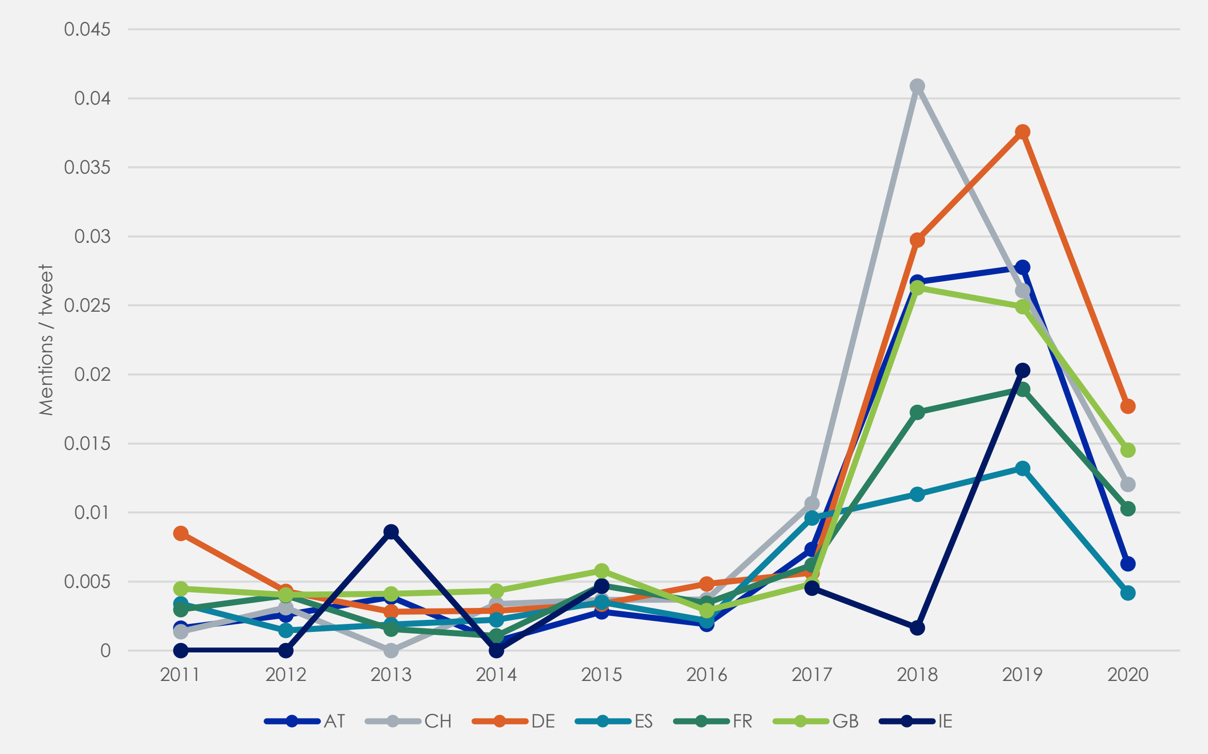

Fig. 15: Importance of climate change per country in heatwave tweets.

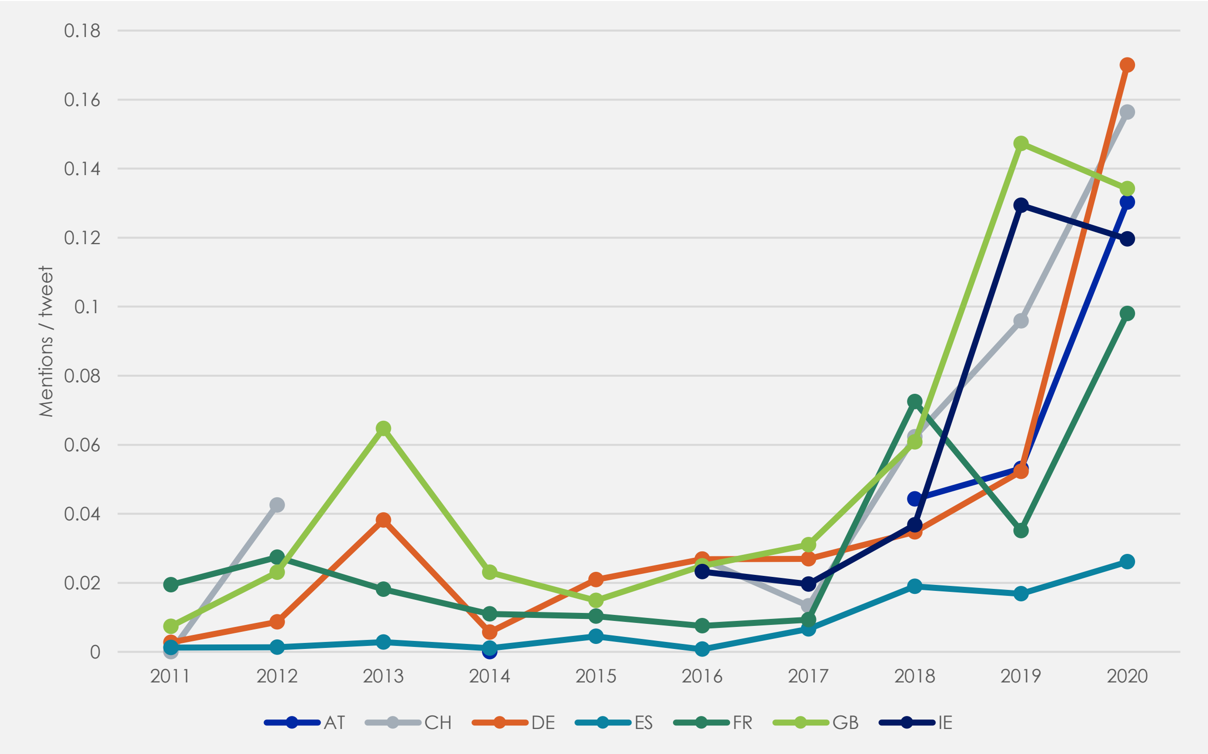

Fig. 16: Importance of climate change per country in wildfire tweets.

Fig. 15 highlights this massive increase of climate change mentions in heatwave tweets between 2017 and 2019 compared to earlier years. During this time period, the usage of climate change-related keywords multiplied by factors ranging from 2 all the way up to 8 with Switzerland jumping from 0.5% in 2016 to 4% in 2018.

A somewhat similar pattern to heatwave tweets can be observed for wildfires (Fig. 16), with the peak period shifted by one year to 2018-2020 and no decrease visible for most countries. The range in peaks is more spread apart between the countries, with Spain and France remaining lower at close to 3% and 8%, respectively and all other countries falling between 13% and 16%. This is, however, still a strong growth for all regions and it can be concluded that climate change has been acknowledged a lot more regularly in the Western European wildfire discourse on Twitter in recent years. The recent increase in occurrence and intensity of many large-scale wildfires such as those in Australia (2019 and 2020), the Amazon rainforest (2019 and 2020) and western North America (2017, 2018 and 2020) (Xu, Yu, et al., 2020: 2173) suggests that Twitter users' concern about climate change in connection with wildfires is dependent on their extent and magnitude. This also applies to the smaller peak observed in Great Britain and Germany during 2013.A later section (4.2.4) will further explore how users react differently to local, national and global events.

For the LTETs glacier, permafrost and sea level, the country-by-country comparison is much more erratic and inconsistent, with countries generally following the overall trends discussed above but showing high variance from year to year, once again suggesting the individual importance of related publications during the year in each country. This literary influence is further supported by the fact that similar trends between countries are usually tied by language in the few cases where they are observable. For LTETs, a deeper investigation into the actual contents of the tweets each year should therefore give more insight into the discourse than this generalised theme overview (see section 4.3.3).

4.1.2 News

News media are argued to have a strong influence on the general public regarding their opinion on and concern about climate change (Ruiz et al., 2020: 114; Segev & Hills, 2014: 67). Thus, the inclusion of this key driver of climate change perception is crucial in any investigation into the climate change discourse online. Investigating retweeting behaviour for climate change-related tweets, Moernaut et al. (2020: 10) found that authoritative sources including legacy media, weather(wo)men and scientists earn a higher number of retweets compared to tweets composed by 'regular', non-authoritative users – highlighting the important role these actors play in the opinion-making of the online sphere.

However, defining the importance of news in tweets solely on keyword mentions (Fig. 17) leads to a systematic underestimation. In the climate change discourse on Twitter, users frequently adopt web links when sharing information and a majority of these external sources are news articles regarding the topic – both in my data and in literature (Kirilenko & Stepchenkova, 2014: 177; Veltri & Atanasova, 2015: 733). Such posts sharing news articles, however, do not necessarily include any specific news keywords and often display a similar pattern of an article headline followed by the corresponding web link:

> *Arctic sea ice is melting at its fastest pace in almost 40 years [link to article]*
>
> *European heat wave gives Germany record temperature: BERLIN - Europe's heat wave has pushed… [link to article]*
>
> *Wildfires blaze across parts of Britain after hottest April on record [link to article]*

Therefore, examining users' linking behaviour can give a more informed insight into the importance of news. Exploring web links in climate change tweets, Kirilenko & Stepchenkova (2014: 177–179) categorised web pages into *traditional news, news aggregators, science, blogs, social networks, media sharing sites* and *advertisements*, focussing on the first four for their research. Within their selection, the authors found traditional news to be most frequently shared and *The Guardian* being the most authoritative source overall. They further identified a correlation between the frequency of climate change-related tweets and the publishing of climate change-related articles in major newspapers (Pearson's r = 0.62 with a one-week lag). Similarly, I can look at the correlation between URL categories and tweet frequency to discover which resources are most popular amongst users when contributing to any given discourse (Tab. 13).

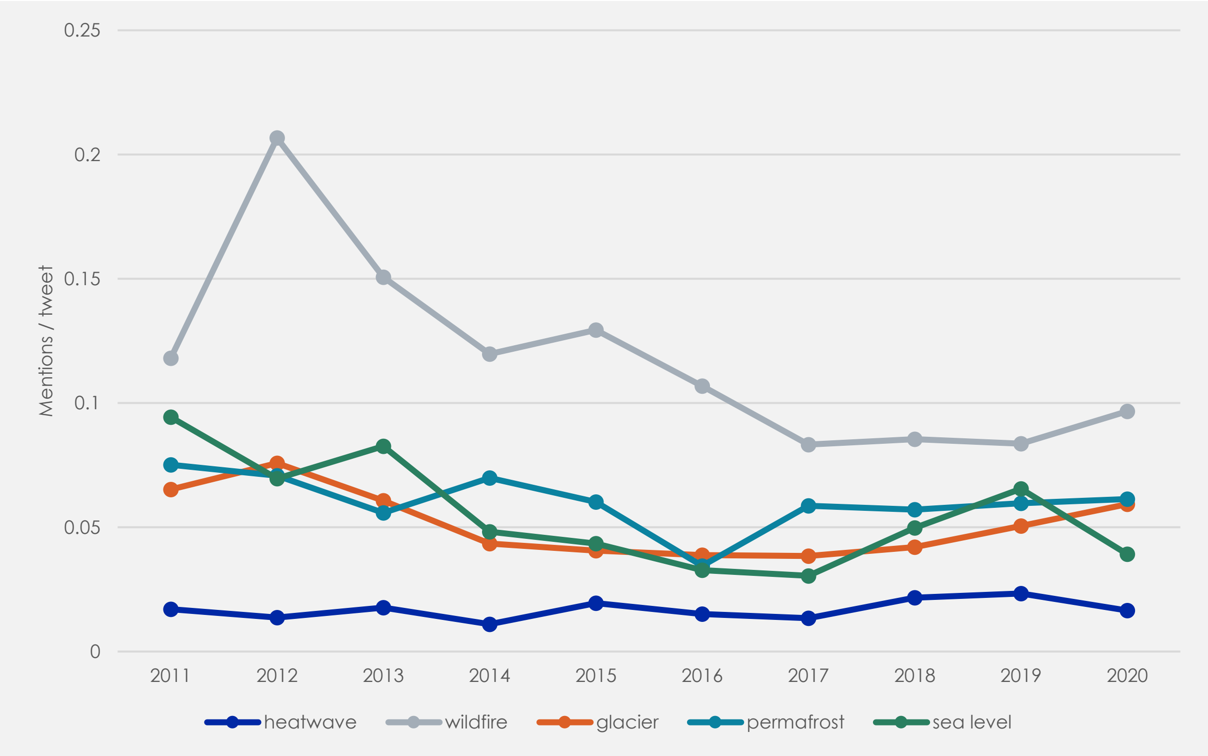

Fig. 17: Importance of news in event tweets.

When looking at shares held by URL categories during an event, there seems to be a symmetry observable across the different event types. The size of an individual event (measured by number of tweets) somewhat correlates with news and social media, whereby the proportion of news links increases, and the proportion of social media links decreases with increased event size. This seems logical, as one would expect a bigger event discussed amongst a larger number of people to be more 'newsworthy', with the newsworthiness being reflected back into the tweet corpus. The fact that these correlations are stronger for the long-term events implies that the number of users tweeting about glaciers, permafrost and sea levels are more strongly tied to increased sharing of news resources. This is in accordance with the expectation that the publishing of relevant news stories for these event types has a considerable impact on the extent of their discussion on Twitter.

Tab. 13: Correlation between number of relevant tweets and share of URL category (relative to other URL categories) during an event.

Event Type	News	Science	Social Media
Heatwave	0.26 (p = 0.00404)	0.45 (p = 0.00000)	-0.19 (p = 0.04393)
Wildfire	0.15 (p = 0.05069)	-0.05 (p = 0.48696)	-0.13 (p = 0.09740)
Glacier	0.49 (p = 0.00002)	0.0 (p = 0.97343)	-0.49 (p = 0.00002)
Permafrost	0.32 (p = 0.00678)	0.15 (p = 0.21797)	-0.39 (p = 0.00094)
Sea level	0.34 (p = 0.00369)	0.18 (p = 0.12930)	-0.41 (p = 0.00041)

The correlations regarding the share of *science* URLs cannot be interpreted due to a lack of significance except in the case of heatwave tweets. Due to the extremely low number of science links in the heatwave corpus (0.08% of tweets), it is expected that they tend to only appear during some of the bigger events, thus leading to a moderate correlation. Overall, URLs linking to explicitly scientific sources are rare compared to news and social networks in all event types (Fig. 18) as a majority of scientific content is shared through news media. The following section will present some more meaningful metrics for defining the importance of science in the various tweet corpora.

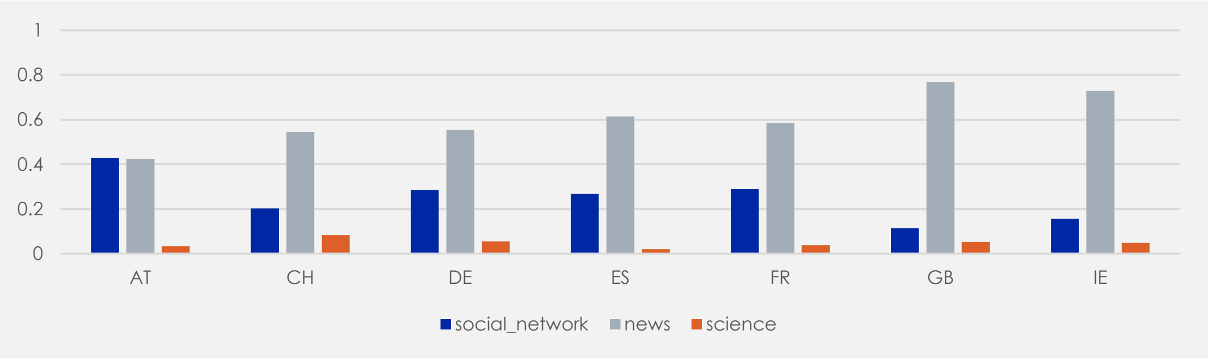

Fig. 18: Shares of URL types per country across all event types.

4.1.3 Science

Scientific content on Twitter can be identified in a variety of ways: Tweets directly mentioning, linking or containing information related to a scientific topic, tweets containing a science-related hashtag, or tweets published by scientists or academic institutions (Weller et al., 2011: 3). This offers a variety of ways to investigate scientific content in event tweets. Looking at a corpus of general climate change tweets, Veltri & Atanasova (2017: 733) found a that science and news are closely tied together in the information sharing networks of the general public, with 74% of web links regarding climate change leading to news articles. Due to the scientific nature of climate change, one might expect that the theme importance of science would follow a similar trajectory.

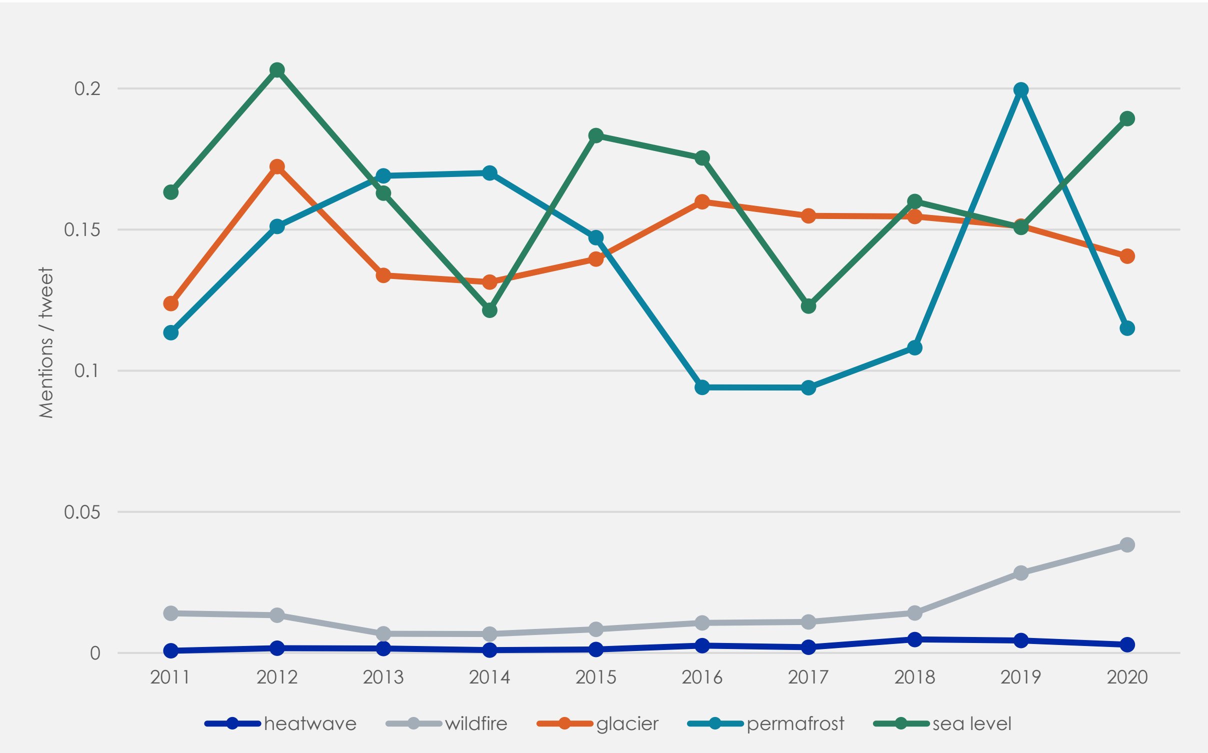

Fig. 19: Importance of science keywords in event tweets.

As Fig. 19 shows (compared to the climate change theme in Fig. 14 above), this assumption does hold true to a certain extent: Overall, the separation between LTETs and STETs is comparable, with the science theme being mentioned roughly half as often as the climate change theme. Most similar are wildfire and heatwave, which once more exhibit their characteristic increase in mentions towards the end of the study period. For the STETs, there seems to be no significant overlaps in individual peaks and the consistent increase of the climate change theme is not represented in the science theme. The closeness of the three event types is, however, rather similar. On the surface level, these observations suggest that the importance of science in the discourse surrounding STETs is somewhat dependent on the importance of climate change in any given year, whereas this dependency is not given for LTETs, where mentions of science do not necessarily seem to coincide with the importance of climate change. This investigation, however, does not take into account that tweets surrounding these more scientific event types might not explicitly mention a science keyword in the tweet as a scientific angle might already be implied by the context of the topic. Subsequently, we can once again look at the web links used in tweets in an attempt to compensate for this deficiency.

Interestingly, the distributions of science links (Appendix, Fig. 52) broadly follow the general trends observed with the keyword-based method in Fig. 19. Although science URLs are very sparse in heatwave and wildfire tweets, the same increase around 2017/18 can be observed. For the LTETs, there is a similar lack in continuous development throughout the study period, with potentially a slight upwards trend overall. These observations seem to back up the previously made conclusion that science has a solid level of importance in LTETs which varies from year to year but seems rather consistent overall.

Tab. 14: Mentions of science links and climate change keywords in tweets. Overlap percentages first denote the overlap's share of the former and secondly the latter.

Event type	Total tweets	Science links	Climate change mentions	Overlap
Heatwaves	786827	620 (0.08%)	8370 (1.06%)	56 (9% / 0.67%)
Wildfires	281434	664 (0.23%)	14828 (5.27%)	93 (14% / 0.62%)
Glaciers	255389	4879 (1.91%)	67036 (26.25%)	1466 (30% / 2.18%)
Permafrost	38992	1162 (2.98%)	12066 (30.95%)	486 (42% / 4.03%)
Sea level	116797	2638 (2.26%)	33605 (28.77%)	597 (23% / 1.77%)

Although some overlaps have been observed, a question to be answered is how much of this scientific content relates to climate change, which is explored in Tab. 14. The low number of science links for heatwaves and wildfires puts into question whether these statistics can be used to make any significant statements about STETs. For LTETs, however, plenty of data is available and permafrost tweets seem to display the strongest bond between science links and climate change mentions with 42% ahead of glacier (30%) and sea level tweets (23%).

In summary, the science and climate change themes go somewhat hand in hand in the various discourses, however is not all scientific discussion related to global warming. Science has far higher importance in LTETs, while STET tweets only exhibit minimal interest. The following section will give a more detailed insight into how these distributions came to be.

4.1.4 Referencing of scientific studies in event tweets

The online sphere offers a great opportunity for the general public to interact with science. Subsequently, a growing majority of people uses the internet – and increasingly social networks – when inquiring about a specific scientific topic (Brossard & Scheufele, 2013: 40–41; López-Goñi & Sánchez-Angulo, 2018: 3). Despite the possibilities of the digital world, however, prices of scholarly journal literature have kept increasing massively and access to scientific information for the general public remains limited (Carroll, 2011: 1; Fecher & Friesike, 2014: 29). As a result, alternative sources for scientific information must be utilised.

Due to the observations above that direct science links are only used marginally, this section delves deeper into scientific content in the various corpora. It will investigate the most popular services employed by Twitter users when discussing science-related aspects of the different events. Research into referencing behaviour on Twitter can be carried out using URLs when external sources of information are of interest (Weller et al., 2011: 2). As such, the following elaborates on notably popular science-related web links in each year's tweets. There is some overlap with the timelines in chapter 4.3.3, however does this section solely summarise the referencing of *scientific* content.

Glaciers

2012 was a notable year for glaciology research due to the historically small extent of the arctic sea ice and the unprecedented melting of the Greenland Ice Sheet (Nghiem et al., 2012: 5; Parkinson & Comiso, 2013: 1356), which is reflected in the science-related articles shared on Twitter that year. Different than in other years, the most popular articles do not summarise recently published studies but report on current scientific findings. A popular BBC piece interviews scientists on the rapid melt of the Arctic sea ice (Shukman, 2012), while other news media point towards a NASA report visualising the unprecedented extent of surface melt on Greenland's ice sheet (Viñas, 2012). In the following winter and spring, references to the historic melt are made again when a study by Liu et al. (2012) links extreme weather events to the lack of sea ice covering the Arctic.

In general, studies reporting 'dramatic' or 'shocking' developments enjoyed popularity in the news media covering scientific findings. Ice sheets threatening to collapse and glaciers rapidly retreating were the most popular scientific findings in British glacier tweets between 2013 and 2015 (e.g., Mouginot et al., 2015) . Similarly, in 2018, the leading scientific story was the rapid rate of ice loss in the Antarctic (Shepherd et al., 2018). In 2019, British users frequently shared an article summarising findings from The Hindu Kush Himalaya Assessment report (Wester et al., 2019) highlighting the severity of expected ice melt in the Central Asian region.

French users showed great interest in an article covering the development of the Thwaites Glacier in Antarctica threatening to contribute considerably to global sea level rise (Milillo et al., 2019). Antarctica continued to peak the scientific interest of French users as well as British users in 2020 with a study detailing how a 2.5 m sea level rise is still likely even if the goals of the Paris climate agreement are met (Garbe et al., 2020). These studies did not push the importance of science in French glacier tweets to an extraordinary level, however did they have a large impact within the scientific sub-discourse during said years.

In 2017, a study by Zwally et al. (2015) arguing that the mass gains accumulated on parts of the Antarctic ice sheets compensate for the losses in other sections (contradicting some findings from the IPCC (2014) report) was shared many times by a small set of users employing the social media management software Buffer[11], possibly in an attempt to undermine public confidence in the scientific consensus on glacier shrinkage and sea level rise. Although the findings of the original paper were published in a reputable peer-reviewed journal, their methods have been questioned in the scientific community for delivering results inconsistent with a large body of previous research (Scambos & Shuman, 2016: 599).

Sea level

In 2020, Irish interest in the science surrounding sea level rise peaked when the Irish press picked up on a study defining Ireland – amongst others – as an especially threatened region by coastal flooding due to climate change (Kirezci et al., 2020).

French users showed a general interest in studies delivering updates on the development of global sea level rise with some overlaps to the scientific interests found in glacier tweets. Amongst the most popular sources are two more recent papers dealing with ice sheet contributions to sea level rise (Bamber et al., 2019; Garbe et al., 2020). Similarly, German users used the same source in 2017 for the discussion surrounding the science of both glacier and sea level rise (Reese et al., 2017).

Sections of a climate report by the United Nations World Meteorological Organization (2019: 16) on the development of sea level rise sparked the interest of some German users when the national press picked up on the reaching of a new record high. Analogously, the largest impact on Spanish users was

[11] Buffer is a social media tool allowing users to schedule the automatic posting of tweets.

made in 2019 when national newspaper El País summarised sea level rise-related findings from the IPCC Special report on the ocean and cryosphere in a changing climate (IPCC, 2019). The number of shares for this story far exceeds other science stories in Spanish sea level tweets, with smaller impacts being made by studies updating predictions for expected global sea level rise in 2013 (Cook et al., 2013) and 2018 (Dangendorf et al., 2017).

Sea level rise exceeding previous projections is a common theme amongst the most popular science news in Britain as well during the entirety of the study period, with most of the top science stories of the decade focussing on underestimations and predictions of new sea level rise trajectories (e.g., Cazenave et al., 2018; Horton et al., 2014). A popular study for glacier tweets in 2020 also topped the charts for sea level tweets the same year, as the referenced study strongly focusses on the connection between the two factors (Garbe et al., 2020).

Permafrost

Permafrost tweets mentioning science during the 2019 peak frequently have news articles attached, which summarise a study by Farquharson et al. (2019) stating that permafrost thaw is 70 years premature compared to predictions. The story was shared in all regions and is the main reason for the peak in science interest regarding permafrost in 2019 (visible in Fig. 19). In the previous peak from 2012 to 2015, no such stand-out study can be defined, and it is expected that the low number of permafrost tweets during those years has led to larger variation in the data. The only notable study found in tweets during that period is referenced by British users in 2014, describing the revival of a 30000 year old virus from the permafrost and the consequences permafrost thaw could therefore have on human health (Legendre et al., 2014).

Heatwaves

Due to the extremely low number of heatwave tweets containing references to science (0.25%), a very limited number of web links have been used by multiple users, making them stand out. The most referenced external source does align with the peak of science importance in the exceptionally hot summer of 2018 in Europe: A report published by researchers from the World Weather Attribution initiative (2018) highlighting the strong link between climate change and a heatwave over northern Europe garnered attention from British users. In the same year, German users showed interest in a scientific explanation for the rare co-occurrence of heatwaves across large parts of the northern hemisphere, however does the news article referenced (Ehlerding, 2018) only contain a short statement from a climate scientist and no scientific publication.

Wildfires

Similar to heatwave tweets, the discourse immediately surrounding wildfire events on Twitter is largely devoid of any references to science with only 1.8% of all wildfire event tweets containing a science-related keyword. This ratio has, however, seen an increase in recent years, peaking at 3.2% in 2020 (Fig. 19). As a result, there are also far more references to external sources covering science-related topics during the latter third of the study period. In general terms, it can be observed that science-related wildfire tweets reference a large variety of news sources covering current wildfire developments, which is to be expected with tweets being limited to the immediate time period surrounding events. Actual research publications therefore do not have a large impact on the importance of the science theme in wildfire tweets. The following highlights the few exceptions to this rule, disregarding the much more common basic news reports on current wildfires.

The Amazon fires dominated the wildfire discourse on Twitter during the summer of 2019, with the largest source in the scientific sub-discourse referencing a New York Times articles collecting scientific analyses of the events in the South American rain forest (Lai et al., 2019). The 2019-2020 Australian bushfires also caused some scientific discussion on Twitter especially in Great Britain, Ireland and Germany with popular articles in news media heavily focussing on the science surrounding global warming and stressing the devastating impact wildfires will increasingly have in the future if climate change continues unabated (e.g., Gergis, 2020). Whereas many of the popular articles often focused on new findings surrounding the impact of climate change (making use of the associated 'shock-factor'), Spanish users shared an article in 2020 detailing the scientific advances in combatting forest fires with cutting-edge technology (Varea, 2020).

To summarise, traditional news media has considerable leverage on the scientific discourses as a majority of science-related content was shared through links to online news articles. Especially findings that detail rapid and/or previously unexpected developments regarding the event types, often highlighting the urgency of the climate crisis, seemed to generally elicit the highest tweet activity.

4.1.5 Danger, risk, and health

Previously touched on in the context of the climate change theme, danger keywords are present in tweets talking about severe impacts of the respective events which endanger environment and humans alike. This theme offers an insight into the level of serious concern present in any of the discourses.

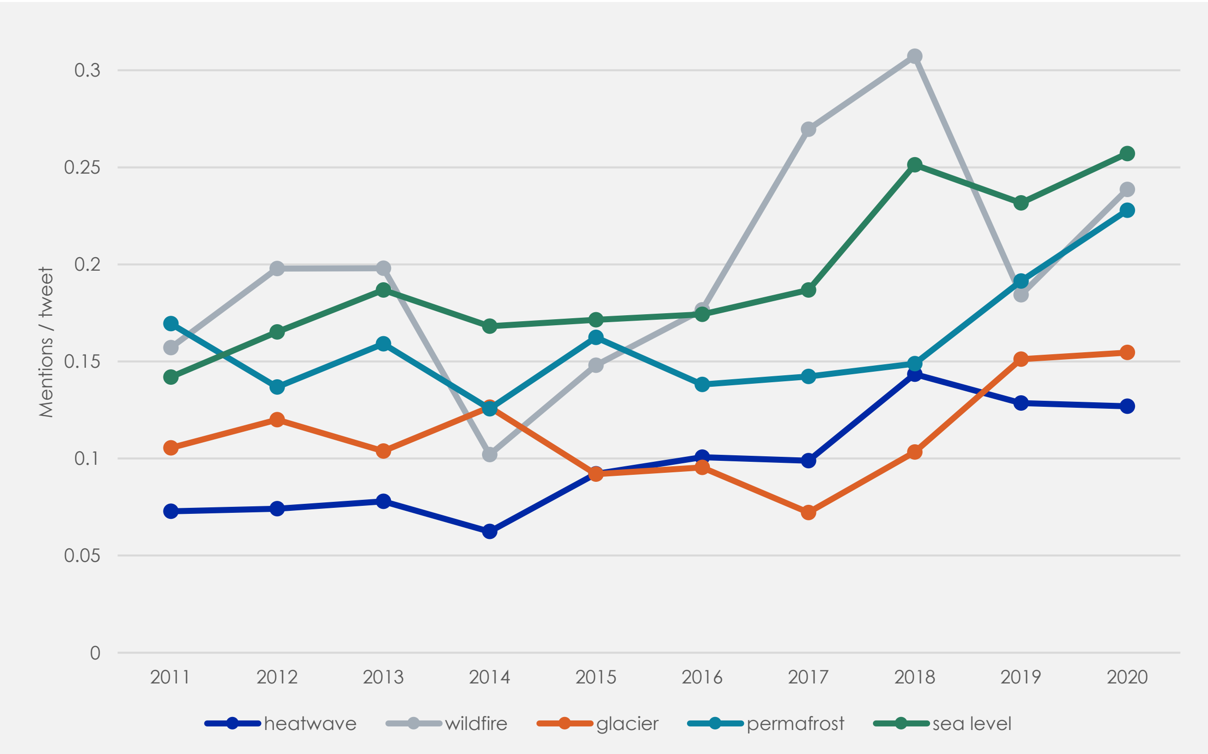

Fig. 20: Importance of danger keywords in event tweets.

While the concern about the dangers related to retreating glaciers has remained somewhat steady throughout the past decade, the other LTETs permafrost and sea level have seen a distinct increase over the more recent years after little change throughout the first half of the study period (Fig. 20). These developments are largely dependent on popular news articles and only rarely due to immediate environmental events, as the analysis in section 4.3.3 will show.

For STETs, however, danger perceptions can be investigated deeper here. The concern about the dangers of heatwaves rose during the hot European summers in 2018 and 2019, before flattening off again in the milder year of 2020. This increase was sharpest in French and English-speaking regions, whereas German and Spanish language areas remained at a level similar to previous years (Appendix, Fig. 53). Upon closer inspection, French users as well as authorities and news media seem very active in sharing warnings about upcoming heatwaves and extreme temperatures. This becomes apparent when looking at the most frequently used words during the heatwave in July and August of 2018 (Tab. 15).

The focus on danger-related themes during this specific event in France immediately stands out with keywords such as *alerte*, *vigilance* (often used in conjunction with *orange* to indicate a heatwave's level of danger) and *santé* being amongst the most important terms of the event. In contrast, German users seemed more concerned about environmental consequences. Especially droughts seemed to be a big topic at the time as there are four or five keywords directly and indirectly related to the issue present in the most important words (*Dürre, Trockenheit, Wasser, Regen,* and potentially *Bauern*). These terms are only partially classified as dangers (specifically *Dürre* and *Trockenheit*) due to the lack of context for the others. The reasons behind the disparity in danger perceptions between the regions can therefore largely be explained by the importance of different sub-discourses during heatwave events.

Tab. 15: Words that deviate the most in frequency during the July/August 2018 heatwave event from the overall word frequencies for all heatwave events of each country. Only words in the countries' main languages are featured as foreign-language words fluctuate much more and are assumed to be reliant on additional outside factors.

France	France (translated)	Germany	Germany (translated)
canicule	heatwave	*hitzewelle*	heatwave
paris	Paris	*dürre*	drought
orange	orange	*klimawandel*	climate change
vigilance	vigilance	*trockenheit*	drought
france	France	*sommer*	summer
orages	storms	*europa*	Europe
départements	Departments (regions)	*wasser*	water
alerte	alert	*regen*	rain
températures	temperatures	*mehr*	more
chaleurs	heat	*bauern*	farmers
santé	health	*ende*	end

The stand-out events when it comes to danger, however, are wildfires, which vary strongly in terms of concern throughout the study period. Overall, wildfires exhibit the highest value of all event types in 2018. This peak is likely due to the California wildfires that year in the United States, which were the deadliest fires of the state's history up until that point (Wang et al., 2021: 252). Similarly, the high danger perception in 2013 can be attributed to the bushfires in the Australian state of New South Wales in October that year, which were at the time the worst fires the region had experienced in recent years (Duc et al., 2018: 2). Not only was the environment influenced, but inhabited areas were affected as well with 248 properties lost (Rea et al., 2016: 151). The same trend can be observed during the high danger perceptions in 2012, when the US state of Colorado experienced one of its worst wildfire seasons in recent history, destroying over 600 homes and causing the evacuation of 32000 people (Alman et al., 2016: 1).

The low point in danger perception in 2014 seems to coincide with a lack of specifically devastating wildfires, as the event that resulted in the most Twitter activity that year – the San Diego county fires in May – 'only' burnt approximately 10% of the area of the above mentioned Colorado fires (NIFC, 2012: 62; Wang et al., 2016: 258). This reinforces the observation that especially destructive fires often threatening both environment *and* people seem to cause the highest danger perceptions.

4.1.6 Calls for action

The action theme includes tweets calling for action against climate change or processes related to it. This does not include specific actions dealing with the immediate consequences of a currently occurring event (e.g., firefighters extinguishing a wildfire), but focusses on solving the larger issue of climate change.

As a general trend, it is observable that calls for action have increased across the board from the start to the end of the study period (Fig. 21). When investigating calls for action more in depth for individual years and events, it becomes apparent that two subtypes exist: (A) Users stating arguments and sharing non-action-specific news articles or other external sources combined with hashtags such as #ActOnClimate or other keywords related to taking action and (B) organised protests/outrage elicited by real-world events. While the former can be attached to essentially any tweet regarding climate change, the latter gives more insight into actual events that elicit calls for action. The following will present some of these events.

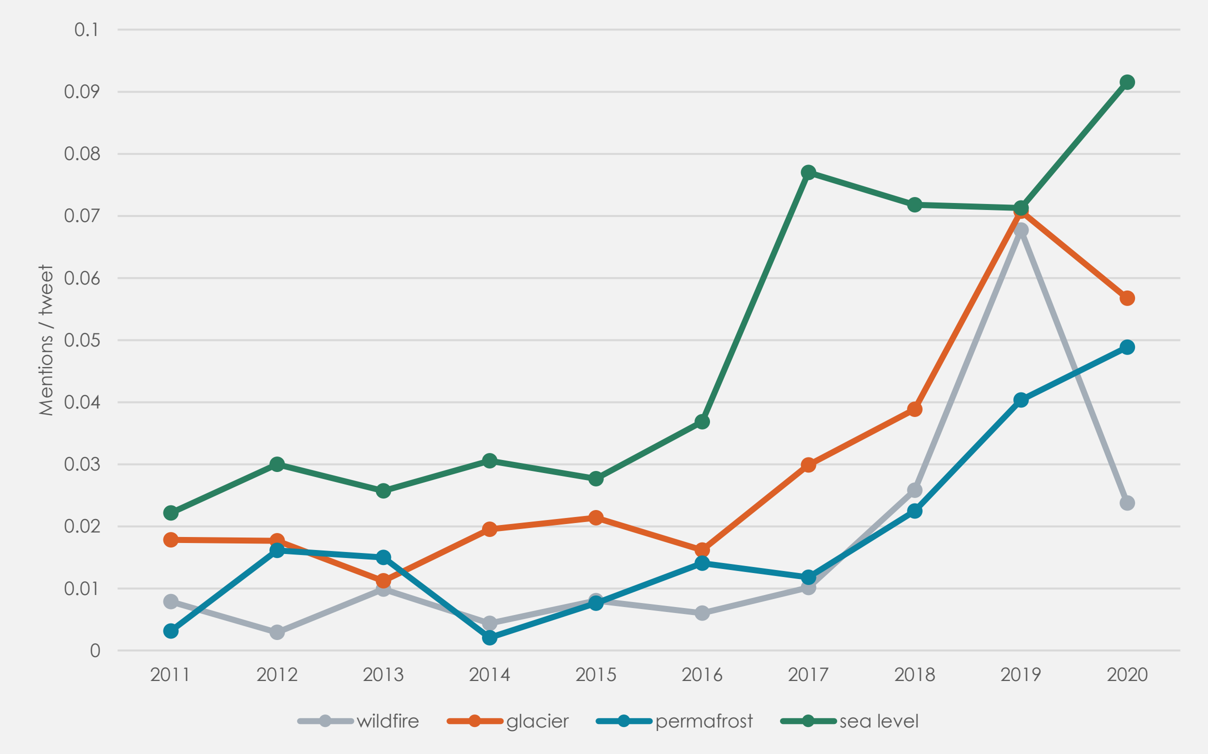

Fig. 21: Importance of action keywords in event tweets. Heatwave tweets are not included due to a lack of keyword mentions.

Wildfire tweets saw its most notable peak in 2019 with especially the Amazon fires causing calls for action across most parts of the study area, excluding Spain. Users frequently shared the #ActForTheAmazon hashtag when contributing to the discourse. Although not making much of an impact in the overall distribution (Fig. 21), Spanish users promoted the glacier protection law ('ley de protección de glaciares', #LeyDeGlaciares), a 2011 law passed in Argentina, in the early years of the study period. In more recent years, the sharp incline in calls for action related to glaciers leads back to the alpine countries of Switzerland and Austria. Swiss users were highly active in the discourse surrounding the glacier initiative ('Gletscher-Initiative'), a public initiative calling for the reduction of fossil fuel consumption, while Austrian users were concerned about photos depicting construction equipment digging on Austrian glaciers (in context of the merger of the Pitztal-Ötztal ski areas), which had gone viral on social media.

Permafrost and sea level tweets saw no perfectly distinguishable events that elicited calls for action as peaks were usually tied to the wording of news headlines paraphrasing research findings. Instead, it seems that there is just a relatively high number of tweets calling for action outside of a specific event context (i.e., subtype A from above).

In summary, the considerable increase in calls for action observable throughout the latter half of the study period seems to be driven by a mixture of growing numbers in specific political movements and initiatives (subtype B) on one hand and a general rise in general support for climate action (subtype A) on the other hand.

4.1.7 Conversation, controversy, and tweet sharing behaviour

Compared to traditional media, social media platforms lend themselves – especially for minorities – to build so-called ‘counterpublic spaces’ in which opinions and perspectives not shared by the majority of the public can be discussed and propagated (Moernaut et al., 2020: 2). As a result, online debates tend to become polarised and develop into ‘echo chambers’ with ever-diverging viewpoints, which is certainly the case in the discourse surrounding climate change (Williams et al., 2015: 135). Moernaut et al. (2020: 2) suggest that these characteristics should lead to highly contested debates when events such as heatwaves are linked to climate change either by Twitter users themselves or by media reports they share.

This section will investigate these conversation characteristics by first measuring which events and themes generated the most debate before continuing on to quantifying the level of controversy in each of the discourses. Finally, the importance of the discourses is discussed using tweet sharing behaviour.

Conversation and controversy

Conversation tweets (CvT) – as opposed to regular tweets – are posts engaging in conversation by making use of either the ‘reply’ or ‘retweet with reply’[12] functions on Twitter. Overall, Fig. 22 shows that there is a slight split observable between STETs and LTETs, with the former showing lower CvT percentages, meaning users tweet statements rather than engaging in conversation more often. Inversely, one could also say that LTETs are brought up in conversation more often than STETs. Interestingly, this difference is evened out when only looking at tweets belonging to the climate change theme. This suggests that all five event types have a similar argumentative value regarding the theme, as they are brought up in conversations surrounding climate change a similar amount.

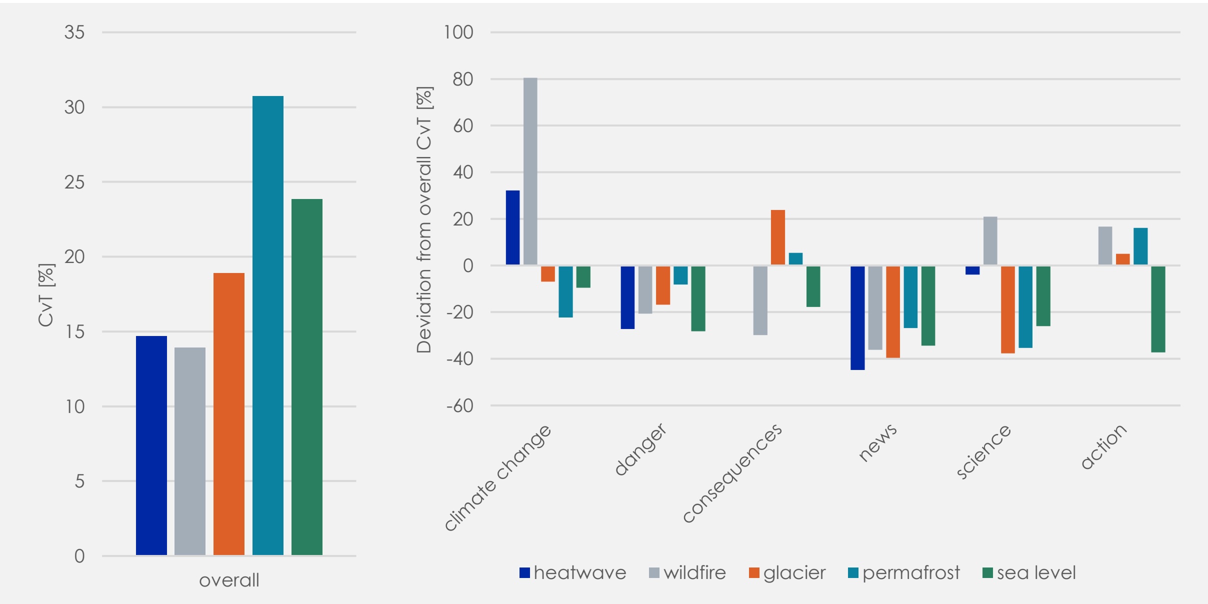

Fig. 22: Percentage of all tweets that are conversation tweets (CvT) for the entire corpus (left) and tweets belonging to individual themes (right). Deviations (right) represent the deviation from the overall CvT in percent.

Further trends show that the sharing of news and science information (as well as dangers) tends to happen less in conversation and is more of a statement or provision of information made by the users. Only for wildfire tweets is the trend reversed when it comes to the science theme. This could suggest that the science surrounding wildfires is more subject to debate amongst Twitter users and not as set in stone as the science regarding LTETs (which have very low values). Although there is considerable scientific consensus regarding wildfires, the corresponding politics are strongly influenced by opinions and largely

[12] Also known as ‘quote’

disregard scientific research (Leverkus et al., 2020: 416–417). These political aspects may in turn affect the public's interaction with science on the topic as the divide between the two becomes apparent.

The consequences[13] of especially glacier melt stand out as they are used more often in conversation than the baseline. This is potentially connected to the fact that the consequences of glacier processes could be considered to be an effective argumentative tool (Borick & Rabe, 2010: 785–786). The generally higher values for conversation regarding calls for action might indicate that users react to information provided by other users with the realisation that action is necessary.

Having only looked at the percentage of conversational tweets so far, there is also the possibility to approximate a tweet's level of controversy. Fig. 23 shows the percentage of such controversial tweets (CtT) in the respective corpora, demonstrating that only a very small fraction of tweets can be deemed controversial overall with the maximum being set at 2.6% in heatwave tweets.

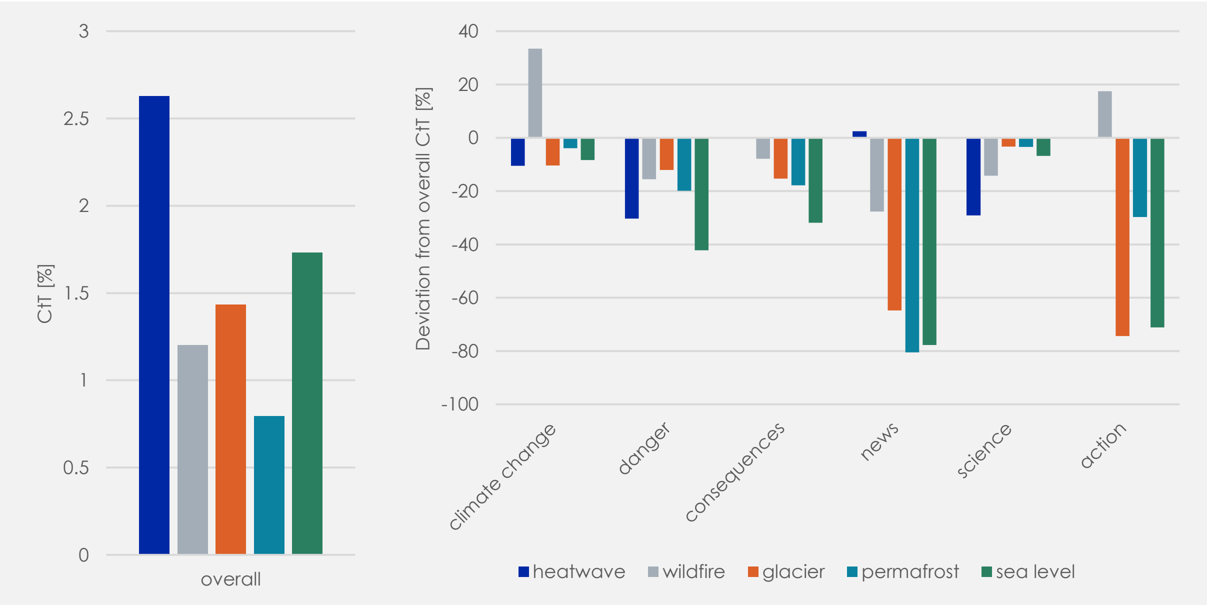

Fig. 23: Percentage of all tweets that are arguably controversial tweets (CtT) for the entire corpus (left) and tweets belonging to individual themes (right). Deviations (right) represent the deviation from the overall CtT in percent.

Looking into the individual themes gives an interesting indication of how controversial different aspects of the discourse on Twitter are. Generally, the conversation surrounding the selected themes remains below the overall values, meaning that none are exceptionally controversial in the bigger picture. When compared between each other, however, some patterns emerge. One of the stand-out values is certainly the relatively high controversy surrounding the aspect of climate change in the wildfire discourse. It seems that the active conversation about this topic found above is indeed also controversial to some degree. Climate change is overall one of the more controversial themes on the list.

A further interesting aspect is the news-science relationship for LTETs as it appears that news tweets are far less controversial compared to science tweets. As both themes scored similarly in terms of pure conversation percentage (Fig. 22), it is unlikely that news articles simply lend themselves more to sharing (due to sensationalistic headlines etc.), while science-related content incites more discussion. It is possible that this disparity represents the interface of science sceptics and pro-science members of the public, which would also align well with the rather controversial values of the climate change theme. Traditional news media tends not to give climate and science sceptics much of a stage (Moernaut et al., 2020: 2), which would explain the other side of the disparity. These results have to be taken with a grain of salt, though, as the measure for CtT is not absolutely definitive.

[13] 'Consequences' is a theme encompassing direct consequences resulting from the various effects (i.e., the release of methane during permafrost thaw).

Tweet sharing

We can judge the importance other members of the public besides the author ascribe to a tweet by measuring how often a tweet is shared through a retweet or quoted tweet. Fig. 24 shows that permafrost tweets generate the biggest response, while heatwave tweets often do not result in many interactions.

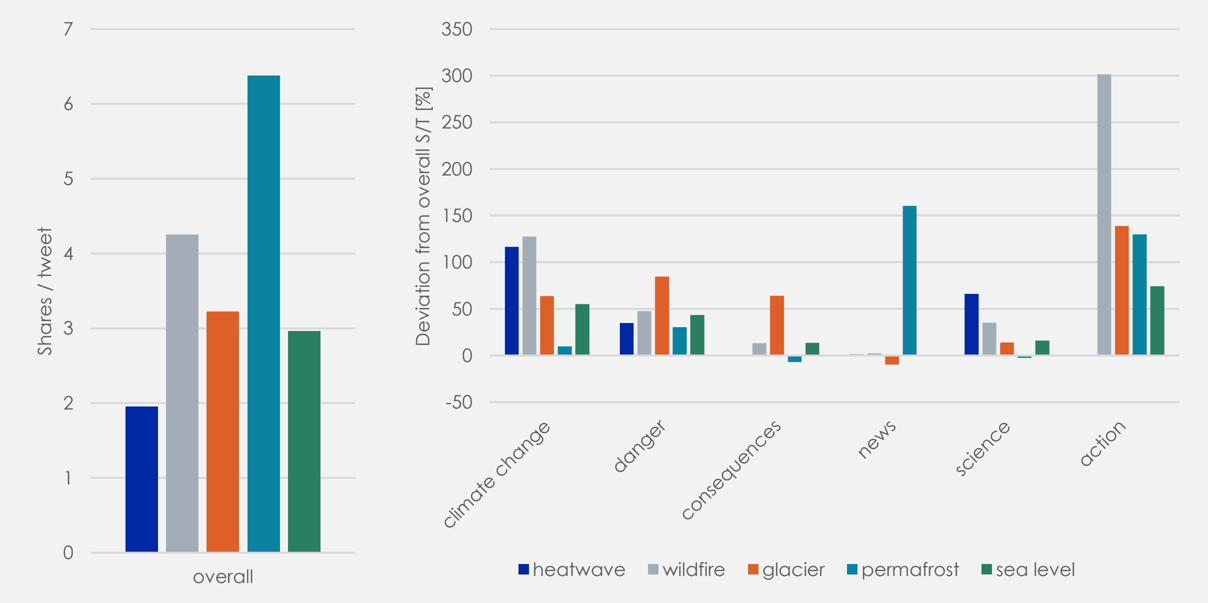

Fig. 24: Mean number of shares (retweets or quote tweets) per tweet in the entire corpus (left) and the deviation from this overall mean for individual themes (right). Deviations (right) represent the deviation from the overall shares per tweet in percent.

Interestingly, tweets which include one of the selected themes seem to generally elicit a much stronger response than the overall values. This is especially true for climate change and calls for action, which result in much higher sharing activity across the board. Consequences of glacier melt were defined as a point of conversation above and they once again separate themselves from the other event types here as they also seem to be a factor in tweet sharing. In all cases except for permafrost, science tweets are shared more often than news tweets. Whether ascribing importance to a tweet is the main motivation behind the sharing behaviour ultimately, however, remains unclear.

4.2 Spatial dimension

The study area encompasses not only seven different countries, but also spans a variety of biomes from the high alpine regions of the Alps to the deserts of Spain and the agricultural lands and forests in much of central-western Europe. The following chapter will dissect spatial trends in tweet activity regarding national, subnational and environment-driven distributions in an attempt to define which spatial components elicit the strongest public reactions to climate change effects.

After an overview of the general spatial distributions, their change over time will be quantified. Due to the global nature of online networks, interests crossing international borders must also be considered and are measured by the corresponding degree of transnationality. A final section investigates the influence of environmental factors relevant to the event types.

4.2.1 Spatial distribution

Heatwaves

The spatial distribution of heatwave tweets (Fig. 25) exhibits a decently strong southwest to northeast directional trend with tweet frequency generally decreasing along this axis. Tweeting activity surrounding heatwaves is therefore the highest in Spain, followed by parts of France. This distribution shows quite strong similarities to trends in extreme temperatures, as will be discussed in section 4.2.4 (Fig. 36). As a general trend across the study period, it can therefore be concluded that people experiencing hotter temperatures more often also react more strongly. This aligns well with expectations that tweet activity surrounding heatwaves is strongly tied to a user's immediate environment.

A second trend emerges within the 'colder' countries, where city regions such as Hamburg, Berlin, Vienna and London show generally higher tweet frequencies then surrounding regions that include a much higher proportion of rural areas. The reasoning for this could be twofold: Firstly, urban areas have a higher density of Twitter users than rural areas (Johnson et al., 2017: 1169; Perrin, 2015: 9), therefore leading to a slight overestimation of tweet frequency in city states. The strength of this bias is, however, weakened by the fact that many of the other regions also include one or more larger urban centres, making them not purely rural regions. The second factor explaining the distribution is the urban heat island effect, a well-known process leading to excess temperature – amongst other issues – in densely populated areas compared to their rural surroundings (Kleerekoper et al., 2012: 30).

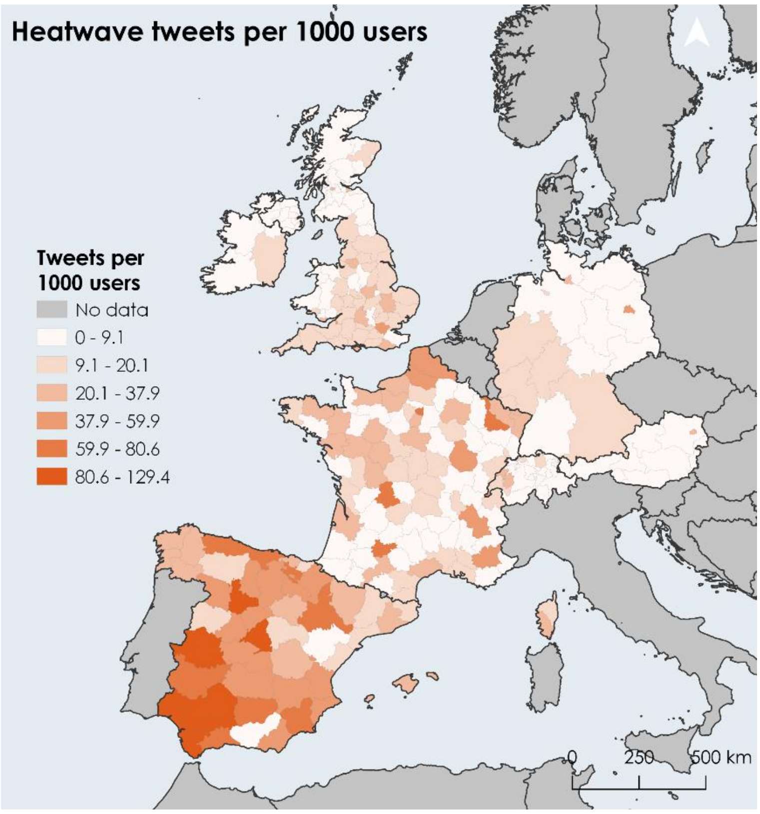

Fig. 25: Heatwave tweets per 1000 Twitter users during the entire study period in each subnational division.

This also explains the higher activity in the central west of Germany, which contains quite a high number of cities, and the overall distribution in the British Isles, where tweet frequencies are generally higher in urban regions than majority rural areas. Whereas much of England is quite active regarding heatwaves (with peaks in London, Bristol, Manchester, Oxford, Cambridge), activity only spikes in the urban areas of Scotland (Edinburgh, Glasgow, Aberdeen), Wales (Cardiff), and Ireland (Dublin). Switzerland is somewhat of a mixed bag with the French-speaking southwest generally being the most active besides the canton of Zurich.

Wildfires

The distribution of wildfire tweets (Fig. 26) appears rather scattered in comparison to heatwave tweets. Although Spain as a whole still displays the largest number of tweets per user, the values are not quite as uniform across the country. The absolute peak can be observed in the Galicia region in north-western Spain. Especially in the years 2016 and 2017, tweet numbers peaked in the region due to large wildfires in the area. The Baleares islands experienced multiple wildfires in the study period such as the 2011 fire on Ibiza and the 2013 fire on Mallorca, leading to a quite high tweet activity surrounding the Mediterranean islands. Corsica is also strongly represented as it experienced extensive wildfires in 2017. North-eastern Germany shows a clear deviation from the remainder of the country, which is mostly due to large forest fires in the state of Brandenburg in 2018 and 2019 (BLE, 2021).

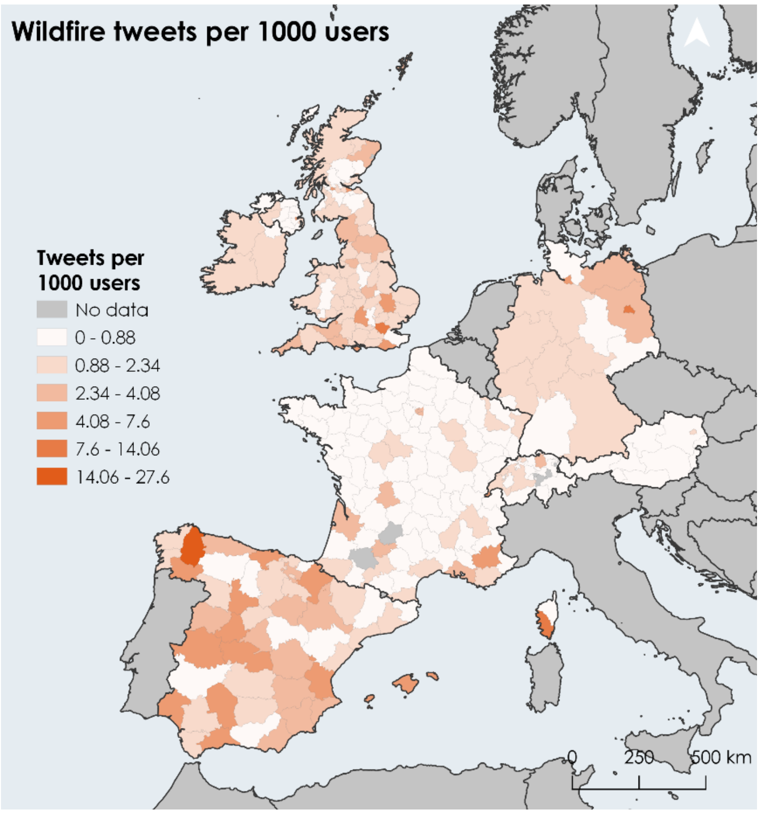

Fig. 26: Wildfire tweets per 1000 Twitter users during the entire study period in each subnational division.

The alpine countries of Austria and Switzerland show distinctly different behaviours. The common lack of larger scale wildfires in the two countries seems to have led to two different types of reaction: In Austria, users only reacted to domestic fires in 2014, when a discarded cigarette caused the largest forest fire in 20 years covering roughly 50 ha of forest (BOKU, 2020). For comparison, the Galicia fires mentioned above burned an area of 42314 ha during a single week in October 2017 (Chas-Amil et al., 2020: 4). Besides this event, international wildfires only caused very limited reactions amongst Austrian users on four other occasions between 2018 and 2020. The difference with Swiss users is that despite a similar lack in large domestic fires, interest in international events is much higher, and moderate to high tweet activities can be observed in regions such as Zurich, Bern and especially Geneva. This is likely because Swiss wildfires mostly occur in the Italian-speaking regions of the country, which are not covered language-wise in this thesis. For the other cantons, the burnt areas from fires are restricted to hectares in the single or double digits, meaning that there is not much local wildfire activity to react to for the Swiss population (Pluess et al., 2016: 228).

Similar to Switzerland, British users seem to strongly react to international events with fires in English-speaking countries such as Australia, the United States and Canada eliciting the three largest reactions of Londoners, which overall show quite high interest in wildfires. Regarding the list of other British regions with high tweet activity, the same trend can be observed for Oxfordshire, Cambridgeshire, Edinburgh, and Glasgow. Belfast is the only city in this list where a domestic fire made it into the top three most discussed events of the study period with the fire in Trollymore, Northern Ireland in 2019.

While the locations and importance of domestic versus international fires differ across the study area, it is very clear that wildfire tweet activity is directly caused by the occurrence of wildfires. Investigation of the tweets within detected tweet clusters shows that this direct connection is undeniable, further validating previous expectations that wildfire tweets – much like heatwave tweets discussed above – are strongly dependent on events occurring on a short time scales surrounding the corresponding tweets (see also section 3.3.5, Wang et al., 2016: 529).

Glaciers

Two groups of areas seem to stand out when it comes to a high interest in glaciers: Firstly, regions containing some of the largest cities in Europe tend to exhibit some of the highest values, especially when they are closely tailored around the cities (e.g., Paris, London, Berlin, Hamburg, Vienna). When cities are part of larger areas (e.g., Munich, Madrid), the trend does not seem to translate as values might be averaged out by the surrounding population.

A second observation is that there does seem to be a loose cluster around the Alps with an unusually dense set of high-value subdivisions spanning from south-eastern France over large parts of Switzerland to western Austria. This suggests that the distance to a mountain range containing glaciers is one factor determining the level of contribution to the glacier discourse on Twitter. It is noteworthy that the Pyrenees on the border between France and Spain do not elicit the same trend, although they do contain some glaciers. This is likely due to the fact that the glaciated area in the Pyrenees is minimal and several orders of magnitude smaller than in the Alps: At the end of the 20th century, a total area of only approximately 5 km^2 contained glaciers in the Pyrenees (Rico et al., 2017: 5), while glaciers in the Alps were extrapolated to cover more than 3000 km^2 (Paul et al., 2004: 2).

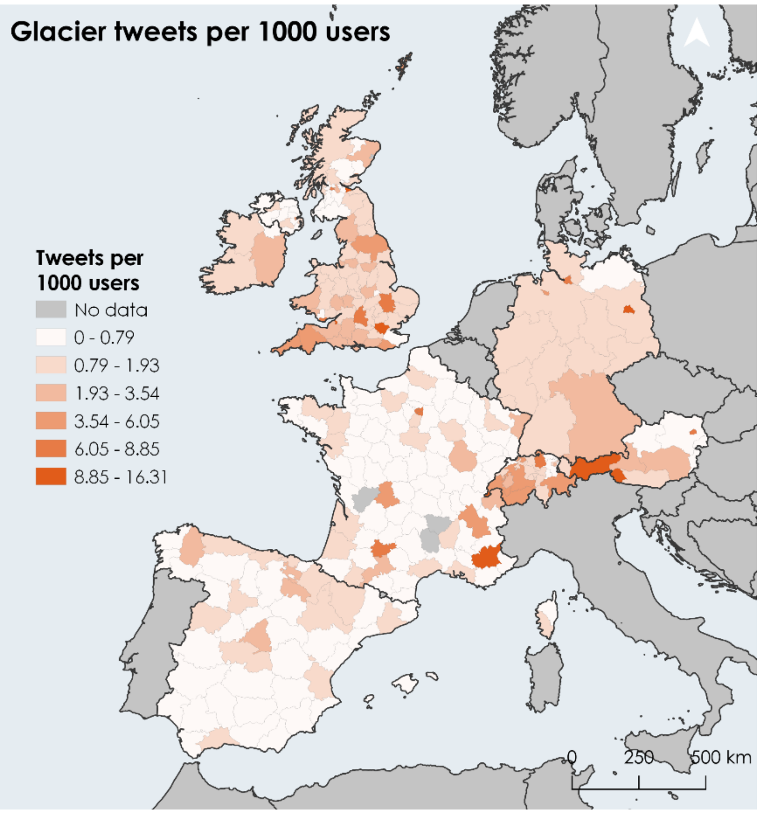

Fig. 27: Glacier tweets per 1000 Twitter users during the entire study period in each subnational division.

Interestingly, the British Isles remain strongly represented when it comes to glacier tweets despite their lack of local glaciated areas. Observations of the UK's and Ireland's placing in the various themes shows that they rank amongst the highest for mentions of consequences and dangers (Appendix, Fig. 54), implying that for example glaciers' contributions to sea level rise and the resulting risks for the British Isles could be the reason why this event type is discussed so frequently. This conclusion is likely as sea level rise not only plays an important role in the glacier discourse on Twitter (see section 4.1.1), but will also turn out to be especially important to British and Irish users (later in the current section).

Permafrost

The distribution of permafrost tweets (Fig. 28) is generally focussed on the bigger cities as they take up a large majority of the 20 highest tweet frequencies observed in the study area. It is worth noting that the range of values is quite small and there are no extreme outliers.

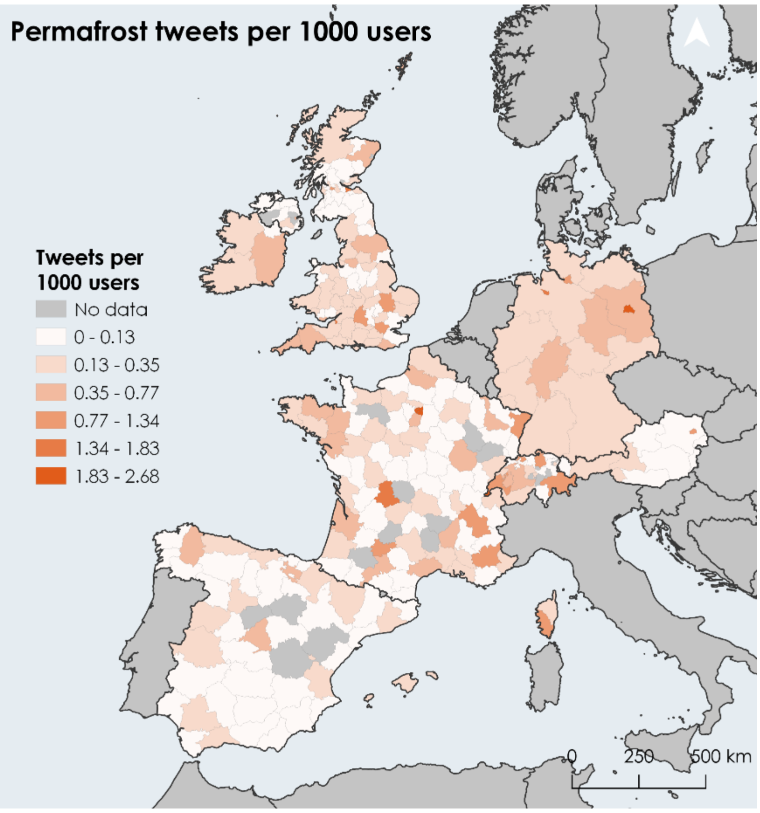

Fig. 28: Permafrost tweets per 1000 Twitter users during the entire study period in each subnational division.

Similar to glacier tweets, but not quite as strongly pronounced, there is somewhat of a cluster around the Alps, which is especially driven by a quite consistently high tweet activity in Switzerland. Inversely, especially Spain shows a general lack of interest in permafrost with some areas of France also showing a similar trend. For high-scoring cities such as Berlin, Paris, and Bristol, there is no indication of local influences on tweet frequency as a majority of tweets discusses global issues associated with permafrost melt reported in news media. Interestingly, the outlier in Haute-Vienne, central France is attributable to only a handful of prolific users that continuously tweeted about permafrost throughout the study period[14]. Overall, the distribution of permafrost tweets as represented in Fig. 28 seems to indicate that spatial aspects only have a very limited impact on the public's reaction to permafrost.

Sea level

Besides the bias towards the highly urbanised regions discussed previously, there seems to be a clear trend in the spatial distribution of sea level tweets as the British Isles exhibit some of the highest activity rather continuously across most of its regions. While Germany also has a continuous distribution, the only notably high values are found in the northern city states. Landlocked Austria and Switzerland exhibit generally low values, with exceptions once again in metropolitan cores. France and Spain are mixed bags regarding tweet activity surrounding sea levels. When investigating the individual tweets of regions in these countries with higher values, there is no trend observable that links the high values to a specific aspect of the regions. The peak in Haute-Vienne is once again due to the same set of prolific users that showed high activity regarding permafrost.

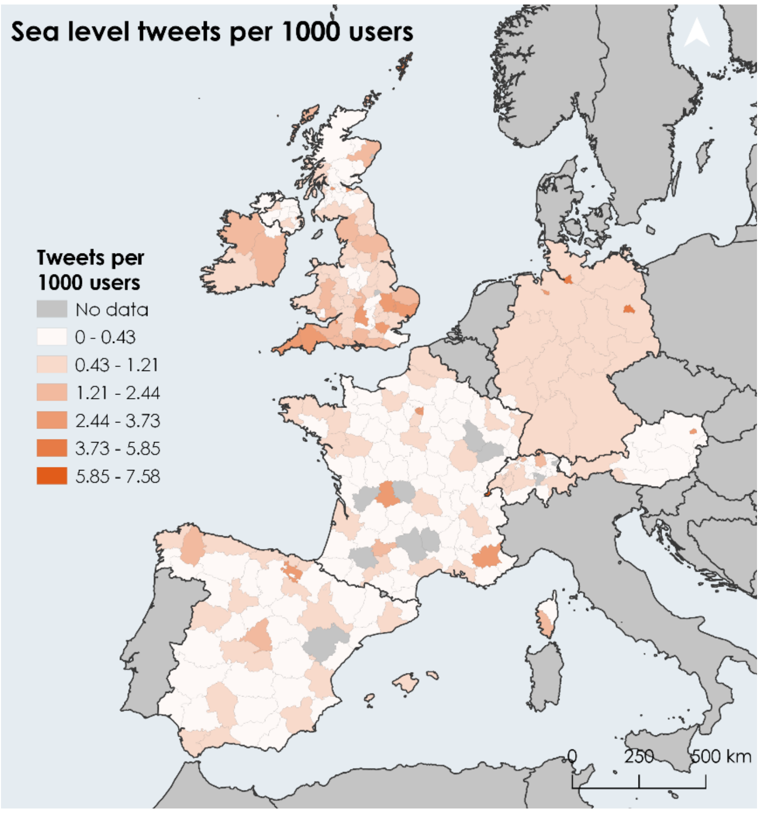

Fig. 29: Sea level tweets per 1000 Twitter users during the entire study period in each subnational division.

Only in some British regions such as Cornwall, Devon, and Suffolk can a direct connection between coastal flooding events and high tweet frequency be drawn. It is highly likely that these types of events also contributed to higher tweet frequencies across the rest of the country.

An obvious spatial question regarding sea level rise is the influence of coastal proximity on public reactions, as one might assume that people living closer to the coast might be more concerned about potential consequences. When comparing the tweet frequency (normalised by population) between all coastal and landlocked subnational regions, there does in fact seem to be a coastal bias regarding the magnitude of the public's reaction (Fig. 30). The median number of 0.61 tweets per 1000 users is in coastal areas is considerably higher than the one in landlocked regions (0.24).

This trend also holds true when the comparison is applied to the national level. Similar proportions to those in the overall medians can be observed in Spain and France, where coastal areas exhibit considerably higher values. In Great Britain and Germany, the difference is not quite as pronounced with coastal areas only showing a smaller lead over landlocked areas. In the case of the latter, it is possible that the

[14] The tweets are not duplicates of each other and the tweet sources indicate no bot (automated posting carried out by software) interference, thus making them all valid tweets.

large spatial extent of German states influences the values for coastal regions as they tend to extend much further inland than subnational divisions in other countries of the study area. However does there not seem to be a large difference between the coastal and landlocked areas of the country in general (Fig. 29).

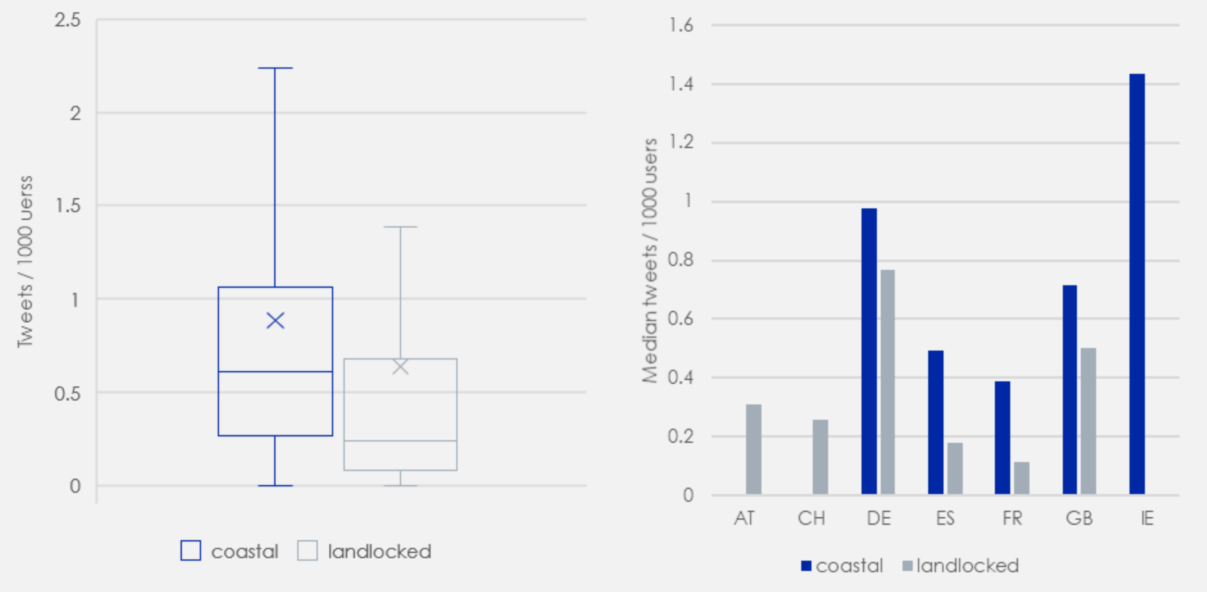

Fig. 30: Sea level tweet frequency separated by coastal and landlocked regions in a) the entire study area and b) individual countries. The outliers in plot a) were omitted to increase readability.

4.2.2 Spatiotemporal patterns

As shown above, some spatial patterns emerge from the different distributions of event tweets. As the study period encompasses a rather long time span in which themes and interests can develop and vary strongly, the question remains whether regions prevail close to the average values displayed in the maps above or if they experience considerable variation from year to year. This variation can be quantified by creating an annual ranking of regions (sorted by number of tweets per users) and then observing how much this relative ranking changes over time for each region.

Tab. 16: Variations in subnational region rank over the study period. Mean rank standard deviation (MRSD) is defined as the standard deviation in annual rankings (by number of tweets per user) of a region, averaged across the study area. The correlation between RSD and overall mean tweets per user in individual subnational areas is listed with the corresponding Pearson's r and p-values.

Event type	MRSD	MRSD (top 20)	Pearson's r	p-value
Heatwave	24.54	17.15	-0.14	0.01602
Wildfire	27.02	21.07	-0.16	0.00733
Glacier	22.33	4.76	-0.27	0
Permafrost	27.39	14.65	-0.22	0.00012
Sea level	25.20	11.00	-0.24	0.00003

As Tab. 16 shows, there is some considerable annual variation in the approximately 300 subnational regions. Over the ten years of the study period, their relative ranking amongst each other changes on average with a standard deviation of roughly 25 ranks. This suggests that regions' rankings do vary over time, however do they tend to stay in a certain range and display some level of consistency. When investigating the top 20 regions in terms of overall tweet frequency, the consistency is decisively higher, which means that regions with high interest in certain event types tend to remain at this high level much more consistently than regions in the middle or the bottom of the ranking. As a result, it can be concluded that there seems to be some baseline spatial dependence driving high tweet activity.

This dependence varies considerably across the different event types, however. The fact that LTETs exhibit a higher consistency than STETs for the top 20 is unsurprising for multiple reasons: First, the spatial distribution of heatwave and wildfire events in the real world is decidedly more random considering that glaciers, permafrost and sea level are all bound to fixed natural features or regions. While the spatial influence of the latter would therefore remain even across the study period, the former could introduce considerable variation as events occur either in a variety of locations or across large swaths of land, further limiting the small-scale spatial influence. Second, there is potentially a demographic component regarding the users tweeting about the different event types. While reacting to heat and fire events could be considered universal, tweeting about glaciers, permafrost and sea levels arguably requires a more specific interest in the topics. If these interests are spatially tied to institutions such as universities or demographic aspects of regions, the spatial influence on tweet activity is strengthened. As research on climate change has shown, economic factors as well as urban/rural differences can certainly impact a person's interaction with climate change and related issues (Hamilton & Keim, 2009: 2348; Ruiz et al., 2020: 114; Semenza et al., 2008: 481–482).

When looking at the correlation between MRSD and overall tweet activity, only a small negative trend can be observed. This suggests that across the entire study area, changes in ranking are hardly dependent on a region's position in the ranking. As shown above, however, this dependence is much stronger in the higher ranks, therefore suggesting that it strongly decreases in the mid and lower ranks.

4.2.3 Transnationality

Until this point, I have investigated where people *are* when they join the discourse on the different climate change effects. Transnationality augments this approach, as it identifies what people are talking *about* in the spatial sense by drawing a connection between their assumed home and places they mention in their tweets. This degree of transnationality can be measured using the *E-I index*, which determines the ratio of domestic links to transnational links in the network (Krackhardt & Stern, 1988: 127; Reber, 2020: 3, 7). Ranging between -1 and 1, positive values indicate more transnational than domestic connections and negative values the inverse.

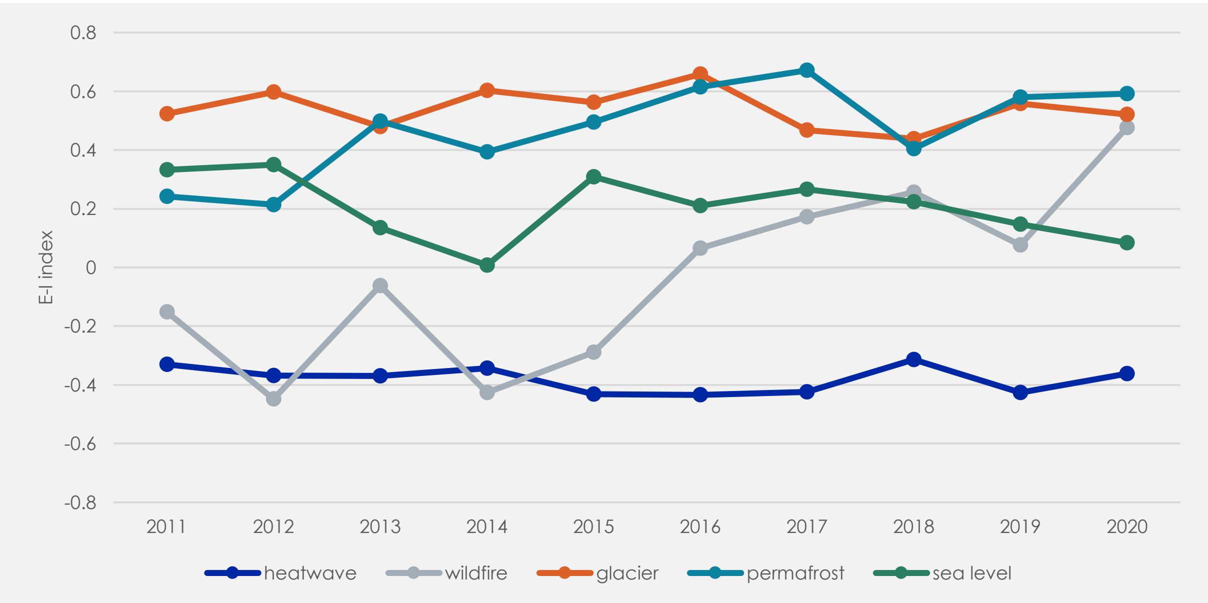

Fig. 31: Transnationality measured by the E-I score for all tweets of every event type. Positive values indicate more transnational than domestic connections and negative values the inverse.

A broad overview of annual E-I scores (Fig. 31) gives an interesting insight into some general characteristics of tweeting behaviour associated with the different event types. Once again, the separation between LTETs and STETs becomes apparent, with the much stronger domestic focus of the latter. This distinction interestingly diffuses throughout the second half of the study period with wildfire tweets increasingly referencing international events, even surpassing sea level tweets in this aspect. Glacier and permafrost tweets remain rather constantly at a highly transnational level, while sea level tweets hovering only slightly above an even split between domestic and transnational mentions. This reinforces the idea that people generally react more strongly to events in their proximity, however does the transnational development of wildfire tweets seem to indicate that international events can be discussed more heavily depending on the circumstances. The differences in LTETs also highlights the fact that across the study area, sea level rise is the LTET that is most frequently discussed in a domestic setting. When looking into the distributions within each event type, clear differences between countries can be observed. The following sections will summarise the corresponding findings.

Heatwaves

Interestingly, there does seem to be a distinct difference between the countries regarding transnationality in heatwave tweets (Fig. 32). While Switzerland and Germany hover around 0, their German-speaking compatriot Austria exhibits the strongest domestic trend of all with values around -0.8. In between them lies a cluster of France, Spain, and Great Britain, which stay between -0.4 and -0.6 and do not deviate strongly from each other.

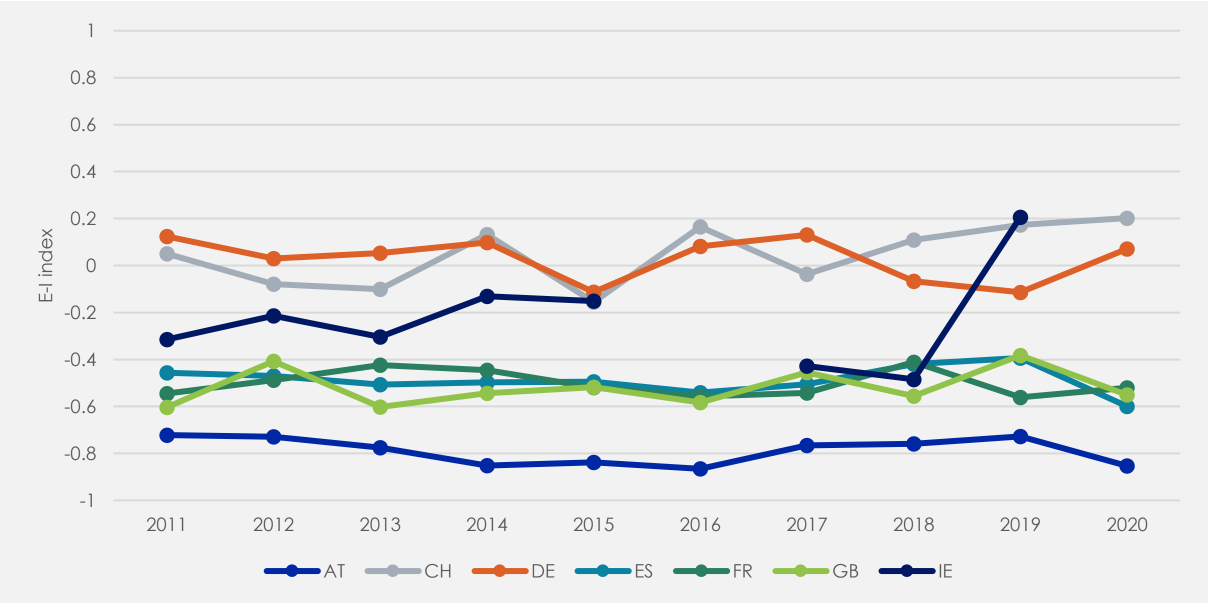

Fig. 32: Annual transnationality (E-I score) per country for heatwave tweets.

A look at the countries mentioned in the transnational connections shows similar trends for most regions in the study area. The strongest connections typically exist in western Europe, with especially strong ties between countries sharing an official language. Northern Europe is mentioned much less (possibly due to lower temperatures) and Eastern Europe is similar. When it comes to other continents, it is usually the largest countries that are mentioned most frequently, with the United States being exceptionally important. Australia, India, China, and Brazil are common connections for most countries, while Spain, as expected, has very strong ties to large parts of Spanish-speaking Latin America.

Wildfires

While values essentially remained the same for each country regarding heatwave tweets, strong variations can be observed in some cases for wildfire tweets (Fig. 33). This might be due to the fact that location tagging is an essential part of wildfire discourse. Vieweg et al. (2010: 1084) found that people affected by wildfires were referencing geo-locations more frequently in their tweets compared to those experiencing floods.

French users start off with a strong transnational interest for the first three years, before the value flips and goes strongly domestic between 2015 and 2017. The following year, it jumps back to a strongly transnational position and stabilises there. An investigation into word frequencies of tweets published during the domestic phase shows that fire events in or near Gironde, Marseille, and Corsica (all southern France) were of high interest during this time and elicited extensive discussion on Twitter. These events also account for some of the largest single-day tweet peaks in France during the study period. The following years, fires in California, the Amazon and Australia pull the trend back into a transnational position in the absence of large reactions to French fires. Only when another extensive wildfire breaks out near Marseille in 2020 is the score somewhat trending down again towards more domestic values. That year, however, the parallel occurrence of multiple large-scale wildfires around the world outweighs the domestic event.

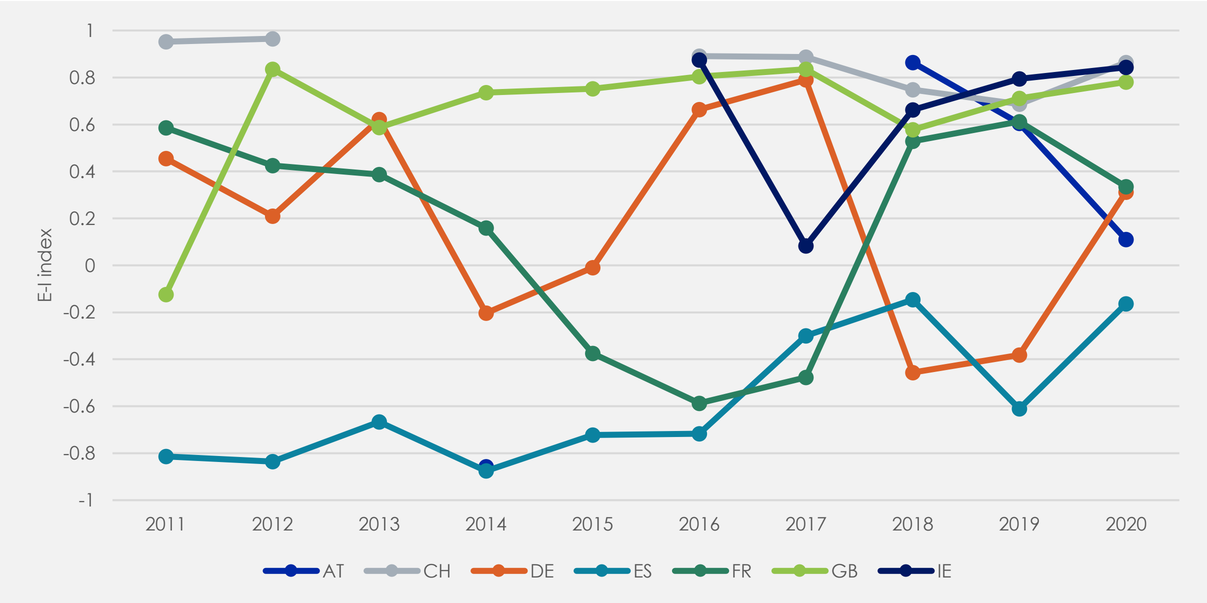

Fig. 33: Annual transnationality (E-I score) per country for wildfire tweets.

Germany is another country which experienced considerable variability in E-I score throughout the study period. The generally more transnational leaning country broke slightly below 0 on two occasions in 2014-2015 and 2018-2019, before returning to strongly transnational scores the following years. While the investigation of tweets during these phases shows that some smaller local fires helped the first domestic trend, it was especially the more extensive Brandenburg wildfires in 2018 (Treuenbrietzen) and 2019 (Jüterbog) that caused increased domestic interest. The much quieter German wildfire season in 2020 similarly explains the subsequent return to more transnational tweets that year. These observations are supported by a decently strong negative correlation (Pearson's $r = -0.71$, p-value = 0.02) between the annual area burnt in Germany (BLE, 2021) and the corresponding annual E-I scores.

Spain, which sees some of the largest domestic wildfires in the study area (Eriksen, 2020: 2), unsurprisingly exhibits the strongest domestic trend. This, however, increasingly develops towards a more transnational discourse starting in 2017 and almost reaches a neutral score in some of the following years. This is likely due to the previously mentioned occurrence of large-scale fires in Australia and the Americas during this timeframe. As previously discussed, Switzerland experienced minimal wildfires in the linguistic regions covered in this thesis, thus – as expected – showing a strong transnational score.

Regarding the spatial distribution of transnational relationships, the United States, Australia, Brazil and Canada (often in this order) are consistently the biggest intercontinental connections for any of the countries in the study area. This is expected, as these countries experienced some of the most devastating wildfires in recent years (Xu, Yu, et al., 2020: 2173). Within Europe, Greece and the Iberian Peninsula stand out with some interest in Sweden as well, which aligns very well with larger wildfire events in

Europe during the study period (Eriksen, 2020: 2). Transnational interests are therefore directed towards regions experiencing extreme or frequent fire events.

Glaciers

It is unsurprising that the alpine countries Austria, Switzerland, and France exhibit the lowest E-I score, thus showing the strongest interest in domestic locations (Fig. 34). The remaining countries all show distinctly transnational tendencies. In terms of the spatial distribution, transnational connections most often lead to Greenland, which seems to be of far greater interest than Antarctica in all countries except Spain. Besides the ice sheets, North America is strongly represented, whilst South America is once again largely limited to the interest of Spanish users. In the European context, there are increased connections to the countries encompassing the Alps as well as Iceland, while other glaciated regions such as Scandinavia or the Caucasus see much less interest overall. Countries in the Himalayas consistently lag behind Europe and North America in terms of mentions.

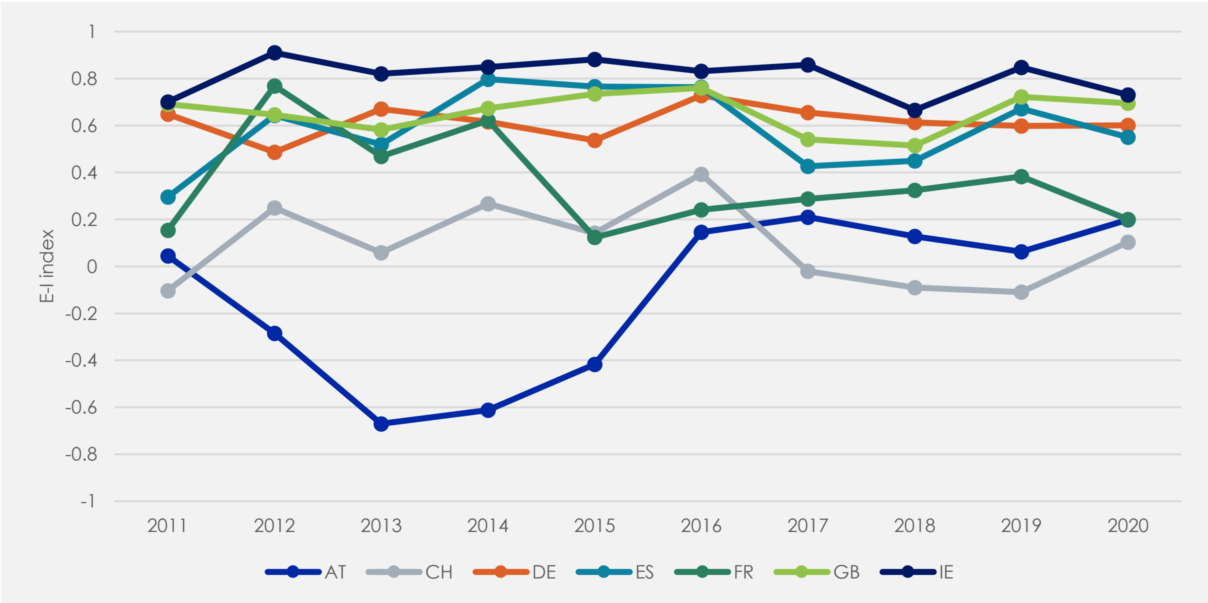

Fig. 34: Annual transnationality (E-I score) per country for glacier tweets.

Permafrost

A similar trend separating alpine countries from non-alpine countries can be observed in the permafrost tweets (Appendix, Fig. 55), however does the general data scarcity regarding the subject lead to a much less smooth distribution with harder to decipher trends. The spatial distribution of transnational connections is unsurprising as a large majority leads to Russia, Canada, the United States, and Greenland, which contain a large majority of the earth's permafrost.

Sea level

The individual E-I scores for sea level tweets offer a clearer picture with some expected trends (Fig. 35). Landlocked Switzerland and Austria both maintain strongly transnational values throughout the study period, while Germany and France start out similarly but then develop towards an almost even E-I score (near zero) throughout the second half of the study period. Although there is certain variation, France, Ireland, and especially Great Britain have the strongest domestic tendency as they mostly hover between a score of 0 and 0.3.

Regarding the spatial distribution of transnational relationships, there appears to be a strong cause-and-effect trend. On the one hand, Antarctica and especially Greenland score very highly as their contribution to global sea level rise is frequently discussed. On the other hand, especially the Western European countries are frequently mentioned with regards to how they will be impacted. As one of the low-lying countries especially vulnerable to sea level rise (Van Koningsveld et al., 2008: 367), it is hardly surprising that especially the Netherlands are often the main subject. France, Germany, Italy, and the United Kingdom also belong to this list of frequently mentioned areas. Besides the usual frequent mentions of the United States, China, and India, there is some increased interest in Bangladesh on the intercontinental level, which is another low-lying country particularly threatened by sea level rise (Huq et al., 1995: 44).

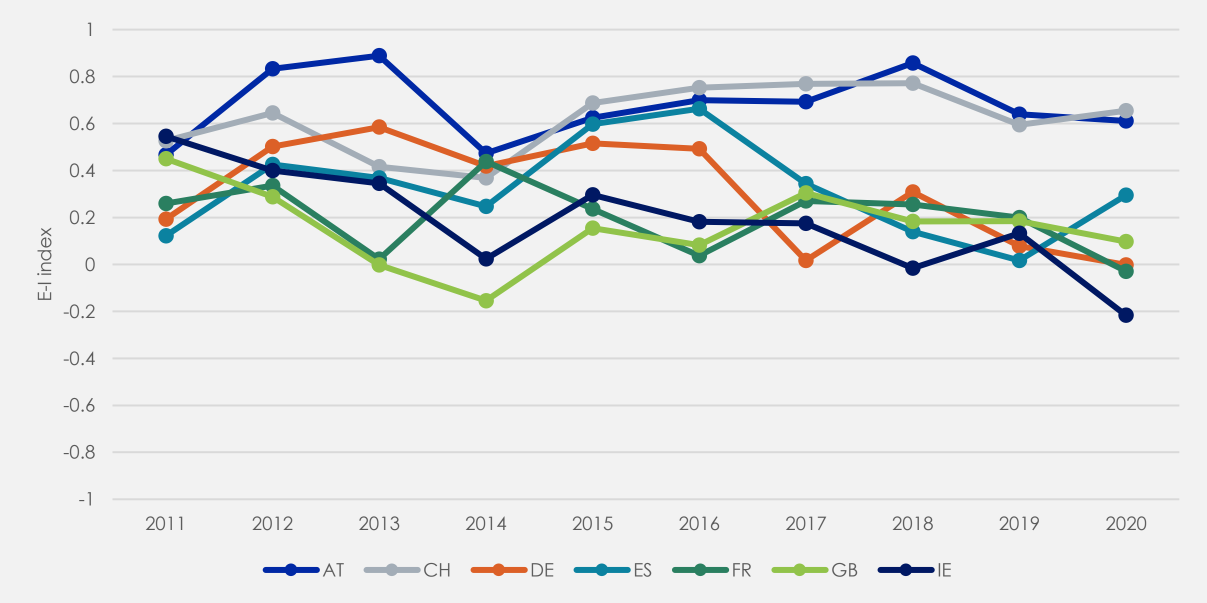

Fig. 35: Annual transnationality (E-I score) per country for sea level tweets.

4.2.4 Environmental context: local, national, and global events

Some of the sections above have already touched on the fact that despite tweets belonging to the online sphere, strong connections can sometimes be drawn between content published by Twitter users and simultaneous occurrences in their real-world environment. Scientific observation of such environmental events collected in databases can be used to validate events detected from social media (de Bruijn et al., 2019: 2) or compare trends in climate change perception to trends in the environment (Hamilton & Keim, 2009: 2350).

Tobler's First Law of geography famously states that "*everything is related to everything else, but near things are more related than distant things*" (Tobler, 1970: 236). Several studies found that this law also applies to online spheres and social media to a certain extent, with the correlation decreasing after a certain distance threshold and eventually even reversing (Li et al., 2014: 515; Ostermann et al., 2015: 316). Regarding the climate change effects in this study, one might once again expect to see a divide between the short-term and long-term processes with the general public more likely to vividly experience heatwaves and wildfires than glacier shrinkage, permafrost melt and sea level rise in their local area. But the question remains whether this truly is the case, or if the reporting on and concern for global events can overshadow the interest in local processes.

Heatwaves

The direct connection between heatwaves and air temperature simplifies the investigation into the impact of real-world parameters. The influence of temperature on heatwave tweet activity is unquestioned as all events were detected in the warmer months between May and September, with a significantly higher number of overall tweets in the hot summer months of June, July, and August (see also Fig. 41c). Even when just investigating tweets *during* these heatwave events, however, temperature still retains its strong influence on tweet activity as most regions of the study area exhibit moderate to strong relationships between the two variables. The following will demonstrate these connections on local, national, and global levels.

Overall, the correlations between tweet frequencies and temperatures (Tab. 17) show a clear divide separating global temperatures from their local and national counterparts. This suggests that the magnitude of tweet activity during a heatwave event on Twitter is hardly impacted by maximum temperatures around the globe. Instead, it is local and national temperatures that really drive the intensity of reactions. This result fulfils the expectation that the local environment has a high importance in eliciting reactions to extreme temperatures, whereas heatwaves on the global scale did not seem to cause reactions strong enough to be detected alongside domestic events.

Tab. 17: Correlations (Pearson's r) between tweet frequency and daily maximum temperature on the local, national, and global scale during detected heatwave events. Local correlations represent national means of all areas. Local and national correlations are all statistically significant (p-value < 0.05). Global maximum temperature is calculated as the mean of the top 1% hottest grid cells.

Country	Local max temp	National max temp	Global max temp
AT	0.46	0.58	0.22 (p = 0.01)
CH	0.57	0.60	0.06 (p = 0.48)
DE	0.60	0.75	0.22 (p = 0.01)
ES	0.54	0.20	-0.13 (p = 0.09)
FR	0.63	0.38	-0.01 (p = 0.96)
GB	0.50	0.50	0.15 (p = 0.11)
IE	0.57	0.54	-0.21 (p = 0.07)

The differences between countries on the local level are not huge and most areas fall into a range of moderate to somewhat strong correlation. On the national scale, there is a little more variation. Spain is the only country that exhibits a very low correlation, which might be due to the common occurrence of heatwaves on the Iberian Peninsula, leading users only to care about local events as heatwaves on the national level are no rarity. The opposite might be the case for Germany, where extreme temperatures anywhere in the country elicit reactions from users anywhere in the nation.

Not only the direct impact of temperatures on public reactions, but also the general spatial distribution of tweet frequencies across the study area (Fig. 25) seems to follow the trends in extreme temperatures quite well. When compared to Fig. 36, the general trend of high tweet frequencies in much of Spain and parts of France is reflected in observed summer temperature anomalies for these regions, whereas the lower temperatures in the remainder of the study area agree in broad terms with lower tweet activity (Kjellström et al., 2007: 253–255).

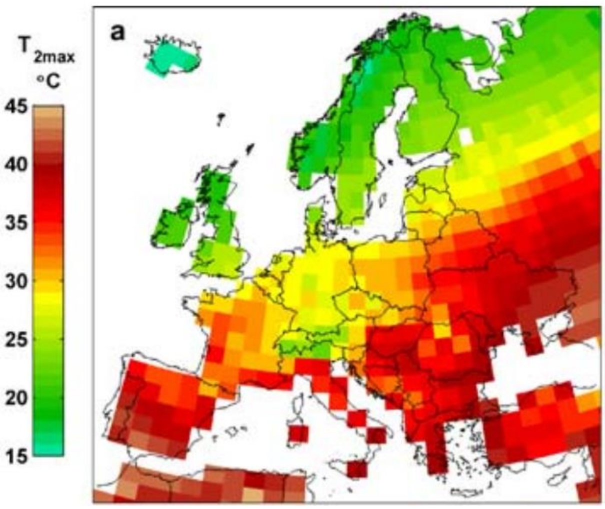

Fig. 36: 95th percentile of daily maximum summer temperatures (Kjellström et al., 2007: 254).

When it comes to the actual temperatures at which the different populations tweet about heatwaves, it is interesting to see that these thresholds (Fig. 37a) align very well with the 95th percentile of summer temperatures in each country (Fig. 36).The only exception is possibly Austria, where the higher temperatures in the country's east are contrasted with an overall lower threshold. This is highly relevant as the 90th to 95th percentile is frequently used as a threshold for real-world heatwaves (Nairn & Fawcett, 2014: 228; Pezza et al., 2012: 211; Stefanon et al., 2012: 2) – meaning Fig. 36 roughly displays heatwave temperature thresholds for Europe[15]. Therefore, temperature thresholds derived from tweeting behaviour (Fig. 37a) align remarkably well with the scientific threshold for heatwaves.

[15] Fig. 36 takes into account the temperatures of all summer temperatures, while the mentioned authors (Nairn & Fawcett, 2014: 228; Pezza et al., 2012: 211; Stefanon et al., 2012: 2) typically investigate temperature on a monthly basis.

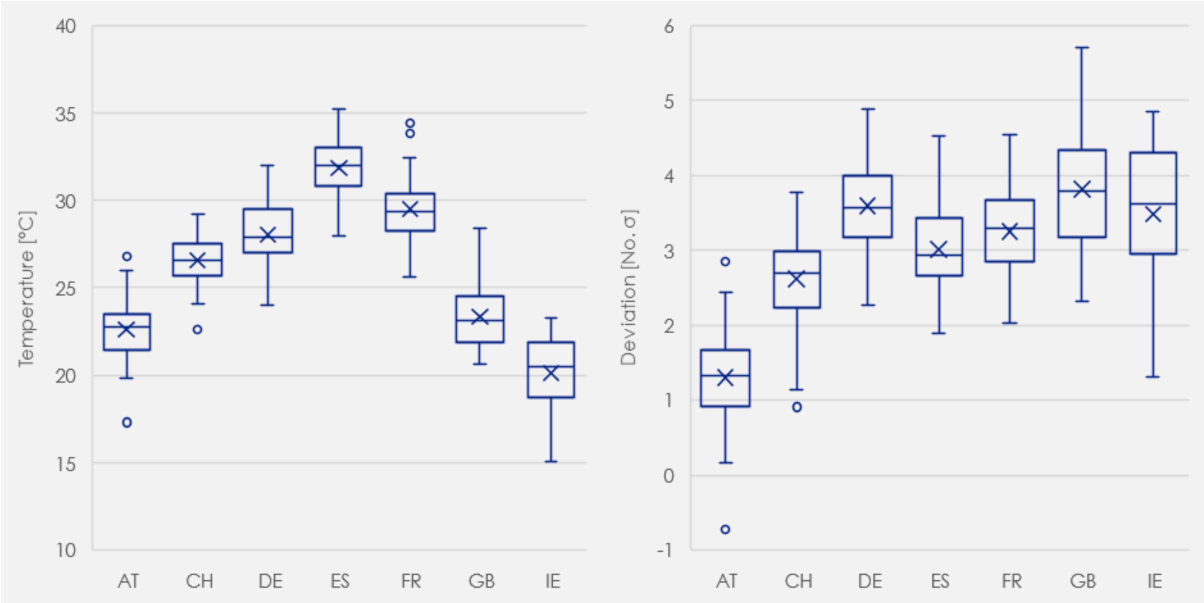

Fig. 37: Distribution of days within heatwave events derived from tweets with the corresponding maximum daily temperature recorded in the country (left) and with the deviation from the 1950-1980 historic mean (right; measured in number of standard deviations).

When comparing the deviation of daily temperatures within heatwave events from the daily historical mean (1950-1980) in Fig. 37b, differences between the countries can be observed as the data is normalised. Interestingly, it is the alpine countries of Austria and Switzerland that somewhat stand out as users react to smaller temperature anomalies than the populations of the other countries in the study area, which all exhibit rather similar values. It seems that a temperature anomaly of 3-4 standard deviations above the historical mean is generally needed to elicit a strong reaction, with the values ranging from 1-2 for Austria and 2-3 for Switzerland.

In conclusion, it can be stated that temperatures and heatwave tweet frequencies are strongly tied to each other. Not only does higher overall frequency align quite well with the spatial distribution of temperature anomalies in Europe, but there is also a moderate to strong correlation between the magnitude of a public reaction and the corresponding temperatures on the local and national scale at that point in time. Furthermore do thresholds in heatwave events derived from tweets align very well with temperature thresholds for real-world heatwaves.

Wildfires

Observing the correlation between daily wildfire tweets and burnt area, varying patterns emerge between different spatial granularities and ways of measuring wildfire extent (Tab. 18). Overall, the number of daily tweets during an event (normalised for twitter users) correlates the most with the total area burnt globally during the event (Pearson's $r = 0.51$, $p < 0.001$), however is this connection only moderate. The correlation on the national level only lags slightly behind, whereas the distance to the local (subnational) level is more considerable. Interestingly, on the local and national scale, the correlations are somewhat similar across the three different statistics with a peak in the maximum burnt area, while there seems to be no correlation at all for average and maximum burnt area on the global scale.

This suggests that wildfires on a global scale only become consistently important once they maintain a large burning/burnt area over a longer duration resulting in a large-scale event. On a local or national scale however, the overall extent of wildfires does not seem to influence tweeting behaviour quite as much as users tend to react the most when a wildfire exhibits an unusually large single-day extent at one

This should lead to an overestimation of thresholds in Fig. 36 as there is a bigger pool of temperatures. With the age of the Kjellström et al. (2007) study, this effect may have, however, already been negated due to increasing summer temperatures.

point during the event. Thus, *almost any* local or national wildfire will elicit a reaction, whereas *only large-scale* wildfires in the global context cause increased discussion.

Tab. 18: Correlation (Pearson's r) between burnt area and daily tweets per 1000 users at different spatial scales. Each identified wildfire event in the tweet corpus represents a datapoint. Wildfire data only spans from January 2011 until June 2018[16].

Granularity	Average burnt area (per day)	Maximum burnt area (single day)	Total burnt area (during event)
Local (subnational)	0.35 (p < 0.001)	0.42 (p < 0.001)	0.37 (p < 0.001)
National	0.48 (p < 0.001)	0.49 (p < 0.001)	0.45 (p < 0.001)
Global	-0.04 (p = 0.71)	-0.07 (p = 0.5)	0.51 (p < 0.001)

When comparing the tweet activity in wildfire events and the area burnt during each event in a country (Fig. 38), it becomes rather clear that Spanish users tend to tweet when there are active wildfires in the country, whereas other populations frequently tweet about wildfires in their absence on the national level. This is one hand due to the fact that Spain, in general, experiences larger and more frequent fires than the rest of the study area (Eriksen, 2020: 2), but might also point towards a more domestic focus regarding wildfires amongst Spanish users, which has already been shown in terms of transnationality (section 4.2.3).

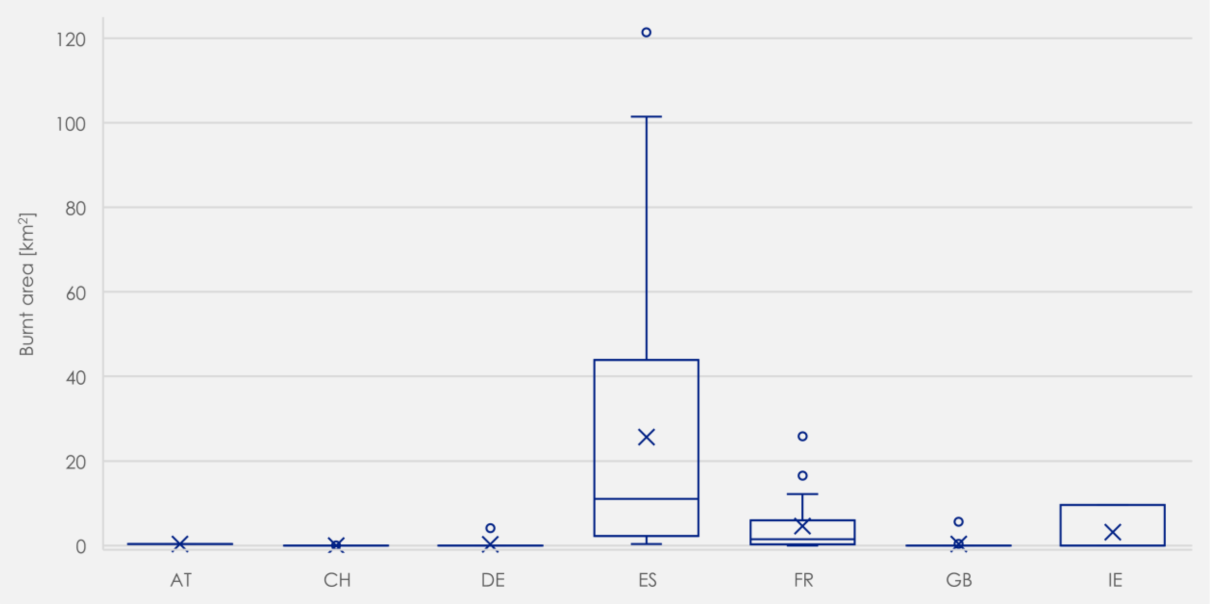

Fig. 38: Distribution of wildfire events derived from tweets with corresponding total area burnt in each country during the event.

Interestingly, this all results in a pattern contradicting Tobler's First Law as correlation seems to increase with distance. Indeed, it seems as though the reversing effect found by Li et al. (2014: 515) is evident in European wildfire tweets as well. The main reason for this is likely related to the fact that a lot of the recent large-scale wildfires occurred outside of the study area in areas such as Australia, the Amazon and Western North America (Xu, Yu, et al., 2020: 2173). This tendency for tweets to gravitate towards extensive international wildfires is further confirmed when looking into the most important keywords for each event, as the locations mentioned mostly reference regions in different countries except when a significant event is occurring in the own country.

[16] Global correlations for the entire study period are expected to be stronger due to extraordinarily large wildfires after 2018 coinciding with increasing tweet frequencies.

Glaciers

For LTETs such as glacier shrinkage, it is more difficult to contextualise tweet frequency with real-world events as it is a process spanning decades and centuries (Zemp et al., 2015: 749). In fact, the 'events' related to glacier melt that spark conversation on Twitter can mostly be seen as emblematic incidents representing glacier shrinkage (e.g., calving events, ice shelf collapses), which however have little impact on the development of the process as a whole.

Nevertheless, measures determining the development of the retreat (and in rare cases advance) of glaciers such as the change in annual mass balance can be used as an indicator of the severity of the shrinkage in any given year (Zemp et al., 2013: 1227). When plotted against the annual tweet frequency of every region in the study area, a slight negative correlation between global mass balance and tweet frequency can be observed (Fig. 39, Pearson's r = -0.29, p = 0.016), however is the signal rather weak. This relationship does meet the expectation though, that years with increased glacier retreat (mass balance loss) tend to elicit more tweets. A direct influence, however, is questionable as mass balance developments can be summarised and published over a year after the measurements (e.g., WGMS (2021b) reporting on the 2018/19 season) and thus a factor such as temperature causing both an increase in ice melt as well as glacier tweets could be a more likely cause.

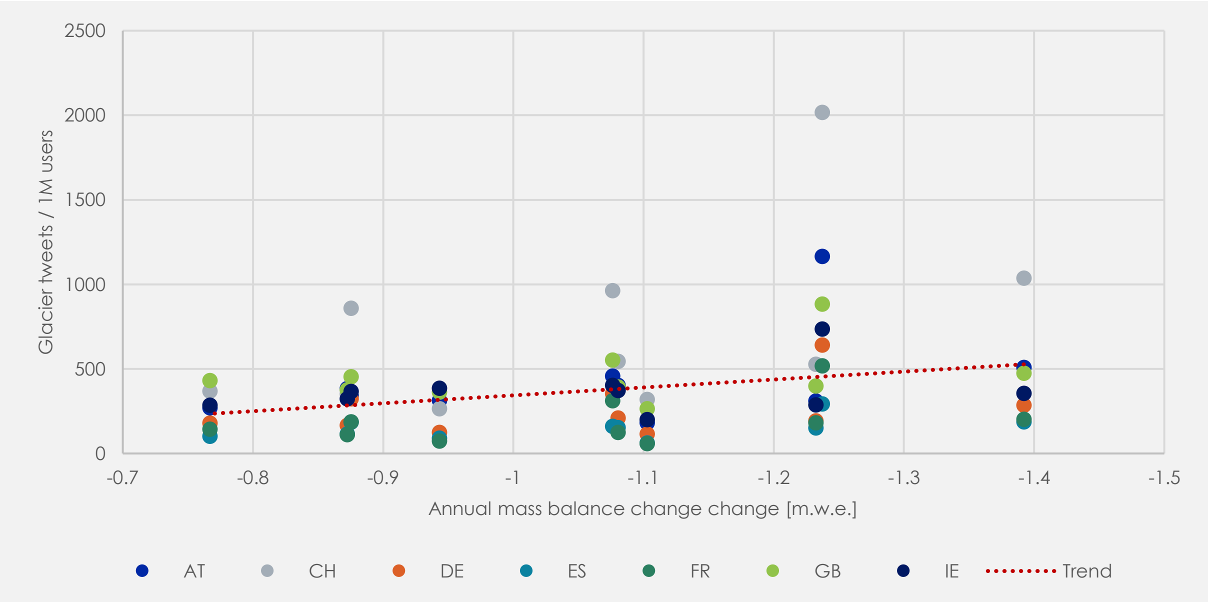

Fig. 39: Correlation between mean global annual mass balance change and annual glacier tweet frequency in every country (Pearson's r = -0.29, p = 0.016).

It is worth noting that no correlation can be found for glacier tweet frequency and European (Pearson's r = 0.01, p = 0.962) or Central European annual mass balance (Pearson's r = 0.07, p = 0.576). There thus seems to be no indication that the mass balance of glaciers closer to the study area has a stronger influence than the global mass balance.

Permafrost

Similar to glacier retreat, permafrost melt can be quantified as the change in annual extent, resulting in a corresponding measure. When comparing global annual permafrost loss to annual tweeting activity, no significant correlation can be found, as the p-value in any Pearson's r calculation far exceeds the 0.05 threshold for significance. This is the case for both correlations within the individual countries as well as the correlation across the entire study area. These observations lead to the conclusion that there is no direct correlation between annual losses in global permafrost extent and public reaction as measured by tweet frequency. Therefore, a set of drivers independent of permafrost extent are entirely responsible for fluctuations in public reaction.

These results do not necessarily come as a surprise as the concept of permafrost is first of all likely not very tangible to the general public (especially since it is not present in large swaths of the study area) and secondly is the slow-paced and long-term nature of permafrost melt not expected to elicit any large-scale reactions without further external influences.

Sea level

Sea level rise can vary strongly on the local scale and is thus best measured using a global average from satellite altimeter data (Church & White, 2011: 586). Thus – even though data is available – using local sea level measurements as context data to juxtaposition against tweet frequency does not serve as an honest investigation into the connection between sea level rise and the magnitude of users' contributions to the discussion on Twitter. As a result, only large-scale datasets are used to check for a correlation. Sea level rise can be measured in relative change (e.g., in monthly or annual intervals) or as an absolute change relative to a set baseline from the past.

Much like it was the case with permafrost extent, relative changes in sea level do not result in any significant correlations. Alongside a global mean, I used study area-relevant measurements of the Atlantic, Mediterranean, North, and Baltic Sea levels as references to compare national as well as study area-wide tweet activity against, both on an annual as well as monthly scale. None of these country-sea or study area-sea combinations exhibited a significant correlation as indicated by the p-values. These results apply to datasets with both the seasonal signal of sea level rise removed as well as with the signal retained.

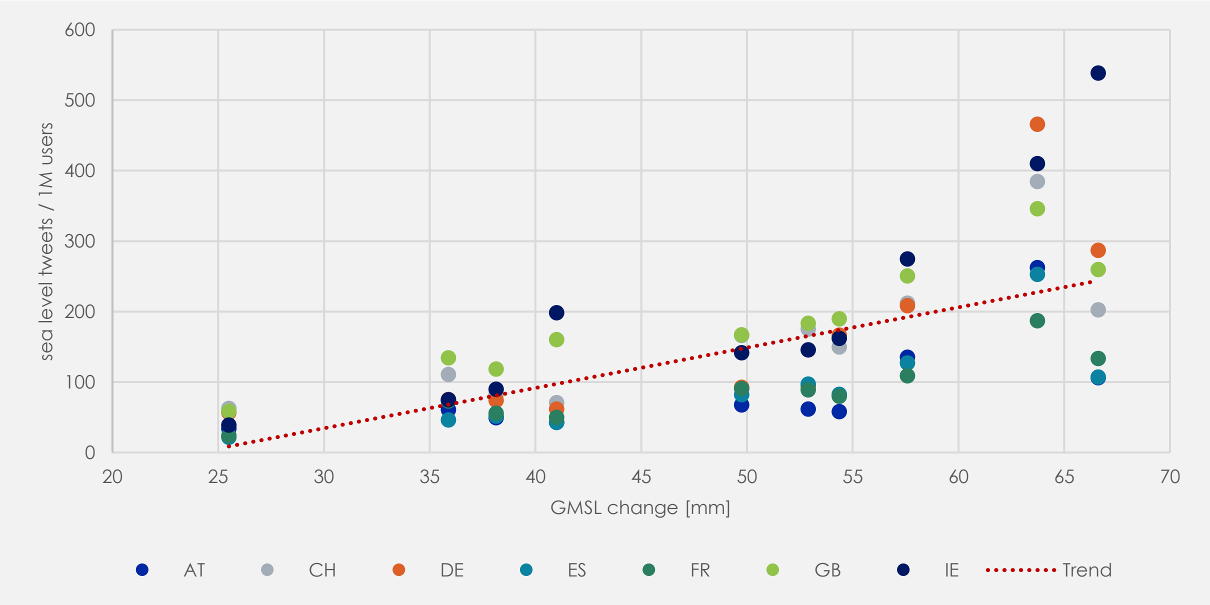

Fig. 40: Correlation between absolute global mean sea level (GMSL) change (compared to a historic baseline) and annual sea level tweet frequency in every country (Pearson's r = 0.68, p-value < 0.001).

A further investigation with coastal and landlocked subnational areas separated was carried out to see whether tweeting activity in regions near the ocean potentially correlates more with annual or monthly sea level change. This, however, lead to similar results with no indication of significant correlations.

When it comes to absolute changes in sea level compared to a historic baseline, however, the picture is entirely different. Not only are the correlations significant, but a strong signal can also be observed in the global case (Fig. 40, Pearson's $r = 0.68$, p-value < 0.001). The correlation is especially strong in the United Kingdom (Pearson's $r = 0.92$, p-value < 0.001), followed by France (Pearson's $r = 0.87$, p-value

< 0.001) and Ireland (Pearson's r = 0.85, p-value = 0.002)[17]. Interestingly, it is exactly these countries that contain some of Europe's most threatened coastlines (Kirezci et al., 2020: 4), which not only suggests that the increased tweet activity in general is partly due to the real-life development of the global sea level, but also that inhabitants of higher-risk areas potentially react more strongly to increasing sea levels.

4.3 Temporal dimension

With the temporal analysis, the goal is to determine *when* a reaction from the public is elicited. Thus, the frequency distribution of tweets over the study period becomes the crucial metric, both in terms of the temporal distribution and the magnitude of the reaction. After introducing some general temporal statistics on tweets and users, this chapter will closely investigate temporal characteristics of tweet activity in two ways: First, timelines of important events are constructed from peaks in the frequency distribution marking large public reactions. In a second step, the approach is reversed and events relevant to climate change are investigated to determine if their occurrence matches any peaks in tweet activity. For these steps, the focus solely lies on the LTETs glacier, permafrost and sea level, as literature and previous sections (4.2.3 – 4.2.4) in this thesis have already established a good understanding on the temporal influences on the STETs heatwave and wildfire (Wang et al., 2016: 529). Finally, the chapter concludes with a discussion of the COVID-19 pandemic and its impact on tweet activity.

4.3.1 Periodicity and seasonality

Investigating if there is any pattern to tweet frequencies can give an insight into *when* people engage in conversations surrounding climate change. Are events more discussed at work or during free time? Do the seasons have an impact on people's climate change perception? Previous research specifically on climate change tweets found that there are strong daily and weekly cycles with users tweeting the most while at work. Specifically the morning working hours and the first four days of the working week saw the most activity (Kirilenko & Stepchenkova, 2014: 180).

Daily cycles

In general, the trends in the literature mentioned above are reflected in tweet corpus of my thesis (Fig. 41). After low tweet activity during the night, the distribution quickly reaches its peak during the mid-morning working hours before flattening off during the remainder of the work day with a slow decline towards the evening.

While all other event types exhibit very similar trends, there are two facets differentiating heatwave tweets from the rest: Firstly, high tweet activity begins earlier in the morning. There is a number of potential explanations for this, ranging from the earlier sunrise time during the summer season and the resulting earlier rise in temperatures to people potentially experiencing sleeping problems due to the heat and communicating it in the morning. A second difference in the heatwave tweet distribution occurs in the afternoon when the absolute peak is reached while the tweet distributions of other event types are already declining again. One would assume that this peak roughly matches the time of day around which the maximum temperature is reached.

Overall, these observations seem to indicate once again that users tweeting about heatwaves are very strongly connected to the environment and seemingly discuss heatwaves in the form of immediate reactions, whereas tweeting behaviour surrounding the other event types seems to be conducted in a more distanced and calculated manner. It has also been suggested that a lot of tweeting about climate change is possibly done in conjunction with a user's work duties (e.g., tweeting from accounts of

[17] It is important to stress that tweet activity is quantified as tweets per 1 million users in the given region (with the number of users being adjusted on an annual basis), meaning that this correlation between cumulative SLR and tweet frequency is not simply a result of an increase in users over time.

newspapers or organisations; Kirilenko & Stepchenkova, 2014: 180), which would be one explanation for the observed cycles in the events other than heatwaves.

Weekly cycles

Similarly to the hourly distribution, trends on the weekly scale show strong agreement with those found in literature, supporting the findings that tweeting activity surrounding climate change – whether that be with direct climate change tweets as in Kirilenko & Stepchenkova (2014) or proxy tweets such as the ones in my thesis – generally is strongest between Monday and Thursday with a first drop-off to Friday and another, even stronger one to the weekend. In this case, there are no decisive outliers as all event types follow a rather similar trend.

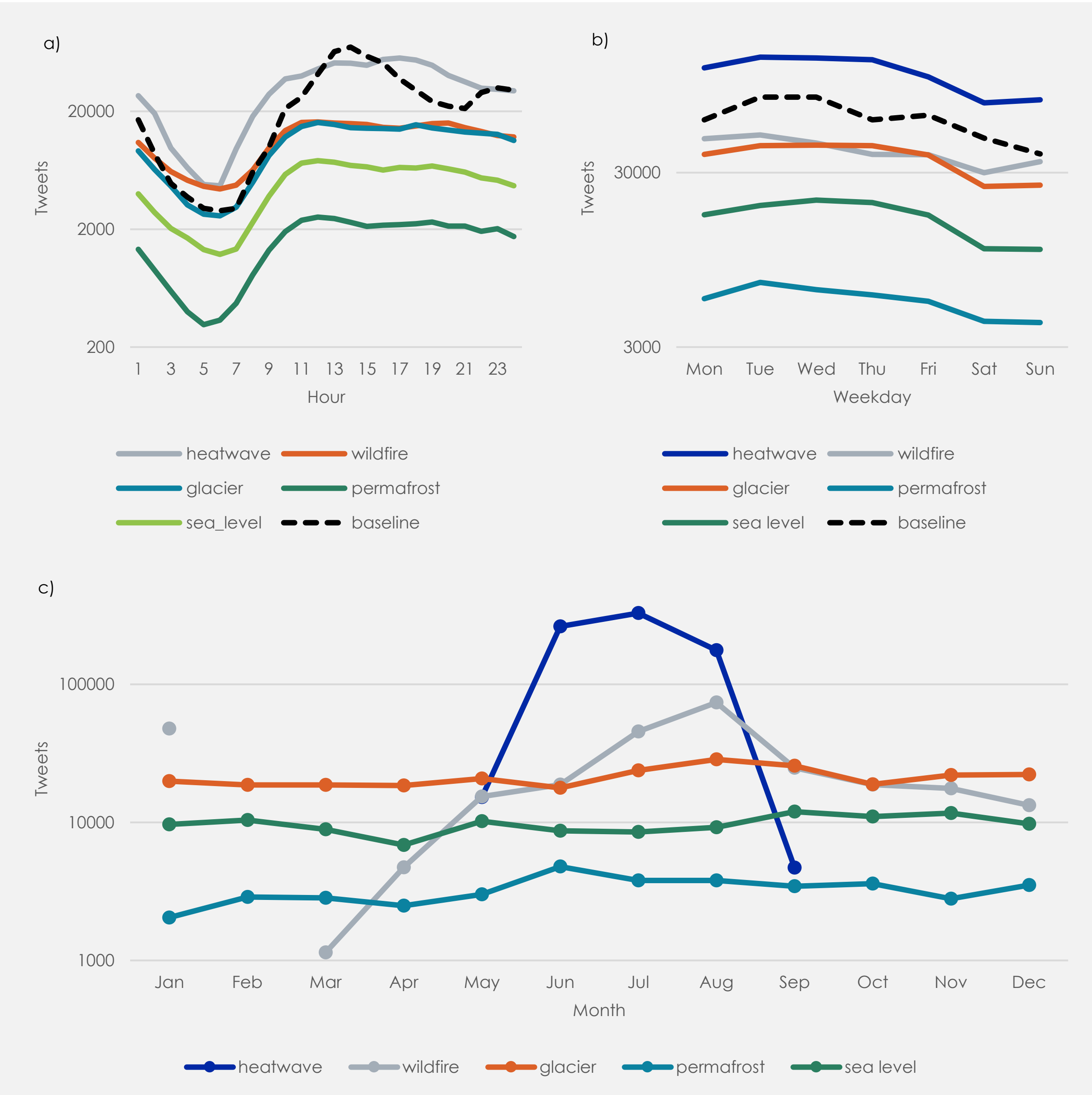

Fig. 41: (a) hourly, (b) weekday, and (c) monthly distributions of all event tweets during the study period. Y-scale is logarithmic. Baselines (Kirilenko & Stepchenkova, 2014: 174) are on an arbitrary secondary axis and are only comparable in their general shape.

Seasonal cycles

When it comes to seasonal signals, the trends are as expected. Heatwave and wildfire tweets exhibt strong seasonal variability, while the LTETs seem to be somewhat randomly distributed across the year. With heatwaves typically being defined by a number of days exceeding a high temperature threshold

(Nairn & Fawcett, 2014: 228; Pezza et al., 2012: 211; Stefanon et al., 2012: 2), it is not surprising to see the only peak during the European summer. Similarly, wildfires are strongly tied to extreme heat (Sun et al., 2019: 129) and thus tweet activity surrounding them increases during summer seasons. Compared to heatwave tweets, however, the decidedly more global discourse surrounding wildfires becomes appareant as the southern hemisphere summer exhibits a similar peak to the northern hemisphere summer – despite none of the studied areas being south of the equator.

Interestingly, this international interest seems to be controlled by language to a certain degree, as Spanish users are by far the most active during the European summer, whereas British users make up the majority of tweets from November to January, the wildfire season of the southern hemisphere and especially english-speaking Australia. French users are more northern hemisphere-centric when it comes to wildfire tweets, while German and Swiss users are spread more evently between the seasons.

Despite their overall rather trend-less distributions, there are some notable aspects about the LTETs. For glacier tweets, the countries Switzerland, Germany, France and United Kingdom show a small, but noticable peak during the late summer months of August and September, coinciding with the end of the ablation period – and thus the most negative mass balance during the year – for glaciers in the northern hemisphere (Huss et al., 2009: 202). This height of the melt season also results in increased news reporting on the issue, which is picked up on Twitter.

Permafrost tweets also exhibit a singular peak in the course of the year in June, when the activity in most countries of the study area is at its heighest. The reason behind this is likely to be the wide-spread interest shown in news articles surrounding the Farquharson et al. (2019) study, as discussed in section 4.1.4. Interest in sea levels seems to be slightly higher during fall than in any other season. One explanation could be the annual United Nations COPs held near the end of the year, which have shown to influence the discourse on sea level rise in some years (see chapter 4.3.4). Events such as the Global Climate Strike could additionally be responsible, as the 2019 event contibuted over 1000 sea level tweets across the study area in a single day at its peak – roughly a tenth of the cumulative September count for the entire study period (section 4.3.6).

4.3.2 User population over time

While the remaining temporal analyses will largely focus on how populations react to different events over time, a short overview of the temporal characteristics of the populations themselves is useful. The turnover of users from event to event (in the case of STETs) or year to year (LTETs) can be examined to detect how much the population of discourse participants changes over time (Williams et al., 2015: 128). The Sørensen similarity (Sørensen, 1957) quantifies this overlap ranging from 0 (no overlap) to 1 (identical populations).

Tab. 19: Mean event-to-event Sørensen similarity of unique users for each country and event type. Country-specific values are only listed for the entire population, while means display the trend over the entire population as well as the top 100 and top 10 most prolific users in each event.

Country	Heatwave	Wildfire	Glacier	Permafrost	Sea level
AT	0.12	0.07	0.22	0.10	0.16
CH	0.16	0.14	0.21	0.15	0.21
DE	0.17	0.14	0.21	0.17	0.20
ES	0.05	0.16	0.18	0.14	0.16
FR	0.17	0.11	0.17	0.14	0.15
GB	0.04	0.12	0.17	0.14	0.20
IE	0.03	0.10	0.16	0.11	0.19
Mean (all)	*0.11*	*0.12*	*0.19*	*0.14*	*0.18*
Mean (top 100)	*0.16*	*0.18*	*0.35*	*0.16*	*0.27*
Mean (top 10)	*0.19*	*0.25*	*0.40*	*0.22*	*0.32*

Generally, the population turnover is rather high, with large numbers of users not carrying over from event to event or year to year. The slightly higher values for LTETs are expected due to the much longer time intervals. Considering Williams et al. (2015: 128) found a 0.19 similarity for climate change tweets in 10-day intervals, the values in Tab. 19 seem reasonable. Compared to their results, however, the index values do not increase nearly as much when selecting the top 100 (19-86% increase) or top 10 (61-114% increase) most prolific users for each event, where the authors found a 189% and 268% increase, respectively. This suggests that the datasets contain a good variety in discourse participants across the board, even amongst the most prolific users, and the divide in turnover rates between the core population (high-activity users) and the peripheral population (low-activity users) is decidedly smaller than found by Williams and colleagues. The most dominant voices steering the respective discourses are therefore expected to vary across the span of the study period.

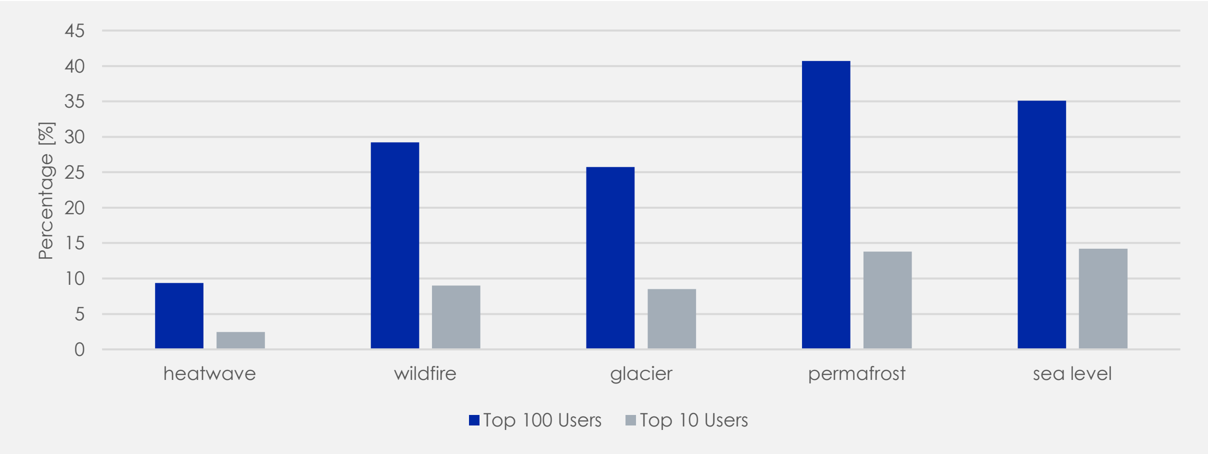

Fig. 42: Mean percentage of tweets made by the top 100 and top 10 most prolific users per event relative to tweets made by all users during the event.

The discourse is further generally diverse with regards to prolific users, as the top 100 users per event only contribute 9-40% and the top 10 users 2-14% of all event tweets (Fig. 42). Therefore, the dominant voices that do remain from event to event (as seen above), do *not* make up a majority of the tweet activity. This is likely due to the filtering of excessive duplicate tweets by individual users during the retrieval process (section 3.3.3).

4.3.3 Event timelines: individual peaks in tweet frequency

As stated before, peaks in heatwave and wildfire tweet frequency align well with the occurrence of corresponding events in real-life. For the LTETs glacier, permafrost, and sea level, however, other factors mostly determine when the process becomes a hot topic on Twitter. To learn what these impacts are and to understand which sources and stories regarding the climate change effects stir up a conversation in the online sphere, the following sections will investigate individual peaks in the corresponding tweet frequencies in more detail. This should give a better idea of what causes the various temporal variations in tweet frequency across the study period.

The timelines are compiled in Tab. 20, Tab. 21, and Tab. 22. The following sections will summarise the findings and explore general trends and special cases for each of the LTETs.

Glaciers

Generally, spikes in tweet activity surrounding glaciers tend to result from news articles summarising recently published scientific papers or ongoing research on global developments regarding glaciers and ice sheets (GSN). The theme of these studies is largely related to processes in the polar regions and their potential to cause significant sea level rise due to melt and collapse. Unsurprisingly, article headlines often focus on large numbers and use ‘alarmist’ terminology to various degrees, a trend that has been increasingly perceived by readers in climate change reporting (Borick & Rabe, 2010: 791; Lorenzoni et al., 2007: 452; Whitmarsh, 2011: 697).

Tab. 20: Timeline of most notable (high anomaly) peaks in glacier tweet frequency for individual years and countries. Daily anomalies are calculated as number of standard deviations from annual tweet frequency mean (per country). Only days with a discernible cause for the high tweet frequency were included. *Ano = Anomaly, C = Country.

Year	Mth	C*	Ano*	Type	Cause and associated study (if available)
2011	Jan	FR	+9.0	NSN	Excerpts from a sea level contribution study stating that three quarters of glaciers in the Alps could disappear by 2100 (Radić & Hock, 2011)
	Dec	DE	+7.2	GN	Images showing dramatic melt in Chile
2012	Jul	CH	+10.3	GSN	Unprecedented melt of Greenland ice sheet (Viñas, 2012)
		DE	+6.3		
		ES	+14.1		
		FR	+13.1		
		GB	+13.0		
		IE	+11.5		
2013	Jan	GB	+8.6	GSN	Study detailing tremendous melt of Andean glaciers (Rabatel et al., 2013)
	Apr	CH	+9.7	NSN	Glacier retreat creates glacier lakes in Swiss Alps (Bojanowski, 2013)
	Nov	ES	+12.6	NSN	SLR will swallow Barcelona in 5000 years due to glacier melt (Pazos, 2013)
2014	May	AT	+9.2	GSN	Potential ice shelf collapse in West Antarctica due to possible irreversible ice loss (Joughin et al., 2014; Rignot et al., 2014)
		CH	+10.9		
		DE	+14.9		
		ES	+16.6		
		FR	+16.1		
		GB	+13.4		
		IE	+9.1		
2015	Aug	CH	+8.4	GSN	Switzerland-based WGMS publishes study on rapid glacier melt in 21st century (Zemp et al., 2015)
		DE	+7.0		
	Nov	ES	+8.2	GSN	Once stable Greenland glacier retreats rapidly, could contribute 0.5m to SLR (Mouginot et al., 2015)
		FR	+7.6		
		GB	+8.6		
2016	Oct	ES	+7.0	GSN	Unprecedented retreat of Antarctic ice shelves due to submarine melting (Khazendar et al., 2016)
	Nov	GB	+11.6	GSN	Report detailing potential of Arctic ice melt leading to uncontrollable climate change (Carson & Peterson, 2016)
		IE	+9.6		
2017	Jul	DE	+8.2	GSN	Giant iceberg calves from Larsen C ice sheet
		FR	+6.2		
		GB	+13.7		
		IE	+10.1		
	Aug	CH	+10.4	NSN	Comprehensive interactive newspaper report on extent of glacier melt in Switzerland (Lutz & Brupbacher, 2017)
2018	Jun	DE	+8.3	GSN	Major study published detailing Antarctic ice loss and resulting sea level rise (Shepherd et al., 2018)
		FR	+7.5		
		GB	+7.9		
		IE	+8.2		
	Oct	CH	+8.6	NSN	Press release from Swiss Academy of Sciences reporting on accelerated glacier retreat during heat summer of 2018 (SCNAT, 2018)
		ES	+16.7	NSN	Spain's last glacier goes extinct (Monte Perdido)
2019	Feb	FR	+11.7	GSN	Potentially disastrous impact of Thwaites Glacier melt (Milillo et al., 2019)
	Jun	ES	+8.0	GSN	Study detailing acceleration of ice loss in Himalayas (Maurer et al., 2019)
	Aug	DE	+8.4	GN	Message on memorial plate for melted Iceland glacier goes viral
	Sep	CH	+7.8	NN	Funeral for Pizol Glacier, Switzerland is held
	Nov	AT	+6.3	NN	Construction equipment sighted digging on Austrian glaciers in context of the merger of the Pitztal-Ötztal ski areas
		GB	+8.2	NN	British Channel 4 replaces prime minister Boris Johnson with melting ice sculpture as he avoids climate change debate
2020	Aug	DE	+7.5	GSN	Imbalance of Greenland ice sheet leading to increased ice loss (King et al., 2020)
		FR	+6.8		
	Sep	GB	+6.2	NN	Petition for Bank of England to "stop funding climate chaos" (Anonymus, 2020)

GN = Global News
GSN = Global Science News
NN = National News
NSN = National Science News

Amongst the most notable events on the timeline is a plethora of studies detailing (1) long-term developments surrounding the Greenlandic and Antarctic ice sheets (Carson & Peterson, 2016; Joughin et al., 2014; Khazendar et al., 2016; King et al., 2020; Milillo et al., 2019; Mouginot et al., 2015; Rignot et al., 2014; Shepherd et al., 2018; Viñas, 2012) and (2) glacier retreat in other regions (Maurer et al., 2019; Rabatel et al., 2013; Zemp et al., 2015). News reporting on these studies dominate large parts of the timeline and are responsible for some of the strongest reactions. Especially in 2012 and 2014, as well as 2018 to a certain degree, single studies caused the highest peak in tweet activity across all countries of the study area despite potential language barriers and diverging national interests.

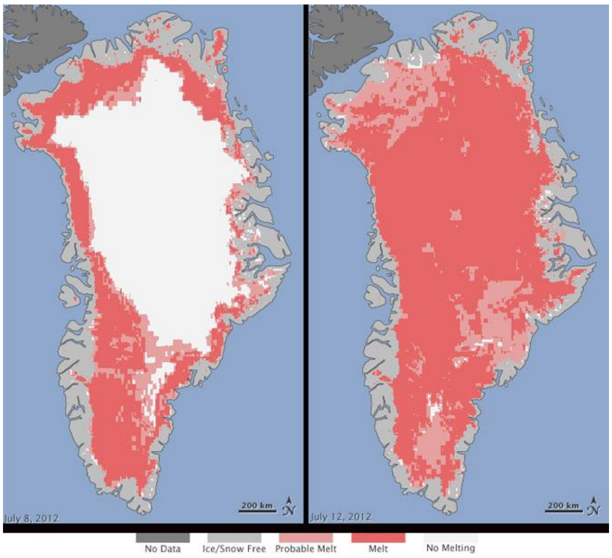

Fig. 43: Surface melt on the Greenlandic ice sheet on 8.7. and 12.7.2012 (Viñas, 2012).

The 2012 peak surrounding a report on unprecedented ice melt in Greenland is largely constrained to a single graphic demonstrating melt occurring on the Greenland ice sheet at the time with the use of two maps (Fig. 43, Viñas, 2012). It is potentially exactly this simple nature and universal language of the imagery that made it relevant in all regions of the study area, further underlining the strong impact visual clues can have on the public in the context of glacier retreat (Borick & Rabe, 2010: 785–786). Imagery is argued to be generally effective regarding climate change, however can it also lead to people feeling overwhelmed and underequipped to adapt to and mitigate these large-scale processes in nature (Metag et al., 2016: 219).

An interesting temporal development from the start towards the end of the study period is an increase in both domestic topics and the memorials or funerals for lost glaciers. Whilst the early years of the study period saw a rather consistent pool of more abstract studies focussing on global glacier retreat and subsequent sea level rise, there has been significantly more variety amongst the most important topics in recent years. Large public reactions were more often caused by political and environmental developments on the national scale and topics became more specific and tangible in certain cases. Nevertheless, some of the global studies of course still remained of high importance on several occasions during this time span.

One special case seems to be Switzerland, whose population tends to focus more on national issues, even early in the study period as well as in years such as 2017 where the reports of a large-scale calving event at Antarctica’s Larsen C ice shelf elicited large responses in many other countries. The reasoning for this phenomenon seems to be twofold: (1) The Swiss Alps contain a considerable amount of land ice, thus domestic glaciers are of higher interest and (2) glacier monitoring is very active in Switzerland as it is home to organizations such as the World Glacier Monitoring Service (WGMS). Some concern about national glaciers was also shown by French users in 2011, when a study detailing the disappearance of a large majority of glaciers in the European Alps (Radić & Hock, 2011) became wide-spread.

Further domestic interest can also be observed in Spain. Specifically, articles reporting on the extinction of Spain’s last glacier in 2018 elicited one of the strongest reactions of the country’s public, which is especially unrivalled in terms of its annual anomaly. Glaciers suffering a similar fate appear several times on the timeline with the funeral of the Pizol Glacier in Switzerland being the most important event in 2019 amongst Swiss users and the memorial for the Icelandic Okjökull ice sheet becoming the hottest topic amongst German users in the same year.

As the United Kingdom does not contain any land ice on its European territories, it might not be surprising that the two occasions on which domestic news elicited the largest public reaction were of political rather than environmental nature. Both peaks in 2019 and 2020 seem to originate from climate action standpoints with the former being a critique on prime minister Boris Johnson's failure to attend a climate change debate and the latter a petition for the Bank of England to halt funding related to activities worsening climate change – including glacier melt (Anonymus, 2020).

A somewhat political issue, however still closely tied to the local environment, sprung up in Austria during 2019 when construction equipment operated on glaciers to implement the merger of the Pitztal-Ötztal ski areas. The resulting reaction marks by far the largest peak in Austrian tweet activity throughout the study period, especially when considering the fact that this topic remained widely discussed for multiple days. Users drew the arguably ironic connection between global climate action efforts and the widely shared imagery of heavy-duty machinery digging into the ice of the Austrian glaciers.

Overall, it can be concluded that news reporting on globally relevant developments of the polar ice sheets as well as other land ice – especially when an unprecedent rate or extent of ice loss is found – most consistently elicited the largest reactions across the study period. In more recent years, this dominance has increasingly been augmented and in some cases even replaced by a larger variety of topics related to glacier retreat, with domestic environmental and political affairs as well as more specific and tangible glacier stories taking over from the arguably more abstract studies surrounding sea level contribution and other geophysical processes.

Occurrences related to the local or national environment can elicit considerably large reactions from the public, as demonstrated for example by the Austrian and Spanish cases, where domestic events caused the highest peak in interest over the study period. In general, news reporting on domestic glaciers shows the largest influence in countries containing land ice. A special case is the United Kingdom, where national politics tangential to glacier retreat elicited strong reactions.

Permafrost

Similar to glaciers, the most obvious trend in the permafrost timeline is the importance of news reporting on studies detailing the different dangers associated with increased permafrost thaw. These reports refer to (1) the release of greenhouse gases (carbon, methane) to the atmosphere and the potential feedback loop with global warming (Hodgkins et al., 2014; Schuur & Abbott, 2011; Vaks et al., 2013), (2) the potential emergence of dangerous viruses from the frozen ground (Legendre et al., 2014; 2015), (3) the release of mercury (Schuster et al., 2018), and (4) the potential economic impacts related to permafrost thaw (Hope & Schaefer, 2016). These aspects of danger are especially prevalent throughout the first half of the study period as well as in 2018, with a slightly more varied set of topics characterising the latter half.

Besides studies focussing on these dangers, Farquharson et al.'s (2019) study elicited a strong reaction across the study area with their findings of Arctic permafrost already thawing at rates not expected before 2090. This was also the only occasion where Austrian users – who exhibited a general lack of tweet activity related to permafrost – produced more than 10 tweets in a single day.

Non-scientific news also resulted in considerable public reactions on separate occasions. In contrast to the glacier timeline, natural disasters and their direct impacts on humans and the environment are already prevalent in the permafrost timeline. These events include the release of anthrax from thawing permafrost killing 1500 reindeer and injuring citizen of Yamal (Guarino, 2016) and a large-scale diesel spill in the Norilsk River (BBC, 2020) caused by infrastructure collapse due to thawing permafrost , both in arctic Russia. Related to infrastructure, the 2017 flooding of the Arctic seed vault in Svalbard elicited a large response in most regions of the study area and a Guardian report on the consequences of melting permafrost for buildings in arctic cities (Luhn, 2016) was popular amongst British users. The latter is largely focussed on images demonstrating these consequences and the report's popularity could point to the fact that such visual clues – found to be so effective in communicating the severity of glacier retreat due to climate change – might also work well in conveying the dangers of permafrost thaw in a less abstract manner.

Tab. 21: Timeline of most notable (high anomaly) peaks in permafrost tweet frequency for individual years and countries. Daily anomalies are calculated as number of standard deviations from annual tweet frequency mean (per country)). Only days with a discernible cause for the high tweet frequency were included. *Ano = Anomaly, C = Country.

Year	Mth	C*	Ano*	Type	Cause and associated study (if available)
2011	Dec	FR	+6.7	GSN	Dangers of melting permafrost: GHG (Schuur & Abbott, 2011)
2012	Nov	FR	+7.1	GSN	Calls for the inclusion of permafrost emissions in climate change models (linked to COP18)
		ES	+5.7		
		GB	+11.1		
2013	Feb	DE	+8.4	GSN	Dangers of melting permafrost: GHG (Vaks et al., 2013)
		GB	+13.0		
2014	Mar	DE	+10.1	GSN	Dangers of melting permafrost: viruses (Legendre et al., 2014)
		GB	+13.2		
	Apr	ES	+6.8	GSN	Dangers of melting permafrost: GHG (Hodgkins et al., 2014)
	Oct	FR	+6.2	GSN	Dangers of melting permafrost: GHG
2015	Sep	FR	+11.3	GSN	Dangers of melting permafrost: viruses (Legendre et al., 2015)
	Sep	ES	+7.4	GSN	Dangers of melting permafrost: economics (Hope & Schaefer, 2016)
	Oct	GB	+15.3	GSN	Accelerated permafrost melt in Alaska (interview)
2016	Feb	CH	+6.5	NSN	Record permafrost temperatures recorded in 2015 in Switzerland (PERMOS)
	Aug	FR	+6.3	ND	Permafrost releasing anthrax kills 1500 reindeer and injures citizens in Yamal (Guarino, 2016)
	Oct	GB	+10.3	GN	Infrastructure damage in Arctic cities due to permafrost melt (Luhn, 2016)
2017	May	DE	+11.4	GN	Arctic seed vault floods due to permafrost melt.
		ES	+9.9		
		FR	+8.9		
		GB	+13.7		
		IE	+5.9		
	Aug	CH	+5.0	ND	Rockslide caused by thawing permafrost damages a mountain village in the Bregaglia valley, Switzerland
2018	Feb	ES	+8.2	GSN	Dangers of melting permafrost: mercury (Schuster et al., 2018)
	Dec	FR	+9.2	GSN	Dangers of melting permafrost: carbon / GHG
2019	Jun	AT	+5.4	GSN	Permafrost is responding rapidly to climate change (Farquharson et al., 2019)
		DE	+8.6		
		ES	+7.3		
		FR	+7.2		
		GB	+12.9		
		IE	+8.0		
2020	Jun	FR	+9.4	ND	Diesel spill in Arctic River in Russia caused by infrastructure collapse due to thawing permafrost (BBC, 2020)
		GB	+7.7		
	Jul	CH	+7.1	NSN	Accelerated permafrost thaw in Switzerland (PERMOS)

GN = Global News
GSN = Global Science News
NN = National News
NSN = National Science News
ND = Natural Disaster

Similar to the observations on the glacier timeline, Swiss users assign a much higher importance to domestic developments than other countries in the study area. The alpine country saw a fifth of its 2016 tweets occur on a single day in February, when national news outlets reported on a statement by Swiss Permafrost Monitoring Service PERMOS[18] revealing that the permafrost in the alpine regions of the country had reached record temperatures in the years prior. In 2017, a rockslide caused by thawing permafrost damaging a mountain village in the Bregaglia valley initiated the most discussion on a single day (adding to the overall list of discussed natural disasters). A further PERMOS report in 2020 with similar findings of accelerated permafrost melt in Switzerland once again had the largest single-day impact that year.

Overall, these science-based news stories on the state of the national permafrost seemed to elicit the largest reactions throughout the study period in Switzerland, including years not mentioned above. Interestingly, these findings coincide with the fact that the nation's users are consistently the most active contributors to the permafrost discourse in the study area (Fig. 44). This could suggest that a combination

[18] http://www.permos.ch/

of (1) the presence of permafrost in a user's country of residence as well as (2) news outlets reporting on the status of said domestic permafrost potentially leads to a much more wide-spread awareness of permafrost including its associated dangers and a more active discussion surrounding the topic on Twitter.

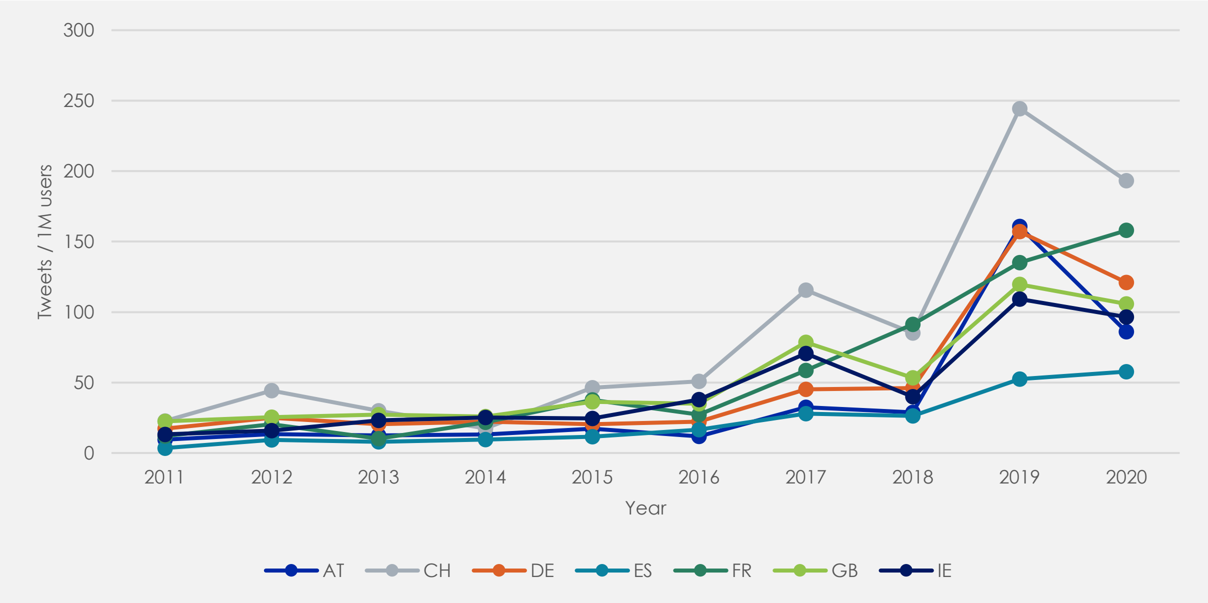

Fig. 44: Permafrost tweet frequency throughout the study period, normalised by twitter users on an annual basis.

A final point of interest in the permafrost timeline is the 2012 COP18 conference on climate change hosted in Doha, Qatar, which was solely responsible for any notable peaks in permafrost tweet activity during that year. The influence of these conferences will be discussed in more detail in section 4.3.4, however is it worth mentioning that the event elicited such a strong reaction in the discourse due to the controversial exclusion of permafrost from the upcoming IPCC AR5 report.

In summary, news reporting on the various dangers of permafrost found in a scientific context have the overall largest impact on users in the study area. Especially popular are reports on greenhouse gases and viruses emerging from the permafrost as well as a 2019 study identifying the thaw process in some artic regions to be 70 years premature. Natural disasters and infrastructure failure due to permafrost melt occurring during the study period also elicited a strong response and this immediate impact on humans and nature interestingly is a theme that did not emerge throughout the glacier timeline.

Once again, the local relevance of permafrost and related news reporting in Switzerland propels the country's tweet activity higher than the remainder of the study area. In contrast to the glacier timeline, however, is this trend not observable for any other regions containing large high alpine areas. Permafrost therefore remains a topic that is primarily treated as a somewhat distant and abstract global issue in large parts of the study area.

Sea level

In general, the largest annual reactions in most parts of the study area are elicited by news surrounding the consequences of sea level rise on the global scale. There are two rather distinct branches of how sea level research is framed, both by authors themselves and the news reporting on it: Research making (1) statements about the rate of sea level rise and its acceleration (Dangendorf et al., 2017; Hay et al., 2015; Kemp et al., 2011; Lambeck et al., 2014; Rahmstorf et al., 2012; Shepherd et al., 2018) and (2) the potential increase in sea level until the end of the 21st century (DeConto & Pollard, 2016; Garbe et al., 2020; Levermann et al., 2020; Nerem et al., 2018). It could be argued that the former inspires more urgency in the short-term as the latter's reference to the future might not have the same immediate impact on people. There does not seem to be a consistent spatial trend regarding the distribution of the two framings.

Tab. 22: Timeline of most notable (high anomaly) peaks in sea level tweet frequency for individual years and countries. Daily anomalies are calculated as number of standard deviations from annual tweet frequency mean (per country). Only days with a discernible cause for the high tweet frequency were included. *Ano = Anomaly, C = Country.

Year	Mth	C*	Ano*	Type	Cause and associated study (if available)
2011	May	FR	+11.0	GSN	SLR of 1m by 2100 according to Australian Climate Council
	Jun	DE	+10.5	GSN	SLR fastest in past 2000 years according to study, methodology is questioned (Kemp et al., 2011)
2012	Oct	GB	+9.3	NN/ND	Coastal flooding alert for Devon and Cornwall
	Nov	CH	+8.8	GSN	SLR 60 percent faster than predicted by AR4 (Rahmstorf et al., 2012)
		DE	+12.7		
		ES	+8.0		
		FR	+13.1		
2013	Aug	ES	+10.3	GSN	2012 was record-breaking year for ice loss and SLR (2012 state of the climate report) (Blunden & Arndt, 2013)
	Sep	CH	+7.9	GSN	Reactions to AR5 WG I release (IPCC, 2013)
		DE	+16.6		
		FR	+9.8		
	Dec	GB	+15.6	NN/ND	Severe coastal flood warning due to extreme storm surge on UK's East Coast
2014	Jan	GB	+7.7	NN/ND	Coastal flooding alert for Northern Ireland and Ireland
		IE	+5.4		
	Oct	ES	+8.2	GSN	Rate of SLR unprecedented in recent history (Lambeck et al., 2014)
2015	Jan	FR	+7.6	GSN	SLR faster than predicted (Hay et al., 2015)
		GB	+9.7		
	Aug	CH	+8.2	GSN	NASA data shows SLR could amount to 1m in next 100-200 years
		DE	+10.5		
		ES	+7.0		
2016	Jan	IE	+11.5	NN/ND	Coastal flooding alert for Ireland's East Coast
	Mar	ES	+11.5	GSN	Antarctica could contribute 1m to SLR in 21st century (DeConto & Pollard, 2016)
		FR	+9.6		
2017	May	ES	+12.0	GSN	SLR twice as fast as predicted (Dangendorf et al., 2017)
	Jul	DE	+7.1	GSN	Giant iceberg calves from Larsen C ice sheet
	Oct	IE	+8.4	NN/ND	Coastal flooding alert due to Storm Ophelia
2018	Jan	IE	+8.3	NN/ND	Severe weather warning and coastal flooding alert for Ireland
	Feb	AT	+8.1	GSN	Sea level will rise 65cm by 2100 (Nerem et al., 2018)
		CH	+12.0		
		DE	+16.4		
		ES	+7.7		
	Jun	GB	+6.0	GSN	Increase in Antarctic ice loss accelerates SLR (Shepherd et al., 2018)
2019	Sep	AT	+11.0	GSN	Reactions to SROCC release (IPCC, 2019)
		CH	+12.4		
		DE	+13.6		
		ES	+14.5		
		FR	+10.2		
		GB	+9.9		
2020	Feb	DE	+10.2	GSN	Record temperatures in Antarctica; Antarctica could contribute significantly to SLR in 21st century (Levermann et al., 2020)
	Sep	GB	+5.9	GSN	Contribution of Antarctic ice loss could raise sea levels by 2.5m (Garbe et al., 2020)
	Dec	IE	+7.3	NN/ND	Coastal flooding alert for Ireland

GN = Global News
GSN = Global Science News
NN = National News
NSN = National Science News
ND = Natural Disaster

One clear spatial trend is the coastal flooding threat perceived and experienced in the United Kingdom and Ireland. On multiple occasions in 2012-2014, 2016-2018, and 2020 is the highest peak in discussion caused by weather warnings related to storms and subsequent coastal flooding in one of the countries. In contrast, none of the countries in the study area on the European mainland show any regard for such domestic impacts and coastal flooding per se is never the main topic.

This distribution is not unexpected as it aligns rather well with various research on sea level threats in Europe, including the strongly increasing probability for compound flooding – the type generally discussed during the UK and Ireland peaks – in this specific area (Bevacqua et al., 2019: 3–4; Kirezci et al., 2020: 4, Fig. 45). Although the German North Sea coast also belongs to this affected area, only a small proportion of the country lives in close proximity to the sea, which might in turn explain a lack of interest in coastal flooding. France, where incidents along much of the northwestern coast are expected to strongly increase, is maybe the only outlier here as one would expect some interest in coastal threats. This lack of discussion on domestic coasts is not a bias introduced by the methodology (selecting only the highest peak in tweet activity per year), as the percentage of all tweets containing the keyword 'coast' tells a similar story: While tweets from France only mention the coast roughly 4% of the time (a similar value to landlocked Austria), the percentages in the United Kingdom and Ireland are considerably higher at 25% and 38%, respectively.

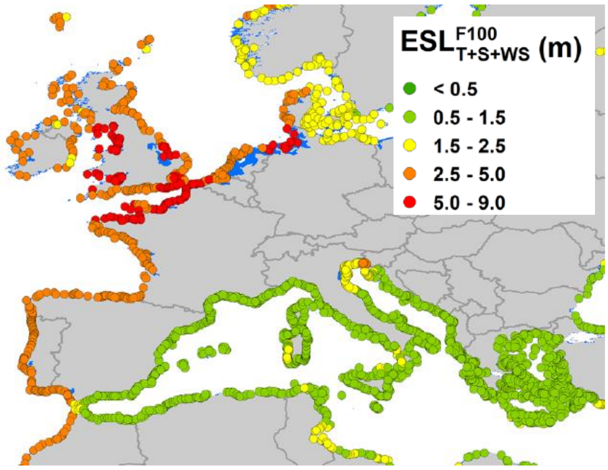

Fig. 45: Projected flooding in 2100 for an extreme sea level event (ESL) with a 100 year return period under the RCP8.5 scenario (Kirezci et al., 2020: 7).

Interestingly, the IPCC also has considerable influence on the public's discussion of sea level rise as especially its special report on oceans in 2019 (IPCC, 2019) drew strong interest from across the study area, far overshadowing any other sea level-related discussions on Twitter that year. The 2013 release of the first sections of the Fifth Assessment Report (IPCC, 2013) – which focusses on various climate change effects – saw a more limited impact restricted mostly to German and French-speaking regions. Particularly British users were far more concerned with coastal flooding on the country's East Coast in December that year. This prioritization is interesting as it suggests that domestic and immediate impacts of sea level rise cause a much larger reaction in the public than the maybe more abstract concept of global sea level rise.

In summary, the largest reactions regarding sea level rise on Twitter are rather consistently elicited by news reporting on studies' findings stating either that the rate of sea level rise is accelerating or that the expected increase in sea level by the end of the century will amount to some metric value. A major difference to trends in the glacier and permafrost timelines is the significant impact of certain IPCC reports, which had not been the case for the other LTETs. Possibly the most telling finding, however, is the high importance of local coastal flooding in the United Kingdom and Ireland, which consistently outperformed global science news in years with notable events. This continues the trend from the previously discussed timelines, further underlining the significance of local event relevance regarding LTETs and the associated increase in interest.

4.3.4 Impact of United Nations Climate Change Conferences (COPs)

Previous research has found that the occurrence of the annual United Nations Climate Change Conference (COP) can lead to significant spikes in climate change-related discussion on Twitter (Abbar et al., 2016: 5; Holmberg & Hellsten, 2015: 819; Kirilenko & Stepchenkova, 2014: 176–177; Stier et al., 2018: 1918). I will therefore investigate the impact of these key dates and others in following sections to further understand what elicits a public response on Twitter besides the main topics found in previous sections.

The impact of COPs on tweet activity surrounding the LTETs glacier, permafrost and sea level seems to be mixed and vary strongly from year to year. As Tab. 23 shows, high tweet activity coinciding with COPs seems to have only occurred at the beginning and end of the study period, while events held during the middle of the decade incited little response on Twitter.

Tab. 23: Tweet activity (TA) during the days of United Nations Climate Change Conferences (COPs). TA is quantified by the deviation from the mean number of daily tweets published during the period including the month before and after each conference. No COP was held in 2020 due to the COVID-19 pandemic.

Conference	Glacier TA	Permafrost TA	Sea level TA
COP17 Durban (2011)	High	High	Moderate
COP18 Doha (2012)	High	High	High
COP19 Warsaw (2013)	Low	Low	Low
COP20 Lima (2014)	Moderate	Low	Moderate
COP21 Paris (2015)	Low	Moderate	Low
COP22 Marrakech (2016)	Low	Low	Moderate
COP23 Bonn (2017)	Low	Low	Low
COP24 Katowice (2018)	High	High	Moderate
COP25 Madrid (2019)	High	Moderate	High

It should be noted that any connections between tweet activity and occurrence of COPs could be purely coincidental, however are some of the peaks observed clearly linked to the conferences, as closer inspection of tweets posted during these dates shows. For example, news stories covering the dangers of permafrost melt brought to the foreground during COP18 in Doha caused the largest reaction regarding this event type across the study area in 2012.

This strong connection between high tweet frequency and the occurrence of the 2012 COP18 conference in Doha, Qatar seems to be in agreement with previous literature. Kirilenko and Stepchenkova (2014: 176–177) found the event to be amongst the most important influences in 2012 with #COP18 being among the top five hashtags for climate change tweets that year. The subsequent lull in interest regarding COP19 (Warsaw, Poland) and COP20 (Lima, Peru) is also comparable with existing research, as Abbar et al. (2016: 6) found a similar lack of tweet activity surrounding the two events in their study analysing climate change tweets originating from Qatar. Holmberg and Hellsten (2015: 819) state that #COP19 was a frequently used hashtag in their study of climate change tweets, however is there unfortunately no quantitative value provided which could give an idea of the scale of its overall importance.

Maybe most surprisingly, the 2015 COP21 conference in Paris, France resulting in the 'Paris Agreement' did not elicit an exceptionally strong response, either during the conference nor in the weeks following it. This may be due to the fact that the negotiated terms only came into force a year later, when a sufficient number of countries joined the agreement (Savaresi, 2016: 20; Yeo, 2016).

In contrast to these results, Stier et al. (2018: 1918) found COP21 to be one of the most influential events in provoking climate change discussion on Twitter between November 2015 and June 2016. When comparing their findings against tweet frequency regarding the LTETs investigated in my thesis (Fig. 46), it becomes clear that there can be considerable discrepancies in the distributions of *direct* climate change tweets and climate change *process* tweets. Whilst the peaks of the former surrounding COP21 are somewhat discernible in the distribution of the latter, they are largely overshadowed by further peaks that neither align with COP21, nor any of the two other major events found by Stier and colleagues.

Furthermore, Abbar et al. (2016: 6) also found high interest surrounding the 2015 conference in their study of climate change-related tweets.

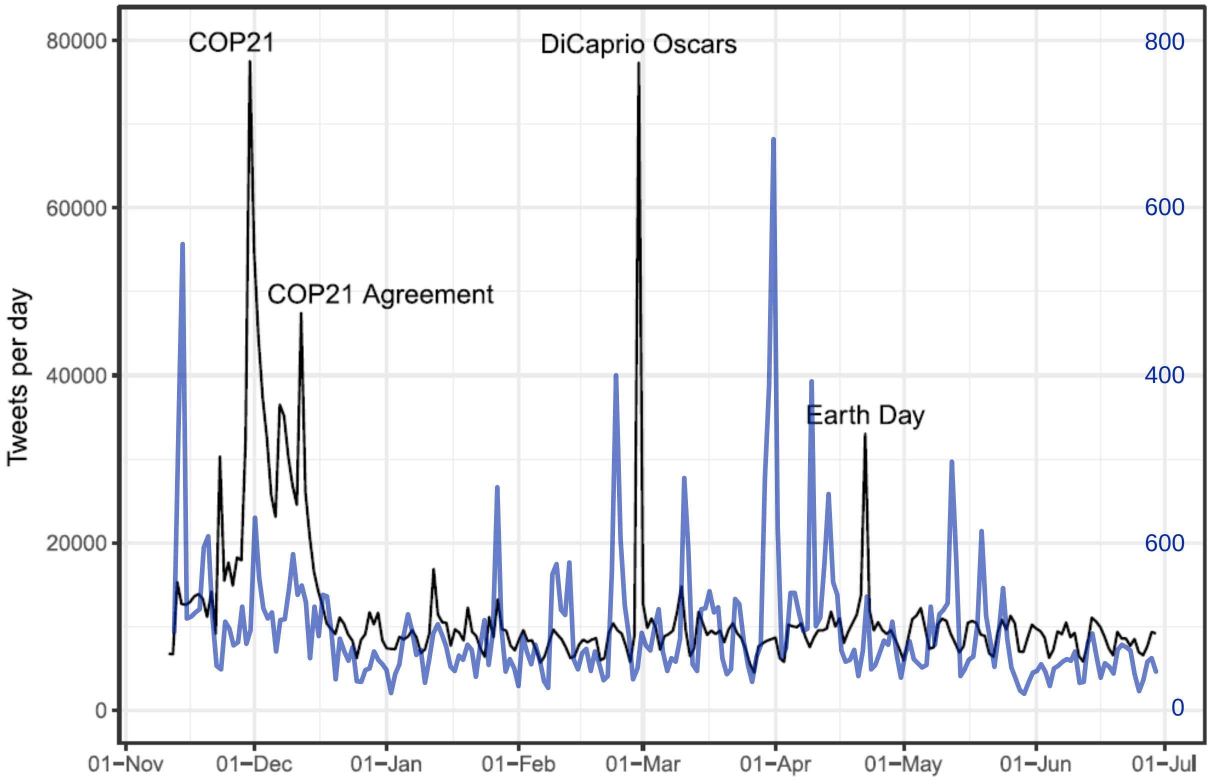

Fig. 46: Overlay of cumulative counts (combining glacier, permafrost and sea level tweets) from this thesis (blue) versus climate change tweet frequency observed by Stier et al. (2018: 1918; black). The authors collected tweets containing the hashtag #ClimateChange.

This is certainly interesting as it suggests that the discussion of LTETs on Twitter is not entirely tied to the discussion of climate change in general and inversely that wide-spread discussion of climate change does not necessarily lead to significantly increased discussion of LTETs. Subsequently, this strengthens the previously made argument that media reports and to some extent domestic events remain the strongest influence on tweet frequency for these event types.

While literature on the connection between tweet frequency and COPs in the latter half of the study period is sparser, the increased activity on Twitter during COP25 in Madrid, Spain might be due to an array of influences. Global climate protests and the declarations of climate emergencies by multiple governments arguably put the 2019 conference on a pedestal (Newell & Taylor, 2020: 580). This effect seems to be reflected in the tweet activity, as an investigation into the impact of the 2019 Global Climate Strikes in section 4.3.6 will show.

Overall, the influence of climate conferences such as the United Nations' COPs on Twitter discourse surrounding glaciers, permafrost and sea levels seems to be moderate at best (across the entire study period), with many years not seeing any direct reaction at all. Furthermore, causation between the conferences and high tweet activity – although observable in some individual cases – does not seem to be given anytime the two coincide, as many of the widely-shared news stories covering scientific findings during such peaks do not mention the UN conference specifically.

As literature on Twitter reactions to COPs largely focuses on climate change tweets directly, their findings of strong impact contrasted with the moderate impact on climate change process tweets observed in this thesis suggest that users might be inclined to tweet about climate change, but not necessarily its specific processes as a reaction to many of the conferences.

4.3.5 Impact of IPCC Assessment Reports (ARs) and Special Reports (SRs)

It has been argued that the release of reports by the Intergovernmental Panel on Climate Change (IPCC) signify an important event for the online discussion of climate change as especially the aspect of scientific evidence is often subject to debate (Kirilenko & Stepchenkova, 2014: 176; Pearce et al., 2014: 1). Similar to the trends observed with COPs, the magnitude of reactions in conjunction with the release of IPCC reports varies (Tab. 24). However are the reasons for this variation more directly tied to the reports themselves and especially their content.

Beginning with the special reports, it should be noted that the occasionally moderate and high tweet activity surrounding SRREN and SREX at the start of the study period are not related to the IPCC reports and instead result from unrelated science news published simultaneously. For the later reports in 2018 and 2019, however, the correlation between high tweet frequency and publishing of IPCC reports is valid upon closer inspection of individual tweets. Regarding the main topics of the Special Reports, these results make sense, as (relevant) peaks in tweet frequency only occur when an event type is the central focus of a report. It can thus be concluded that the publishing of IPCC Special Reports can elicit quite a strong response in the Twitter discourse surrounding individual climate change processes when said event types are crucial in the report.

Tab. 24: Tweet activity (TA) surrounding the days of IPCC Assessment Report (AR) and Special Report (SR) publishing. TA is quantified by the deviation from the mean number of daily tweets published during the period including the month before and after each report.

Report and topic	Glacier TA	Permafrost TA	Sea level TA
SRREN (2011) *Renewable energy and CC mitigation*	Moderate	Low	Low
SREX (2012) *Extreme events and disasters*	Low	High	Moderate
AR5 WG I (2013) *Physical science basis*	Moderate	Low	High
AR5 WG II (2014) *Impacts, adaptation and vulnerability*	Low	Low	Low
AR5 WG III (2014) *CC mitigation*	Low	Low	Low
AR5 SYR (2014) *Synthesis Report*	Low	Low	Moderate
SR15 (2018) *Global Warming of 1.5 °C*	Low	Low	High
SRCCL (2019) *CC and land*	Low	Low	Low
SROCC (2019) *Ocean and Cryosphere*	High	High	High

When it comes to the Fifth Assessment Report (AR5) and its various stages of publishing, the interpretation gets somewhat more complicated due to the comprehensive nature of the report. As a general trend, it seems that when such a document touching on all aspects of climate change is released, the strongest response is elicited in the sea level discourse, suggesting that this might be the main take-away news media and to some extent the general public extract. Interestingly, this heavy bias towards sea level rise as the focal point of AR5 WG I is not apparent in the Summary for Policymakers of the report (IPCC, 2013) most articles are referencing. Instead, sea level rise is listed evenly amongst an array of other climate change processes in the atmosphere, ocean, and cryosphere. This raises the question whether members of the public would themselves single out sea level rise, as news media clearly steer the discourse with 78% of tweets following the publishing of AR5 WG I containing a link to a news article.

4.3.6 Impact of Global Climate Strikes 2019

The formation of the Fridays for Future (FFF) movement in 2018 sparked a new wave of climate activism by younger generations around the world, culminating in four so-called 'global climate strikes' during 2019 (March, May, September, and November), where millions took to the streets worldwide to protest for more action on climate change. The strikes garnered wide-spread media attention and it is thus expected to see their impact on Twitter as well (Boulianne et al., 2020: 208; Laux, 2021: 414, 416).

Going in chronological order, the first strike on March 15th seems to have had very little impact on the tweeting activity regarding glaciers, permafrost and sea levels (Appendix, Fig. 56). In fact, the event was largely overshadowed by reactions to news articles on glacier retreat during the months before and after the strike took place. A similar story can be observed for the days surrounding the second Global climate strike on May 5th, where the tweet activity increases slightly subsequent to the event but remains around an average level (Appendix, Fig. 57).

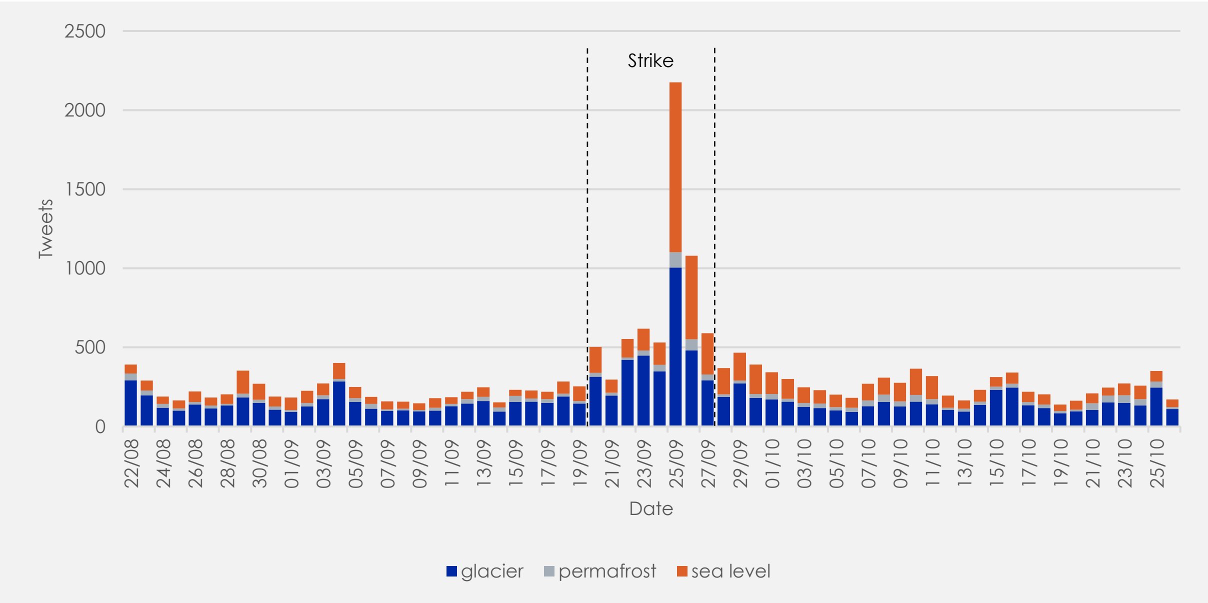

Fig. 47: Total number of glacier, permafrost and sea level tweets across the study area during the days surrounding the third Global Climate Strike (Sep 20-27 2019).

It is only during the third Global Climate Strike in September 2019, when the event is held throughout a week and 6 million people across 185 countries worldwide participate (350.org, 2019) – marking the largest extent of the protests – that a strong influence on tweet activity regarding long-term climate change processes can be observed (Fig. 47). The impact is quite strong, and the number of tweets published during the seven days of the event alone makes up roughly 15% (glacier), 10% (permafrost) and 22% (sea level) of all tweets published in September during the entire ten years of the study period for the respective event types.

The fourth strike on November 29th saw another decent impact on tweeting activity, reaching approximately half the magnitude of the single-day peak observed during the September protests (Appendix, Fig. 58). When compared to the first strike in March, the fourth strike only marked a 43% increase in participants (Barclay & Amaria, 2019; Laux, 2021: 416), but displayed a 325% increase in combined glacier, permafrost and sea level tweets, meaning that the event had a very strong impact online. The strike was held days before the UN COP25 in anticipation of the conference, which might have aided its visibility as the topic of climate change is assumed to have been of above-average relevance in the media and public eye at the time.

Overall, these observations suggest that focused, large-scale protests do have a place in raising climate change awareness and concern, as they can spark online discussion not only on climate change itself, but also on its effects and processes as shown with the strong peaks in glacier, permafrost and sea level discourses on Twitter. It also seems that holding protests in conjunction with other international climate change-related events such as COPs can help increase the visibility of protests and result in much larger tweet activity regarding the LTETs mentioned. In turn, large enough protests also seem to play a part in increasing the visibility of COPs.

4.3.7 Impact of COVID-19 pandemic

The global COVID-19 pandemic that began in December of 2019 in Wuhan, China and subsequently spread across the world throughout 2020 had severe impacts on public health, lifestyle and the general economy (Gautam & Hens, 2020: 4953; Shi et al., 2020: 343). Meanwhile, Twitter experienced considerable growth in its user base during this period, with a 12% increase in daily active users between quarters 1 and 2 of 2020 – the point at which a pandemic was declared by the World Health Organisation (WHO, 2020) – alone, and a general increase in users compared to 2019 (Statista, 2021). Therefore, one would expect – all things being equal – a corresponding rise in tweet numbers. With previous research however showing that climate change is often not the public's primary concern, especially when overshadowed by global events such as economic crises (Ruiz et al., 2020: 114; Whitmarsh, 2011: 691), the question arises how much of an impact the COVID-19 pandemic had on tweet activity surrounding climate change-related events.

When comparing 2019 and 2020 tweet frequencies on a monthly scale, different trends emerge between STETs and LTETs (Fig. 48). Especially the distribution of heatwave tweets seemingly continues to be mainly influenced by environmental variables, as tweet frequency still follows closely to European summer temperatures in 2020. For wildfire tweets, there does seem to be a negative trend starting in March, however is it difficult to quantify how much of this decreased tweet activity is due to COVID-19 instead of variations in global and local wildfire activity. Considering the fact that 2020 was a particularly bad year for wildfires with many large-scale events worldwide (Burke et al., 2021: 5; Xu, Yu, et al., 2020: 2173), it does seem, though, that tweet activity surrounding wildfires was considerably lower than one would expect without external forcings. Interestingly, the compounding effect of wildfires on the severity of COVID-19 symptoms and rate of excess deaths (Meo et al., 2020: 10287; Zhou et al., 2021: 6) did not seem to significantly increase public interest in wildfires on Twitter either during 2020.

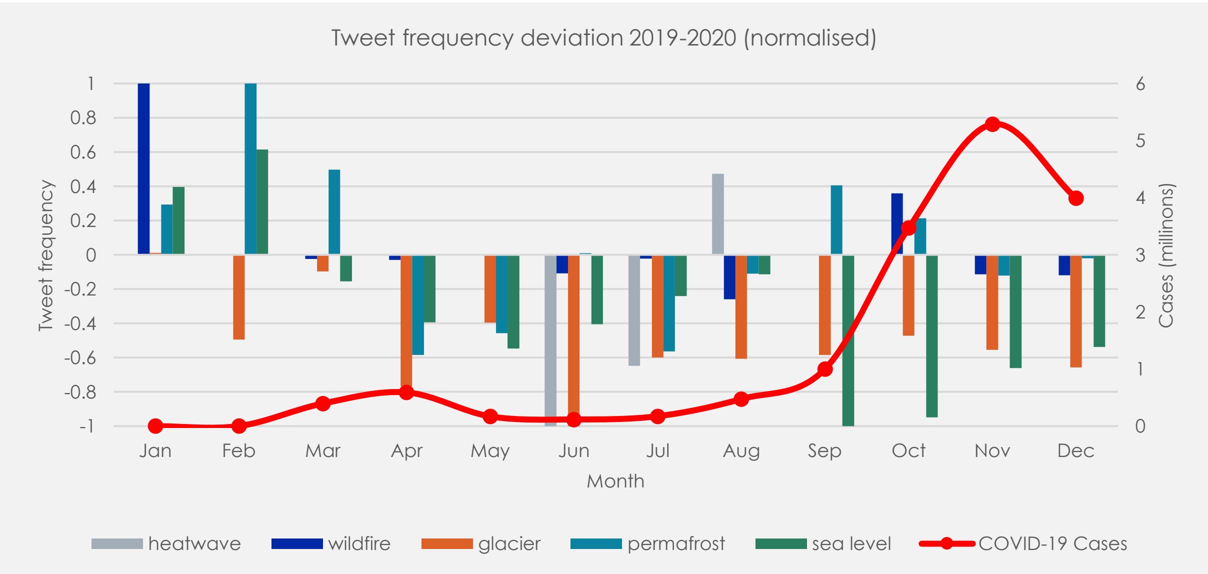

Fig. 48: Monthly deviation of 2020 tweet frequency compared to 2019 values. Deviations have been normalised to allow for comparison between event types. Monthly number of COVID-19 cases in the European Union during 2020 (Ritchie et al., 2020) are plotted on a secondary axis for comparison.

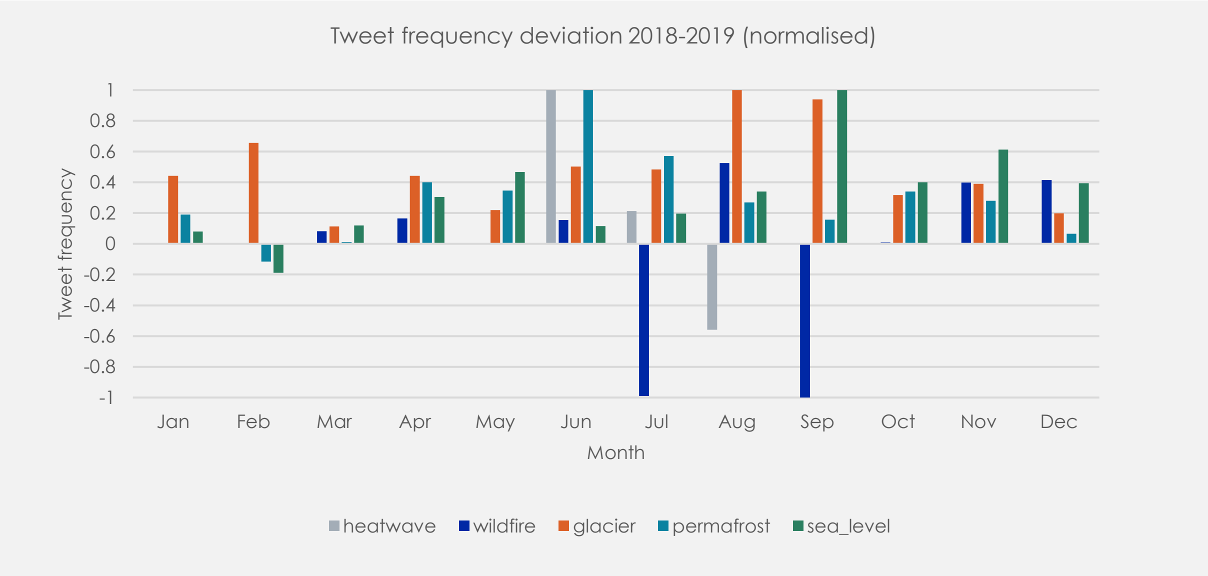

Fig. 49: Monthly deviation of 2019 tweet frequency compared to 2018 values. Deviations have been normalised to allow for comparison between event types.

The picture seems much clearer when looking into LTETs, as all three event types demonstrate similar temporal developments. As these topics are less tied to real-world events, the beginning of the year shows an increase in tweet frequency, which would be expected as a continuation of the trend observable in Fig. 50. Subsequently, however, a strong downwards trend is observable in conjunction with the first wave of COVID-19 cases in Europe. The resurgence of 2020 tweet activity in August and September may either be due to the publishing of wide-spread articles (which is likely the case for glacier tweets) or could come as a delayed result of low case numbers in the months prior and a general de-escalation of the situation. With the second wave in fall and winter, tweet activity once again decreases below 2019 levels.

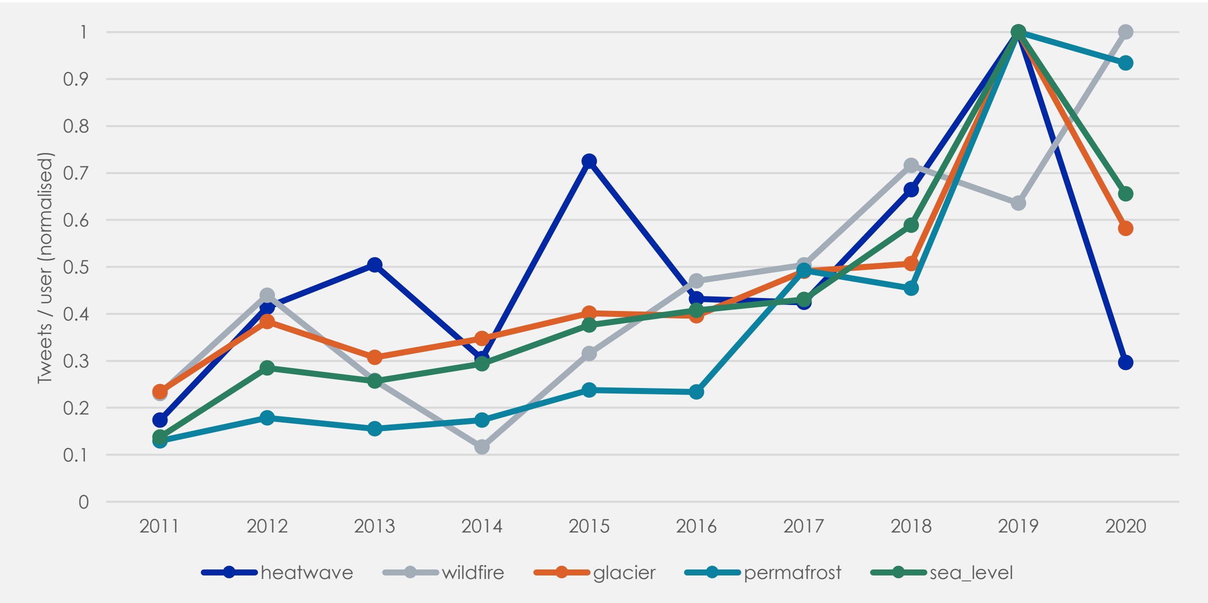

Fig. 50: Annual tweets per user for each event type. Distributions were normalised to enable comparison between the different event types.

In conclusion, the COVID-19 pandemic seems to have had considerable impact on tweet frequency surrounding LTETs, decreasing activity despite a considerable increase in Twitter users during the same time period. Especially when comparing the 2019-2020 deviations to the 2018-2019 deviations (Fig. 49) and the abrupt end to the overall upward trend over the study period (Fig. 50), the impact of the pandemic on tweeting behaviour related to climate change processes is undeniable. The reason why wildfire tweets peak in 2020 instead of 2019 are the Australian bushfires in January before the onset of the first wave in Europe. Tweets during this month make up roughly two thirds of the year's total, meaning that there is also a significant drop in wildfire tweets during the pandemic-affected part of 2020. The trend is therefore consistent among all event types.

For STETs, the true impact is harder to quantify as the environmental influences certainly still play a major role. At least for wildfires, however, there does seem to have been a decrease in tweet activity to some extent as a result of the pandemic despite a very active wildfire season and the added health risk of wildfires increasing COVID cases and deaths. These findings are in overall good agreement with the literature on prioritisation of climate change mentioned above.

5. Discussion

The following chapter will critically assess the results presented in the previous chapter, draw connections between different findings, and put them into context with existing literature. The structure from the results section is largely maintained and the corresponding research questions asking *what*, *where* and *when* public reactions to climate change effects are elicited carry over. The argumentation within the different section, however, draws on the full range of findings to contextualise them optimally. In a final, additional section, I will reflect on the data and methodology used in this thesis and discuss their advantages and limitations.

5.1 Thematic Dimension

The discussion of thematic findings will answer the research question of *what* elicits strong reactions amongst the public's discourses on climate change processes. Specifically of interest is how climate change itself is involved in the debate, how risk perceptions vary between the event types, and how the interplay between news and science materialises.

5.1.1 Prevailing importance of traditional news media

It can be observed that the online versions of traditional news media such as the BBC and The Guardian dominate the science communication regarding all event types on Twitter throughout the study period. This is the case despite many academic publishers, scientific societies, research centres, and scientific journals being active on Twitter (López-Goñi & Sánchez-Angulo, 2018: 1). Within the news articles, direct references to the paraphrased papers via web link are rare and often only the name of the journal or a lead author is stated. In many cases, there is no information identifying the paper, and 'scientists' is used as an umbrella term for the source.

In the context of Twitter, the prevalence of news media is not surprising as trending topics on the platform are often aligned with news tweets (Shariff et al., 2017: 785). Interestingly, my results differ from findings in the United States, where more internet users turned to non-traditional online sources than online versions of traditional news outlets (Brossard & Scheufele, 2013: 40). However, traditional media generally remains an important source of information on climate change for the public (Ruiz et al., 2020: 114).

The above results underline the power traditional news media (still) have over science communication to the general public in a digital world, as studies have shown that especially the tone when delivering and commenting on scientific findings is crucial in forming a reader's opinion on an otherwise balanced science story (Brossard & Scheufele, 2013: 41). More generally still, traditional news media are known to have a strong influence on anyone's perception of the world and the public tends to hold news tweets in high regards, often struggling to separate them from misinformation and accepting their interpretations of science as truth (Brossard & Scheufele, 2013: 40–41; Ruiz et al., 2020: 114; Segev & Hills, 2014: 67; Shariff et al., 2017: 794). This implies that the public's perception of climate change and its individual processes is likely shaped by news media, at least when it comes to the science behind it. Involvement of political influences in this process is also likely as politics have been shown to be a driving factor steering the focus of news reporting on climate change (Schäfer et al., 2014: 169).

One could argue that this dominance over science communication on Twitter is partly due to the fact that direct access to scientific findings remains limited for the general public (Fecher & Friesike, 2014: 29). Especially in the case of tax payer-funded research, the argument is made that the public should be able to access the results free of charge, as they have already contributed to their contribution (Phelps et al., 2012: 2). In general, the free publication of scientific findings would help increase the audience for direct science communication – bypassing news media – by removing the price barrier of the current publication model (Carroll, 2011: 1). The promising potential of social media to become a direct channel of communication between research and the public (López-Goñi & Sánchez-Angulo, 2018: 3) has therefore not yet been fulfilled in the context of climate change processes, indicating room for improvement in this aspect.

There does seem to be, however, a general interest in scientific findings from traditional academic research sources as a large majority of the top-linked web resources in tweets belonging to the science theme led to articles either referencing scientific journals, reports or interviews with scientists themselves. This suggests that the basis for scientific discussion on Twitter surrounding climate change processes is in essence founded upon peer-reviewed research, although communicated through the lens of news media. In turn, this implies that the general public in Western Europe partaking in these discourses is highly susceptible to 'proper' science, which offers a promising basis for introducing and justifying adaptation and mitigation measures through policies.

5.1.2 Climate change and its dangers

The discourses surrounding LTETs are much more focused on climate change than those encompassing STETs, however have the latter seen a recent increase in importance. This is interesting as a wealth of research suggest that the more tangible short-term effects of climate change lead to higher awareness of and concern about the issue (Borick & Rabe, 2010: 785; Frondel et al., 2017: 180; Hamilton & Keim, 2009: 2351; Lee et al., 2015: 1016). However, even when taking into consideration the absolute number of tweets (which is highest in STETs), the trend cannot be compensated for. The gap separating a process' recognition and attribution to climate change therefore seems to be much wider for STETs than LTETs, with the former seeing the overall bigger recognition and the latter exhibiting a higher attribution rate. In other words, more people recognise or tweet about STETs, whereas the percentage of people attributing an event to climate change is much higher in LTETs.

This is not entirely unexpected as the connection between local weather events and climate change is not always made by the general public, even flood victims (Boudet et al., 2020: 72; Lorenzoni et al., 2007: 452). LTETs are furthermore often primarily used to contribute to a climate change discourse (as indicated by high keyword frequencies), while STETs see larger numbers of users discussing immediate personal impacts. This is unsurprising due to the long-term and short-term nature of the respective event types, but also indicates that there is potentially a lack of awareness or concern about the connection between STETs and climate change.

Where attribution is given (i.e., the connection between an event and climate change is drawn), *a considerable overlap between climate change and danger can be observed*. Perceptions of serious risk are distributed somewhat more evenly between the event types, can however vary strongly between regions during the same event. During a heatwave, for example, some populations are more concerned about dangers to human health, while others worry about the consequences for the environment. These variations in risk perceptions both on the national and individual level agree with similar findings in the literature (Frondel et al., 2017: 173). Furthermore, climate change attribution for heatwave and wildfire events in legacy media varies across space and time as well (Hopke, 2020: 504), which could be both cause and effect of the observed public risk perceptions as news media can be a strong influence (Segev & Hills, 2014: 67).

This highlights the fact that different (temporally evolving) value judgments exist within the study area, which could for example be driven by news media and the aspects of an event they deem most important to report on (see also section 5.3.2). This must be kept in mind when applying climate policies, as certain adaptation and mitigation measures align differently with local risk perceptions – which are crucial for the support of such measures (O'Connor et al., 1999: 469–470) – and might therefore fail to carry over successfully from one region to another.

5.2 Spatial Dimension

The discussion of the spatial results revolves around the research question of *where* reactions are elicited (*RQ1*) and why this is the case (*RQ2*). The following sections will answer these questions by presenting the spatial grouping of common interest as well as describing how spatial distributions are linked to tangibility and how domestic and transnational trends give insights into the spatial priorities of populations.

5.2.1 To each their own: distinct regions and their specialisations

Risk perceptions related to climate change vary considerably between countries, which agrees with previous literature (Frondel et al., 2017: 173). Broadly speaking, the study area can be divided into four general regions where results are somewhat similar across many of the event types (Fig. 51). These are the British Isles, the Alps, Germany, and the Southwest (non-alpine France with Spain). It is important to note that bigger cities are somewhat exempt from this classification as they consistently exhibit relatively high tweet frequency across most event types.

In the British Isles, users are among the populations least susceptible to heatwaves as they exhibit the highest median thresholds for tweeting compared to the historic temperatures. In contrast, users are especially active when it comes to the threats of coastal flooding (sea level), which sees very little discussion in the rest of the study area. These results align well with findings by Taylor et al. (2019: 158), which came to the conclusion that Brits far underestimate the future impact of heat extremes while assigning appropriate importance to flooding risks.

The alpine regions unsurprisingly have a special interest in glaciers, which is additionally quite strongly domestic. The same can be said for permafrost tweets, however is the clustering around the Alps less pronounced in this case. Switzerland especially stands out, as interest in the alpine events is very high and domestic developments make up an unusually high percentage of major events in the timeline. When it comes to sea levels, this landlocked region exhibits generally low interest in the subject. This is interesting, as glaciers and sea levels are often connected thematically due to the considerable contribution of the former to the latter (Milne et al., 2009: 472; Nicholls & Cazenave, 2010: 1517), which was also observable in my data (section 4.1.1). It seems, though, that the domestic focus and landlocked nature of the region override this link. It can also be observed that this region has the lowest temperature threshold for heatwave tweets, however is the tweet activity regarding the subject generally rather low.

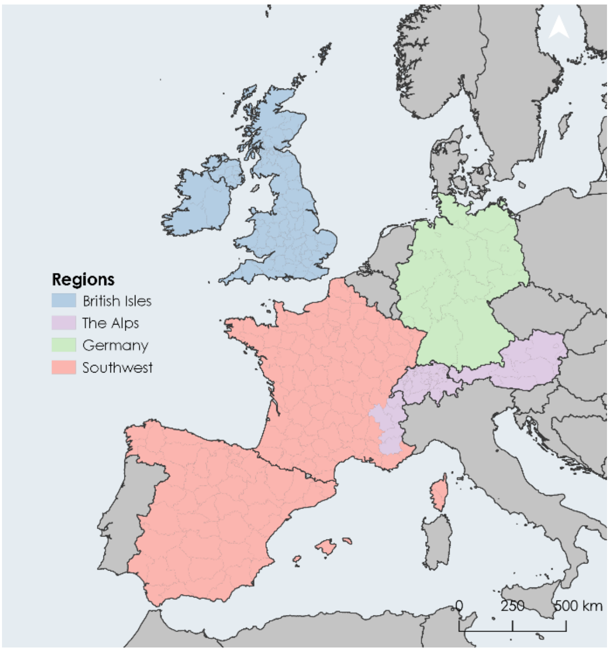

Fig. 51: Regions of the study area showing broadly similar trends across the different event types.

The Southwest consisting of Spain and the non-alpine areas of France stands out when it comes to heat-related events. Heatwaves elicit substantially stronger reactions in this region compared to the rest of the study area, which coincides with the region experiencing the hottest summers of the study area (Kjellström et al., 2007: 254). As wildfires are closely related to extreme temperatures (Sun et al., 2019: 129) and Spain as well as southern France experience the most frequent and largest fires in the study area (Eriksen, 2020: 2), high tweet frequencies in these areas regarding the subject are unsurprising. Interest in the LTETs is mostly low to very low in the Southwest.

Germany is often the odd one out as it does not group itself with the coastal flooding specialisation of the British Isles, the glacier and permafrost focus of the Alps, or the special heat and fire interest of the Southwest. German users show average tweet activity as well as average transnationality for most event types, resulting in a low specialisation. This is possibly due to a general lack in constant large-scale events during the study period, as they did react strongly when massive forest fires occurred in the country's northeast in 2018 and 2019. Even though heatwaves might occur somewhat frequently, German users have one of the highest temperature thresholds for tweeting when compared against the historic mean and they tend to be less concerned about personal impacts (e.g., section 4.1.5). Previous research has shown that heatwaves are indeed perceived as less of a risk than for example storms in Germany (Frondel et al., 2017: 174).

It can be concluded that spatial influences play a role in shaping public climate change perception as populations tend to focus their interest on a genre or specialisation of climate change effects that are especially relevant in their local or national environment. The distribution is, however, rather broad and diffuse, suggesting that further aspects must significantly influence how the public reacts to climate change.

5.2.2 Tangibility as a major driver of reactions

Overall, the largest reactions are elicited where the effects of climate change can either be directly felt or affect areas and people close to a user. In these cases, the spatial component becomes the main driver for reactions. Heatwave tweets thus see the highest peaks of the selected event types, as temperature is a continuous phenomenon that can be perceived constantly. Wildfire tweet distributions follow this logic to a certain extent, as significant local fires elicit the highest tweet activities in the study area. As wildfires are strongly influenced by extreme temperatures, the two event types can be grouped under the same umbrella in this context. These trends are in full agreement with previous research, as local weather and temperature have been shown to be some of the strongest influences on climate change perception in general (Borick & Rabe, 2010: 785; Hamilton & Keim, 2009: 2351; Lee et al., 2015: 1016).

When the effects of climate change processes are less tangible for the population of the study area – as is the cases with LTETs – spatial aspects decrease in importance and tweet activity fluctuations depend increasingly on temporal and thematic influences. Nevertheless, the strongest reactions are still generally elicited where LTETs have some local importance, as the clustering of glacier tweets around the Alps and the higher occurrence of sea level tweets in coastal regions compared to landlocked regions show. The relatively frequent danger of coastal flooding in the United Kingdom and Ireland described above is a further example, as the two countries exhibit some of the highest activity regarding sea levels. Therefore, *elicited reactions are larger where the effects of climate change are stronger and more impactful*. This is further supported by the fact that the intensity of temperatures across the study area aligns well with tweet frequency. Additionally, the high tweet activity in urban areas is likely coupled to the urban heat island effect exacerbating local temperatures (Kleerekoper et al., 2012: 30).

The decreased interest in event types that do not seem to directly affect a user's environment is, however, unwarranted as the LTETs all have potential future consequences for people not directly affected by them today. The extensive retreat of glaciers projected for the European Alps is expected to affect many regions of the study area by influencing agriculture and hydropower amongst others (Beniston et al., 2018: 772–773), while general glacier retreat will impact any coastal regions due to its contribution to sea level rise (Gardner et al., 2013: 857). Meanwhile, permafrost not only threatens to accelerate climate change through its release of greenhouse gases (Jorgenson et al., 2010: 1233; Schuur et al., 2015: 176–178) – leading to tangible temperature increases everywhere – but also is the potential source for the re-emergence of viruses that had remained locked away in the frozen ground for thousands of years. This revival does not necessarily occur immediately upon release, but is increasingly carried out in the context of research (Legendre et al., 2015: E5327; Miner et al., 2021: 814). The impact of sea level rise on landlocked regions will come with the migration of endangered coastal populations, which has been largely neglected in academic literature (Hauer, 2017: 324) as well as Twitter users judging by the spatial distribution of sea level tweets. Users are therefore not nearly as susceptible to the future tangibility of climate change effects as they are to the present one. This could pose as a hindrance when targeting such future issues with climate policies as the public may not yet be concerned about their eventual impacts. It further coincides with the fact that climate change in general is perceived as a distant issue by the public (Ruiz et al., 2020: 114). Simultaneously, media influence on shaping opinions grows when topics are seen as distant and personally unrelated rather than local and immediate (Segev & Hills, 2014: 70).

Both the finding that attribution of STETs to climate change is lacking (section 5.1.2) as well as the above mentioned failure to recognise local impacts of distant long-term climate change processes are problematic as the public's recognition of these links is argued to be instrumental in raising support for climate action and its necessity (Lorenzoni et al., 2007: 452; Ruiz et al., 2020: 114). Strengthening the comprehension of these connections should significantly motivate concern about climate change (Howe et al., 2012: 1).

5.2.3 Differences in domestic and transnational interests

When event types have strong local impacts, users focus less on international events. Unsurprisingly, this results in a split between STETs and LTETs, with especially heatwave tweets showing a consistent domestic-leaning interest. Wildfire events interestingly developed from a similar domestic interest to a decidedly transnational relationship in the latter half of the study period, which is likely due to the increased occurrence of exceptionally large and devastating fires outside of the study area in recent years (Eriksen, 2020: 2; Xu, Wang, et al., 2020: 2173). Generally, the transnationality of wildfires fluctuates according to the occurrence of important events within or outside a country.

For LTETs, some countries show stronger domestic interests than others. The distribution follows the previously mentioned trends with the alpine countries showing increased domestic interest for glaciers and permafrost and countries containing a coastline showing increased domestic interest for sea level compared to landlocked countries. This leads to the conclusion that *the local tangibility of a subject not only increases the tweet frequency, but also results in a more domestic discussion of the subject*. For example, the presence of glaciers nearby therefore does not necessarily increase a user's overall interest in glaciology, but specifically increases the interest in those nearby glaciated areas. Meanwhile, users distant from any such relevant features tend to have a more international perspective on the climate change effects.

Complimentary to these results, studies investigating legacy media have found discourses to generally be domestic first and transnational second but reporting on international politics as well as events such as natural disasters to be common as well. Whether this finding would hold true for web platforms was unclear as legacy media had been observed to better widen the spatial scope of domestic climate change discourses to an integrated transnational perspective compared to the online sphere (Reber, 2020: 2). My findings suggest that this uncertainty has some merit as there is considerable variation in transnationality under the umbrella of climate change. Not only does the transnationality of a discourse depend on the type of climate change process, but also on the temporal and spatial context it is discussed within. It is possible that the transnationality of the discourses is strongly connected to the sharing of legacy media as increased transnationality coincides with less tangible event types[19].

This not only implies that the public develops a more international perspective on events not occurring in their close proximity due to an increased reliance on news media, but also that the public is expected to be less aware of climate change effects in other parts of the world when they are tangible and can be experienced nearby. The knowledge about international climate change developments is therefore seemingly dependent on domestic impacts as the two somewhat complement each other.

[19] Only 34% of the strongly domestic heatwave tweets include a URL, while 72% of the strongly transnational glacier tweets include a URL.

5.3 Temporal Dimension

Results from the temporal analysis form the basis for answering the research question of *when* large reactions are elicited (*RQ1*) and what the driving forces behind these patterns are (*RQ2*). As with the corresponding results section, the purpose of the temporal analysis is to investigate LTETs, since the temporal aspects of STETs are largely explained by tangibility (sections 4.2.3 – 4.2.4). The following discussion will therefore mostly be relevant for LTETs only.

5.3.1 Alarmist or alarming? How news headlines determine public reactions

Almost all of the largest reactions in each country and year are elicited by news reports for LTETs which then spread across Twitter as users share articles with similar headlines and contents from various news outlets, often in different languages. Fluctuations in media attention to climate change-related issues can be impacted by current scientific and political events on one hand and information fatigue associated with attention cycles on the other hand (Ratter et al., 2012: 6–7). In many cases, especially large reactions occur when a subject is either of special domestic importance or descriptors such as ‘rapid’, ‘irreversible’, ‘worse than we thought’ and similar are used in conjunction with contextually large numbers. While the former has already been discussed above (section 5.2.2), the usage of ‘dramatic’ language offers plenty of fuel for debate.

Risbey (2008: 26, 34–35) asks whether such language should be neutrally termed ‘alarming’ as it accurately represents the urgency of the climate crisis or whether it should carry the negative connotation of ‘alarmist’, representing an overreaction inconsistent with science. He states that climatologists are split on the issue but concludes from his review that ‘alarming’ is the more appropriate term as the terminology fits the scale of the problem. This does not necessarily match the perception of the public, however, which has increasingly been referring to ‘exaggerated’ and ‘alarmist’ news reporting regarding climate change (Borick & Rabe, 2010: 791; Lorenzoni et al., 2007: 452; Whitmarsh, 2011: 697).

Seeing as ‘alarming’ headlines did in fact seem to be effective in eliciting large responses from the public in the context of my thesis, it might be in the interest of the scientific community and policy makers alike for this trend to continue as the public becomes more aware and concerned about the potentially drastic impacts of climate change. If the public, however, develops a strong negative connotation to this reporting style or simply forms information fatigue as a result, concern and subsequently support for climate action could see a distinct drop. High increases in media coverage have already been juxtaposed with decreases in serious concern about climate change in the western world (Ratter et al., 2012: 5–6). It is also not distinguishable from my results whether the users reacting to such headlines already belong to the high concern group, or whether new people are added to this group.

Simply bombarding the public with the alarming evidence of climate change’s consequences is further not necessarily very affective due to the ‘perception gap’, which describes the arguably irrational way individuals perceive risks too much or too little despite scientific evidence (Ropeik, 2012: 1222). Ideally, communication related to climate change should therefore be nuanced and take into account how the public *feels* about it. No matter how ‘alarming’ the scientific evidence might be, a discourse perceived as ‘alarmist’ will hardly be successful in gaining new supporters for climate action. Similarly, however, is it questionable whether a less ‘alarmist’ news media would be more successful as the targeted part of the population already resides on the other side of the perception gap anyways (Ropeik, 2012: 1224–1225).

5.3.2 Impact of major political and scientific events

Major political and scientific events related to climate change have varying impacts on specific event tweet activity. The influence of climate conferences such as the annual United Nations Climate Change Conference (COP) on Twitter discourse surrounding glaciers, permafrost and sea levels only rarely reaches high levels and seems to be moderate at best across the entire study period, with many years not seeing any direct reaction at all. This somewhat contrasts literature on general climate change tweets, which found that the occurrence of the United Nations' COPs can lead to significant spikes in climate change-related discussion on Twitter, however is this trend not entirely continuous either (Abbar et al., 2016: 5; Holmberg & Hellsten, 2015: 819; Kirilenko & Stepchenkova, 2014: 176–177; Stier et al., 2018: 1918). Political activism in the form of organised protest is far more efficient in eliciting strong reactions on social media, as for example the Global Climate Strikes in 2019 saw tweet activity soar surrounding the dates of the bigger events. The wide-spread coverage of the protests observed in news media (Boulianne et al., 2020: 208; Laux, 2021: 414, 416) seems to have translated to the online sphere as well.

The releases of scientific reports by the Intergovernmental Panel on Climate Change (IPCC) see an overall similar level of impact as COPs. Once again, this somewhat subverts expectations from literature on general climate change tweets, which predict a higher importance for the reports (Kirilenko & Stepchenkova, 2014: 176; Pearce et al., 2014: 1). Interestingly, reports only seem to influence tweet activity strongly when they are thematically related to an event type. Overall, sea levels saw the most discussion resulting from the IPCC reports, even in instances such as the Fifth Assessment Report (AR5), where sea levels were reported on equally alongside other issues. This indicates a bias towards sea level rise as the emblematic climate change process as especially news media seems to steer towards it when a balanced report with varying event types is presented.

Taking into consideration the above discussed strong influence news media has on both online science communication and public perception of climate change, this is of great interest. The platform of the balanced IPCC report offers an insight into how strongly traditional media can gravitate towards certain aspects of climate change, not only resulting in unbalanced reporting characteristics, but also potentially shaping a strong bias in public climate change perception towards certain facets of global warming.

A secondary conclusion is that the public's perception can vary strongly between climate change as a whole and individual climate change processes. As demonstrated by the above discussed disparities between peaks in tweet frequency related directly to climate change on one hand and its individual effects on the other hand, there seems to be somewhat of a disconnect on Twitter between the general climate change discourse and the more in-depth discourse on climate change effects. Stated simply, when many people are tweeting about climate change, a significantly lower number mention specific processes affected by climate change. Subsequently, findings from research on climate change reactions as a whole are not necessarily representative of sub-discourses on individual effects. This becomes apparent when for example the impacts of political and scientific events do not quite align between the two different tweet corpora. Future research contributing to the growing corpus of literature on public reactions to climate change on Twitter therefore has to make this distinction between climate change and its individual processes.

5.3.3 Priorities during crises: when climate change goes forgotten

Climate change processes are not a top priority for the general public. This becomes abundantly clear in the context of the COVID-19 pandemic, as tweet activity drops significantly with the onset of the first wave in Europe. Despite a sharp growth in the Twitter userbase and a steady increase of event tweets per user during the latter half of the study period, absolute (as well as population normalised) tweet numbers were down on the 2019 numbers in all event types, further showing a coinciding of the tweet activity decline and the first pandemic wave in Europe. These results are in agreement with previous research, which has shown that especially global events such as economic crises lead the public to assign climate change a low priority (Ruiz et al., 2020: 114; Whitmarsh, 2011: 691). With a potentially increasing frequency and intensity of global crises in the future (Biggs et al., 2011: 1), there is a danger that climate change will continuously be overshadowed by such events, especially when they increasingly begin to overlap each other.

5.4 Reflection on methods

Throughout this thesis, a wide range of methodology has been applied to collect and analyse Twitter data in the context of climate change. This section will serve as a reflection on the chosen approaches and discuss their limitations as well as their achievements. As the methodological limitations faced during each step of the analysis have already been detailed in the corresponding sections of chapter 3, this chapter will shortly summarise the main limitations and contextualise them with the results.

5.4.1 Data retrieval and processing

Queries build the foundation not only for the information retrieval, but subsequently also for all findings derived from the resulting data. Simultaneously, query design requires certain compromises to match the scope of a project (Manning et al., 2008: 1; Parsons et al., 2015: 1222). Queries are therefore crucial for my thesis and were carefully crafted using previous literature alongside thesauri and dictionaries as a basis for keywords (e.g., Ostermann & Spinsanti (2012: 22) for wildfires).

A second important phase is the classification of tweets once they had been retrieved, which is made difficult by their short and comprised nature (Ostermann & Spinsanti, 2012: 23). This step required extensive testing and developed from a rigid rule-based algorithm to the more dynamic Random Forest Classifier. Here, the selection of parameters used in the algorithm's decision-making is key, as they not only impact accuracy but can also introduce biases when applied beyond the testing data. Although proving very effective in increasing precision when the training data covers the entirety of the study period at reasonable temporal intervals, the date parameter develops a strong bias when the classifier is used for data outside its (temporal) training range. As such, a loss in accuracy had to be accepted in order to remove the bias. Using tweets from the entire study period as training data would have been difficult as a very large number of tweets would have been needed to be annotated to maintain a decent temporal resolution. Furthermore, each event-language combination requires its own model (i.e., 20 models in total)[20]. In the end, F-scores between 0.8 and 0.85 were targeted for STETS (which contained more ambiguous and off-topic tweets) and values between 0.85 and 0.9 for LTETs.

Problems and limitations of geographic information retrieval, especially when dealing with user-generated content such as tweets, are aplenty: Aspects such as vagueness and ambiguity characterize both the thematic content of the tweets as well as the spatial language used on Twitter, the latter being known as toponym ambiguity (Purves et al., 2018: 214). This impacts, for example, the granularity achievable for my analysis: While country-level toponym recognition and resolution are fairly reliable, the state of the art in geoparsing poses a limiting factor for events occurring on local scales as issues such as toponym ambiguity increasingly have more of an impact (Brunner & Purves, 2008: 26).

Whereas the accuracy of toponym recognition was largely limited by the abilities of the NER library, toponym resolution was considerably improved by developing a grouping approach inspired by de Bruijn et al., (2018: 2), which uses toponyms occurring alongside each other as contextual information when deciding which location to assign to a placename. The result is a fairly good F-score for the resolution at 0.9, which remedies some of the shortcomings of the NER used for recognition (F-score 0.78).

Somewhat related, the approximation of Twitter penetration in each country (for each year of the study period) poses one of the larger uncertainties. Although entirely based on relevant data, the set of assumptions and interpolations necessary to arrive at the estimates likely introduced inaccuracies. Furthermore could only a national value be achieved and subnational penetration had to be scaled by population, disregarding urban-rural differences and other regional characteristics. Nonetheless, the spatial and temporal results affected by this measure can in some cases be validated as they seem to exhibit trends that agree with previous literature or environmental factors.

[20] Models typically required between 300 and 600 annotated training tweets (plus another 100 to 200 for testing) to achieve satisfactory results.

The final step of clustering events from the tweets was carried out using an empirical method that allowed for the fine-tuning of outcomes based on a single parameter. Despite its simplicity, the approach successfully delineated STET events, which were shown to align well with real-world occurrences in the results.

5.4.2 Data analysis

Uncertainties in the data analysis are largely carried over from the retrieval and processing steps. On the thematic side, for example, uncertainties result from the ambiguity in natural language, which are exacerbated by the need for automatic classification and categorisation due to the large corpus size. The extraction of themes is one the analytical approaches that had to be automated. Although topic modelling is frequently used in research dealing with Twitter data (e.g., Bauer et al., 2012; Resch et al., 2018), the results were hardly conclusive and interpretation was further hindered by the impossibility to understand the decision-making process of topic modelling algorithms such as LDA due to their ‘black box’ nature. The final keyword-based approach proved much more concise and meaningful, however was considerable human interference introduced as a result of annotating and categorising keywords. Compared to methods such as LDA, however, this influence is manageable as the annotations determining the outcomes are known and can be taken into consideration when interpreting the results.

The spatial analyses are mostly dependent on above-mentioned issues when it comes to uncertainties and limitations. Importantly, data was aggregated on the national level for a large majority of the analyses to address some of these challenges. In the case of the mapped spatial distributions, where subnational divisions were used, inferences about highly local trends were only made when reasonable explanations were available, with the remaining conclusion being drawn from combined values of larger areas. Transnationality once again is majorly influenced by the geoparsing accuracy.

Regarding the influence of environmental factors, the contextual data was aggregated to best match the corresponding event types on the temporal scale, with the different spatial extents serving as a base for comparison. Correlations here certainly must be taken with a grain of salt as causation is not given by any means, especially in LTETs due to the strong influence of news media and the lack of direct connection between the public and the processes. Nevertheless, some investigations brought forward promising results as for example the distribution of regional temperature thresholds for heatwaves from tweets align remarkably well with the scientific definition of the threshold. This not only validates the specific methodology of this section, but also indicates that the data collection and processing likely produced a relevant and meaningful dataset.

5.4.3 Ethical considerations

Besides the complex nature of working with tweet-derived data, ethical considerations must be contemplated, not least to keep this thesis from simply becoming a large-scale data fishing exercise. The goal of creating a better understanding regarding climate change perception aimed to give my thesis a meaningful, real-world application. The findings indicate that interesting contributions were able to be made by my thesis.

Further ethical aspects affect the handling of the collected information. Although all the data about a tweet or user returned by the API is also visible to the public when reading the post or looking at a user’s profile, no individual person should be identifiable in the presentation of the results or any other part of this thesis. Therefore, data was aggregated over large areas and presented anonymously. Even on the aggregated level, the classification of people (i.e., into income classes based on jobs derived from user profiles) was avoided in this thesis. For the same reason, examples of individual tweets were used very sparingly and only when they were not uniquely identifiable (i.e., sharing a news headline).

6. Conclusion

This thesis applied Geographic Information Retrieval, Natural Language Processing, and spatiotemporal analysis methodologies to derive information from social media data on how the public interacts with specific climate change effects and processes. It has shown that Twitter can be used as a data source to gain insight into such perception characteristics, which can be helpful in making informed decisions regarding climate policies and the populations they target. A framework was developed that went beyond the commonly carried out event detection from social media and investigated the tweets associated with each event more deeply on multiple levels. Such comprehensive long-term studies taking into account various dimensions and individual climate change effects have up until now been rare at this scale to the best of my knowledge.

Public reactions to five individual climate change effects representing different environmental systems were analysed by investing corresponding corpora of tweets spanning the timespan of a decade and the spatial extent of western Europe. The investigation determined thematic trends from tweet content, spatial distributions and influences of environmental factors, as well temporal distributions and impacts of real-world events.

6.1 Implications of main findings

Summarising all results, the main finding of this thesis states that the public's reactions to and perceptions of climate change do in fact differ tremendously in the thematic, spatial, and temporal dimensions. This has several implications: Future research on the public-climate interface needs to be aware of this multidimensional diversity that lies beneath the umbrella term. For example, spatial differences in behaviour patterns related to climate change are not always indicative of the people themselves but could result from the types of environmental processes and events they are exposed to. Additionally has the long-term nature of this thesis uncovered that trends in any of the three dimensions can vary strongly within the span of a decade. This needs to be taken into account when interpreting public reaction data spanning shorter time periods. Especially with major events such as the ongoing COVID-19 pandemic, public concern about climate change can drop significantly and subsequently skew data. This builds on previous research suggesting these temporal dependencies in regards to economic crises and priority-setting (Semenza et al., 2008: 483; Whitmarsh, 2011: 691) as well as seasonal influences on concern about climate change (Howe et al., 2012: 3).

Decision- and policymakers of the climate action field further need to consider this strong variability when implementing mitigation and adaptation measures for their local constituents. The spatial variations in my results suggest that solutions successful in one country will not necessarily succeed in another due to different interests and concerns, the same even applying for different regions within countries. Similarly, interests and concerns evolve over time, meaning that the public could develop diverging stances on the same policy in different years. Finally, people have shown increased interest in processes and events relevant to their national or local sphere (e.g., glaciers in alpine countries), which would suggest that policies following along similar lines and addressing such tangible issues could profit from higher public agreement.

Regarding the driving forces behind the observed findings, this thesis has shown that traditional news media have a rather dominant grasp on the varying discourses on Twitter, especially when it comes to scientific discussions rather than statements of personal experience. This implies that direct science communication on Twitter is severely lacking. The academic community is largely at the mercy of news media when it comes to delivering research to the masses in the online sphere, surrendering the power to select and frame findings in a light that best suits their agenda. In turn, news media can serve as a powerful tool to reach and inform the public about climate change on Twitter and beyond.

6.2 Future work

While the analyses of this thesis have provided answers to the research question set out in the beginning, many more avenues for potential research were exposed. The results presented in chapter 4 merely scratch the surface of the sea of data that can be generated from such a corpus of tweets. Endless combinations of factors and variables from the three dimensions could offer detailed insights into a plethora of connections, creating an even deeper understanding of how the public interacts with climate change. The same corpus could furthermore be investigated with a wide array of additional methodologies that went unused in this thesis such as for example sentiment analysis or a range of additional spatial statistics. The further potential for interdisciplinary research using such a corpus, both within and beyond geography, is of course substantial as well.

Furthermore, questions have been raised about the at times diverging public reactions between general climate change tweets and the event-specific tweets from my thesis, which should further be investigated to determine how representative climate change is of its various processes and vice versa in the domain of public perception. Of special interest is additionally the impact of the COVID-19 pandemic, both in the short and long-term. Significant decreases in the public's attention to climate change processes were already observable in 2020 and it will be important to monitor how this trend develops through later stages and beyond the pandemic as a dwindling concern about the issue could lead to severe ramifications for the implementation of climate policies. Not least, this should further provide a solid knowledge base for the development of the public-climate interface under future global crises.

7. References

Due to the high number of references in sections 4.1.4 and 4.3.4 used for demonstration purposes only, this listing of references is split into two parts. First, a regular overview of references is given, before the references from the above-mentioned sections that did not serve any information-sourcing purpose are listed separately

7.1 Regular References

350.org (**2019**). Over six million people demand action after week of climate strikes, available at: https://350.org/press-release/over-six-million-people-demand-action-after-week-of-climate-strikes/ (accessed 8 December 2021).

Abbar, S., Zanouda, T., Berti-Equille, L. & Borge-Holthoefer, J. (**2016**). Using twitter to understand public interest in climate change: The case of Qatar. *International Workshop on Social Media for Environment and Ecological Monitoring* pp. 1–10.

Acheson, E., De Sabbata, S. & Purves, R.S. (**2017**). A quantitative analysis of global gazetteers: Patterns of coverage for common feature types. *Computers, Environment and Urban Systems* 64, pp. 309–320.

Alman, B.L., Pfister, G., Hao, H., Stowell, J., Hu, X., Liu, Y. & Strickland, M.J. (**2016**). The association of wildfire smoke with respiratory and cardiovascular emergency department visits in Colorado in 2012: a case crossover study. *Environmental Health* 15 (64), pp. 1–9.

Alshaabi, T., Adams, J.L., Arnold, M. V., Minot, J.R., Dewhurst, D.R., Reagan, A.J., Danforth, C.M., et al. (**2021**). Storywrangler: A massive exploratorium for sociolinguistic, cultural, socioeconomic, and political timelines using Twitter. *Science Advances* 7 (29), pp. 1–14.

Anisimov, O.A. (**2007**). Potential feedback of thawing permafrost to the global climate system through methane emission. *Environmental Research Letters* 2 (4), pp. 1–7.

Annoni, A., Luzet, C., Gubler, E., Ihde, J., Augusto, G., Brandenberger, C., Dhee, F., et al. (**2001**). *Map Projections for Europe*. eds. Annoni, A., Luzet, C., Gubler, E. and Ihde, J. European Comission Joint Research Centre.

Artés Vivancos, T., Oom, D., de Rigo, D., Durrant, T.H., Maianti, P., Libertà, G. & San-Miguel-Ayanz, J. (**2019**). A global wildfire dataset for the analysis of fire regimes and fire behaviour. *Scientific Data* 6 (1), pp. 1–12.

Artés Vivancos, T. & San-Miguel-Ayanz, J. (**2018**). Global Wildfire Database for GWIS. *PANGAEA*.

Atefeh, F. & Khreich, W. (**2015**). A survey of techniques for event detection in Twitter. *Computational Intelligence* 31 (1), pp. 132–164.

Barclay, E. & Amaria, K. (**2019**). Photos: kids in 123 countries went on strike to protect the climate. *Vox*, available at: https://www.vox.com/energy-and-environment/2019/3/15/18267156/youth-climate-strike-march-15-photos (accessed 8 December 2021).

Bauer, S., Noulas, A., Seaghdha, D.O., Clark, S. & Mascolo, C. (**2012**). Talking places: Modelling and analysing linguistic content in Foursquare. *2012 ASE/IEEE International Conference on Social Computing and 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust* pp. 348–357.

Beniston, M., Farinotti, D., Stoffel, M., Andreassen, L.M., Coppola, E., Eckert, N., Fantini, A., et al. (**2018**). The European mountain cryosphere: a review of its current state, trends, and future challenges. *The Cryosphere* 12 (2), pp. 759–794.

Berkely Earth (**2021**). Data Overview. *Berkely Earth*, available at: http://berkeleyearth.org/data/ (accessed 1 December 2021).

Bevacqua, E., Maraun, D., Vousdoukas, M.I., Voukouvalas, E., Vrac, M., Mentaschi, L. & Widmann, M. (**2019**). Higher probability of compound flooding from precipitation and storm surge in Europe under anthropogenic climate change. *Science Advances* 5 (9), pp. 1–7.

Biggs, D., Biggs, R. (Oonsie), Dakos, V., Scholes, R.J. & Schoon, M. (**2011**). Are We Entering an Era of Concatenated Global Crises? *Ecology and Society* 16 (2), pp. 1–10.

Black, C., Tesfaigzi, Y., Bassein, J.A. & Miller, L.A. (**2017**). Wildfire smoke exposure and human health: Significant gaps in research for a growing public health issue. *Environmental Toxicology and Pharmacology* 55, pp. 186–195.

BLE (**2021**). Waldbrandstatistik. *Bundesamt Für Landwirtschaft Und Ernährung*, available at: https://www.ble.de/DE/BZL/Daten-Berichte/Wald/wald_node.html (accessed 21 December 2021).

Blei, D.M., Ng, A.Y. & Jordan, M.I. (**2003**). Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3, pp. 993–1022.

Bojinski, S., Verstraete, M., Peterson, T.C., Richter, C., Simmons, A. & Zemp, M. (**2014**). The concept of essential climate variables in support of climate research, applications, and policy. *Bulletin of the American Meteorological Society* 95 (9), pp. 1431–1443.

BOKU (**2020**). Waldbrand-Statistik. *Universität Für Bodenkultur Wien*, available at: https://boku.ac.at/wabo/waldbau/forschung/themen/bewirtschaftungskonzepte/waldbewirtschaftung-und-klimaaenderung/waldbrand/waldbrand-statistik (accessed 21 December 2021).

Borick, C.P. & Rabe, B.G. (**2010**). A reason to believe: Examining the factors that determine individual views on global warming. *Social Science Quarterly* 91 (3), pp. 777–800.

Boudet, H., Giordono, L., Zanocco, C., Satein, H. & Whitley, H. (**2020**). Event attribution and partisanship shape local discussion of climate change after extreme weather. *Nature Climate Change* 10 (1), pp. 69–76.

Boulianne, S., Lalancette, M. & Ilkiw, D. (**2020**). "School Strike 4 Climate": Social Media and the International Youth Protest on Climate Change. *Media and Communication* 8 (2), pp. 208–218.

Bowman, D.M.J.S., Kolden, C.A., Abatzoglou, J.T., Johnston, F.H., van der Werf, G.R. & Flannigan, M. (**2020**). Vegetation fires in the Anthropocene. *Nature Reviews Earth and Environment* 1, pp. 500–515.

Brossard, D. & Scheufele, D.A. (**2013**). Science, new media, and the public. *Science* 339 (6115), pp. 40–41.

de Bruijn, J.A., de Moel, H., Jongman, B., de Ruiter, M.C., Wagemaker, J. & Aerts, J.C.J.H. (**2019**). A global database of historic and real-time flood events based on social media. *Scientific Data* 6 (311), pp. 1–12.

de Bruijn, J.A., de Moel, H., Jongman, B., Wagemaker, J. & Aerts, J.C.J.H. (**2018**). TAGGS: Grouping Tweets to Improve Global Geotagging for Disaster Response. *Journal of Geovisualization and Spatial Analysis* 2 (2), pp. 1–14.

Brunner, T.J. & Purves, R.S. (**2008**). Spatial autocorrelation and toponym ambiguity. *Proceedings of the 5th Workshop on Geographic Information Retrieval*, pp. 25–26.

Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J. & Wara, M. (**2021**). The changing risk and burden of wildfire in the United States. *Proceedings of the National Academy of Sciences* 118 (2), pp. 1–6.

Carroll, M.W. (**2011**). Why full open access matters. *PLoS Biology* 9 (11), pp. 1–3.

Chas-Amil, M.-L., García-Martínez, E. & Touza, J. (**2020**). Iberian Peninsula October 2017 wildfires: Burned area and population exposure in Galicia (NW of Spain). *International Journal of Disaster Risk Reduction* 48, pp. 1–12.

Church, J.A. & White, N.J. (**2011**). Sea-Level Rise from the Late 19th to the Early 21st Century. *Surveys in Geophysics* 32, pp. 585–602.

Dahal, B., Kumar, S.A.P. & Li, Z. (**2019**). Topic modeling and sentiment analysis of global climate change tweets. *Social Network Analysis and Mining* 9 (24), pp. 1–20.

DataReportal (**2021**). Complete Report Library, available at: https://datareportal.com/library (accessed 1 November 2021).

Davis, J. & Goadrich, M. (**2006**). The Relationship Between Precision-Recall and ROC Curves. *Proceedings of the 23rd International Conference on Machine Learning* pp. 233–240.

Doerr, S.H. & Santín, C. (**2016**). Global trends in wildfire and its impacts: perceptions versus realities in a changing world. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371, pp. 1–10.

Dong, T.T.T., Hinwood, A.L., Callan, A.C., Zosky, G. & Stock, W.D. (**2017**). In vitro assessment of the toxicity of bushfire emissions: A review. *Science of The Total Environment* 603, pp. 268–278.

Duc, H.N., Chang, L.T.-C., Azzi, M. & Jiang, N. (**2018**). Smoke aerosols dispersion and transport from the 2013 New South Wales (Australia) bushfires. *Environmental Monitoring and Assessment* 190 (428), pp. 1–22.

Dunkel, A., Andrienko, G., Andrienko, N., Burghardt, D., Hauthal, E. & Purves, R. (**2019**). A conceptual framework for studying collective reactions to events in location-based social media. *International Journal of Geographical Information Science* Taylor & Francis, 33 (4), pp. 780–804.

Eriksen, C. (**2020**). Europe's Fiery Future: Rethinking Wildfire Policy. *Policy Perspectives* 8 (12), pp. 1–4.

Farquharson, L.M., Romanovsk, V.E., Cable, W.L., Walker, D.A., Kokelj, S. V. & Nicolsky, D. (**2019**). Climate change drives widespread and rapid thermokarst development in very cold permafrost in the Canadian High Arctic. *Geophysical Research Letters* 46, pp. 6681–6689.

Fecher, B. & Friesike, S. (**2014**). Open science: one term, five schools of thought. in Bartling, S. and Friesike, S. (Eds.), *Opening Science* Springer, Cham, pp. 17–47.

Finlay, S.E., Moffat, A., Gazzard, R., Baker, D. & Murray, V. (**2012**). Health Impacts of Wildfires. *PLoS Currents* 4.

Frondel, M., Simora, M. & Sommer, S. (**2017**). Risk Perception of Climate Change: Empirical Evidence for Germany. *Ecological Economics* 137, pp. 173–183.

Gardner, A.S., Moholdt, G., Cogley, J.G., Wouters, B., Arendt, A.A., Wahr, J., Berthier, E., et al. (**2013**). A Reconciled Estimate of Glacier Contributions to Sea Level Rise: 2003 to 2009. *Science* 340 (6134), pp. 852–857.

Gautam, S. & Hens, L. (**2020**). COVID-19: impact by and on the environment, health and economy. *Environment, Development and Sustainability* 22, pp. 4953–4954.

GeoNames (**2021**). GeoNames, available at: https://www.geonames.org/ (accessed 1 November 2021).

Gibson, C.M., Chasmer, L.E., Thompson, D.K., Quinton, W.L., Flannigan, M.D. & Olefeldt, D. (**2018**). Wildfire as a major driver of recent permafrost thaw in boreal peatlands. *Nature Communications* 9 (3041), pp. 1–9.

Gimenes, M. & New, B. (**2015**). Worldlex: Twitter and blog word frequencies for 66 languages. *Behavior Research Methods* pp. 1–10.

Goodchild, M.F. (**2007**). Citizens as sensors: The world of volunteered geography. *GeoJournal* 69 (4), pp. 211–221.

Goodchild, M.F. & Li, L. (**2012**). Assuring the quality of volunteered geographic information. *Spatial Statistics* 1, pp. 110–120.

Haines, A., Scheelbeek, P. & Abbasi, K. (**2020**). The health case for urgent action on climate change. *BMJ* 368, pp. 1–2.

Hamilton, L.C. & Keim, B.D. (**2009**). Regional variation in perceptions about climate change. *International Journal of Climatology* 29, pp. 2348–2352.

Hansen, J., Sato, M. & Ruedy, R. (**2012**). Perception of climate change. *Proceedings of the National Academy of Sciences of the United States of America* 109 (37), pp. E2415–E2423.

Hauer, M.E. (**2017**). Migration induced by sea-level rise could reshape the US population landscape. *Nature Climate Change* 7, pp. 321–325.

Häussler, T., Schmid-Petri, H., Adam, S., Reber, U. & Arlt, D. (**2016**). The climate of debate: How institutional factors shape legislative discourses on climate change. *Studies in Communication Sciences* 16 (1), pp. 94–102.

Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P. & Ratti, C. (**2014**). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science* 41 (3), pp. 260–271.

Hepburn, C., O'Callaghan, B., Stern, N., Stiglitz, J. & Zenghelis, D. (**2020**). Will COVID-19 fiscal recovery packages accelerate or retard progress on climate change? *Oxford Review of Economic Policy* 36 (S1), pp. S359–S381.

Hollenstein, L. & Purves, R.S. (**2010**). Exploring place through user-generated content: Using Flickr tags to describe city cores. *Journal of Spatial Information Science* 1 (2010), pp. 21–48.

Holmberg, K. & Hellsten, I. (**2015**). Gender differences in the climate change communication on Twitter. *Internet Research* 25 (5), pp. 811–828.

Hopke, J.E. (**2020**). Connecting Extreme Heat Events to Climate Change: Media Coverage of Heat Waves and Wildfires. *Environmental Communication* 14 (4), pp. 492–508.

Howe, P.D., Markowitz, E.M., Lee, T.M., Ko, C.Y. & Leiserowitz, A. (**2012**). Global perceptions of local temperature change. *Nature Climate Change* 3 (4), pp. 352–356.

Hu, Y. (**2018**). Geo-text data and data-driven geospatial semantics. *Geography Compass* 12 (11), pp. 1–19.

van Huissteden, J. (**2020**). *Thawing Permafrost: Permafrost Carbon in a Warming Arctic.* Springer, Cham.

Huq, S., Ali, S.I. & Rahman, A.A. (**1995**). Sea-Level Rise and Bangladesh: A Preliminary Analysis. *Journal of Coastal Research* Special Is, pp. 44–53.

Huss, M., Bauder, A. & Funk, M. (**2009**). Homogenization of long-term mass-balance time series. *Annals of Glaciology* 50 (50), pp. 198–206.

IPCC (**2013**). Summary for Policymakers. in WorkingGroupI, Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., et al. (Eds.), *Climate Change 2013: The Physical Science Basis* Cambridge University Press, Cambridge, available at:https://doi.org/10.1260/095830507781076194.

IPCC (**2014**). *Climate Change 2014: Synthesis Report*. eds. CoreWritingTeam, Pachauri, R.K. and Meyer, L.A. IPCC, Geneva, Switzerland.

IPCC (**2019**). *Special Report on the Ocean and Cryosphere in a Changing Climate*. eds. Pörtner, H.-O., Roberts, D.C., Masson-Delmotte, V., Zhai, P., Tignor, M., Poloczanska, E., Mintenbeck, K., et al.

Javadinejad, S., Dara, R. & Jafary, F. (**2019**). Taking Urgent Actions to Combat Climate Change Impacts. *Annals of Geographical Studies* 2 (4), pp. 1–13.

Johnson, I., McMahon, C., Schöning, J. & Hecht, B. (**2017**). The Effect of Population and "Structural" Biases on Social Media-based Algorithms – A Case Study in Geolocation Inference Across the Urban-Rural Spectrum. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* pp. 1167–1178.

Jorgenson, M.T., Romanovsky, V., Harden, J., Shur, Y., O'Donnell, J., Schuur, E.A.G., Kanevskiy, M., et al. (**2010**). Resilience and vulnerability of permafrost to climate change. *Canadian Journal of Forest Research* 40 (7), pp. 1219–1236.

Kääb, A., Wessels, R., Haeberli, W., Huggel, C., Kargel, J.S. & Khalsa, S.J.S. (**2003**). Rapid ASTER Imaging Facilitates Timely Assessment of Glacier Hazards and Disasters. *Eos* 84 (13), pp. 117–124.

Karl, T.R., Nicholls, N. & Ghazi, A. (**1999**). CLIVAR/GCOS/WMO Workshop on Indices and Indicators for Climate Extremes Workshop Summary. in Karl, T.R., Nicholls, N. and Ghazi, A. (Eds.), *Weather and Climate Extremes* Springer, Dordrecht, pp. 3–7.

Kaser, G., Großhauser, M. & Marzeion, B. (**2010**). Contribution potential of glaciers to water availability in different climate regimes. *Proceedings of the National Academy of Sciences* 107 (47), pp. 20223–20227.

Kerr, R.A. (**2007**). How Urgent Is Climate Change? *Science* 318 (5854), pp. 1230–1231.

Kirdyanov, A. V., Saurer, M., Siegwolf, R., Knorre, A.A., Prokushkin, A.S., Churakova, O. V., Fonti, M. V., et al. (**2020**). Long-term ecological consequences of forest fires in the continuous permafrost zone of Siberia. *Environmental Research Letters* IOP Publishing, 15 (3), pp. 1–11.

Kirezci, E., Young, I.R., Ranasinghe, R., Muis, S., Nicholls, R.J., Lincke, D. & Hinkel, J. (**2020**). Projections of global-scale extreme sea levels and resulting episodic coastal flooding over the 21st Century. *Scientific Reports* 10, pp. 1–12.

Kirilenko, A.P. & Stepchenkova, S.O. (**2014**). Public microblogging on climate change: One year of Twitter worldwide. *Global Environmental Change* 26, pp. 171–182.

Kjellström, E., Bärring, L., Jacob, D., Jones, R., Lenderink, G. & Schär, C. (**2007**). Modelling daily temperature extremes: recent climate and future changes over Europe. *Climate Change* 81, pp. 249–265.

Kleerekoper, L., van Esch, M. & Salcedo, T.B. (**2012**). How to make a city climate-proof, addressing the urban heat island effect. *Resources, Conservation and Recycling* 64, pp. 30–38.

Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., et al. (**2002**). Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology* 22 (12), pp. 1441–1453.

Van Koningsveld, M., Mulder, J.P.M., Stive, M.J.F., Van Der Valk, L. & Van Der Weck, A.W. (**2008**). Living with Sea-Level Rise and Climate Change: A Case Study of the Netherlands. *Journal of Coastal Research* 242, pp. 367–379.

Krackhardt, D. & Stern, R.N. (**1988**). Informal Networks and Organizational Crises: An Experimental Simulation. *Social Psychology Quarterly* 51 (2), pp. 123–140.

Krippendorff, K. (**2018**). *Content Analysis. An Introduction to Its Methodology*. Sage Publications, Thousand Oaks, CA.

Lansley, G. & Longley, P.A. (**2016**). The geography of Twitter topics in London. *Computers, Environment and Urban Systems* 58, pp. 85–96.

Laux, T. (**2021**). What makes a global movement? Analyzing the conditions for strong participation in the climate strike. *Social Science Information* 60 (3), pp. 413–435.

Lee, K., Ganti, R., Srivatsa, M. & Mohapatra, P. (**2013**). Spatio-temporal provenance: Identifying location information from unstructured text. *2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)* pp. 499–504.

Lee, T.M., Markowitz, E.M., Howe, P.D., Ko, C.Y. & Leiserowitz, A.A. (**2015**). Predictors of public climate change awareness and risk perception around the world. *Nature Climate Change* 5 (11), pp. 1014–1020.

Legendre, M., Lartigue, A., Bertaux, L., Jeudy, S., Bartoli, J., Lescot, M., Alempic, J.-M., et al. (**2015**). In-depth study of Mollivirus sibericum, a new 30,000-y-old giant virus infecting Acanthamoeba. *Proceedings of the National Academy of Sciences* 112 (38), pp. E5327–E5335.

Leverkus, A.B., Thorn, S., Lindenmayer, D.B. & Pausas, J.G. (**2020**). Wildfire debate needs science, not politics. *Science* 370 (6515), pp. 416–417.

Li, T.J.J., Sen, S. & Hecht, B. (**2014**). Leveraging advances in natural language processing to better understand Tobler's first law of geography. *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* pp. 513–516.

De Longueville, B., Ostländer, N. & Keskitalo, C. (**2010**). Addressing vagueness in Volunteered Geographic Information (VGI) – A case study. *International Journal of Spatial Data Infrastructures Research* 5, pp. 1725–0463.

López-Goñi, I. & Sánchez-Angulo, M. (**2018**). Social networks as a tool for science communication and public engagement: Focus on Twitter. *FEMS Microbiology Letters* 365 (2), pp. 1–4.

Lorenzoni, I., Nicholson-Cole, S. & Whitmarsh, L. (**2007**). Barriers perceived to engaging with climate change among the UK public and their policy implications. *Global Environmental Change* 17 (3–4), pp. 445–459.

MacEachren, A.M., Jaiswal, A., Robinso;, A.C., Pezanowski, S., Zhang, X., Savelyev, A., Mitra, P., et al. (**2011**). SensePlace2: GeoTwitter analytics support for situational awareness. *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)* pp. 181–190.

Manning, C.D., Raghavan, P. & Schütze, H. (**2008**). *Introduction to Information Retrieval*. Vol. 39 Cambridge University Press, Cambridge.

Marcos, M., Rohmer, J., Vousdoukas, M.I., Mentaschi, L., Le Cozannet, G. & Amores, A. (**2019**). Increased Extreme Coastal Water Levels Due to the Combined Action of Storm Surges and Wind Waves. *Geophysical Research Letters* 46 (8), pp. 4356–4364.

McGranahan, G., Balk, D. & Anderson, B. (**2007**). The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and Urbanization* 19 (1), pp. 17–37.

Meo, S.A., Abukhalaf, A.A., Alomar, A.A. & Alessa, O.M. (**2020**). Wildfire and COVID-19 pandemic: effect of environmental pollution PM-2.5 and carbon monoxide on the dynamics of daily cases and deaths due to SARS-COV-2 infection in San-Francisco USA. *European Review for Medical and Pharmacological Sciences* 24 (19), pp. 10286–10292.

Metag, J., Schäfer, M.S., Füchslin, T., Barsuhn, T. & Kleinen-von Königslöw, K. (**2016**). Perceptions of Climate Change Imagery: Evoked Salience and Self-Efficacy in Germany, Switzerland, and Austria. *Science Communication* 38 (2), pp. 197–227.

Milne, G.A., Gehrels, W.R., Hughes, C.W. & Tamisiea, M.E. (**2009**). Identifying the causes of sea-level change. *Nature Geoscience* Nature Publishing Group, 2 (7), pp. 471–478.

Miner, K.R., D'Andrilli, J., Mackelprang, R., Edwards, A., Malaska, M.J., Waldrop, M.P. & Miller, C.E. (**2021**). Emergent biogeochemical risks from Arctic permafrost degradation. *Nature Climate Change* 11, pp. 809–819.

Minot, J.R., Arnold, M. V., Alshaabi, T., Danforth, C.M. & Dodds, P.S. (**2021**). Ratioing the President: An exploration of public engagement with Obama and Trump on Twitter. *PLoS ONE* 16 (4), p. e0248880.

Mitchell, L., Frank, M.R., Harris, K.D., Dodds, P.S. & Danforth, C.M. (**2013**). The Geography of Happiness: Connecting Twitter Sentiment and Expression, Demographics, and Objective Characteristics of Place. *PLoS ONE* 8 (5), pp. 1–15.

Moernaut, R., Mast, J., Temmerman, M. & Broersma, M. (**2020**). Hot weather, hot topic. Polarization and sceptical framing in the climate debate on Twitter. *Information Communication and Society* pp. 1–20.

Murthy, D. (**2011**). Twitter: Microphone for the masses? *Media, Culture & Society* 33 (5), pp. 779–789.

Nairn, J.R. & Fawcett, R.J.B. (**2014**). The excess heat factor: A metric for heatwave intensity and its use in classifying heatwave severity. *International Journal of Environmental Research and Public Health* 12, pp. 227–253.

Nelson, F.E., Anisimov, O.A. & Shiklomanov, N.I. (**2002**). Climate change and hazard zonation in the circum-arctic permafrost regions. *Natural Hazards* 26 (3), pp. 203–225.

Newell, P. & Taylor, O. (**2020**). Fiddling while the planet burns? COP25 in perspective. *Globalizations* 17 (4), pp. 580–592.

Nghiem, S. V., Hall, D.K., Mote, T.L., Tedesco, M., Albert, M.R., Keegan, K., Shuman, C.A., et al. (**2012**). The extreme melt across the Greenland ice sheet in 2012. *Geophysical Research Letters* 39 (20), pp. 1–6.

Nicholls, R.J. & Cazenave, A. (**2010**). Sea-level rise and its impact on coastal zones. *Science* 328 (5985), pp. 1517–1520.

Nicholls, R.J., Wong, P.P., Burkett, V., Woodroffe, C.D. & Hay, J. (**2008**). Climate change and coastal vulnerability assessment: scenarios for integrated assessment. *Sustainability Science* 3, pp. 89–102.

NIFC (**2012**). National Report of Wildland Fires and Acres Burned by State 2012. *National Interagency Fire Center*. Available at: https://www.predictiveservices.nifc.gov/intelligence/2012_statssumm/fires_acres.pdf (accessed 24 December 2021).

NOAA Laboratory for Satellite Altimetry (**2020**). Global sea level time series, available at: https://www.star.nesdis.noaa.gov/socd/lsa/SeaLevelRise/LSA_SLR_timeseries.php (accessed 1 November 2021).

Nothman, J., Ringland, N., Radford, W., Murphy, T. & Curran, J.R. (**2013**). Learning multilingual named entity recognition from Wikipedia. *Artificial Intelligence* 194, pp. 151–175.

O'Connor, F.M., Boucher, O., Gedney, N., Jones, C.D., Folberth, G.A., Coppell, R., Friedlingstein, P., et al. (**2010**). Possible role of wetlands, permafrost, and methane hydrates in the methane cycle under future climate change: A review. *Reviews of Geophysics* 48 (4), pp. 1–33.

O'Connor, R.E., Bard, R.J. & Fisher, A. (**1999**). Risk Perceptions, General Environmental Beliefs, and Willingness to Address Climate Change. *Risk Analysis* 19 (3), pp. 461–471.

Obu, J., Westermann, S., Barboux, C., Bartsch, A., Delaloye, R., Grosse, G., Heim, B., et al. (**2021**). ESA Permafrost Climate Change Initiative (Permafrost_cci): Permafrost extent for the Northern Hemisphere, v3.0. NERC EDS Centre for Environmental Data Analysis. *CEDA Archive*.

OECD (**2012**). *Redefining "Urban": A New Way to Measure Metropolitan Areas* OECD Publishing, Paris.

OECD (**2021a**). Functional urban areas by country, available at: https://www.oecd.org/regional/regional-statistics/functional-urban-areas.htm (accessed 23 November 2021).

OECD (**2021b**). Regions and Cities: Metropolitan areas, available at: https://stats.oecd.org/Index.aspx?Datasetcode=CITIES# (accessed 23 November 2021).

Office for National Statistics (**2021**). Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland, available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland (accessed 1 November 2021).

Ostermann, F. & Spinsanti, L. (**2012**). Context Analysis of Volunteered Geographic Information from Social Media Networks to Support Disaster Management. *International Journal of Information Systems for Crisis Response and Management* 4 (4), pp. 16–37.

Ostermann, F.O., Huang, H., Andrienko, G., Andrienko, N., Capineri, C., Farkas, K. & Purves, R.S. (**2015**). Extracting and comparing places using geo-social media. *ISPRS GEOSPATIAL WEEK 2015* II–3 (W5), pp. 311–316.

Parkinson, C.L. & Comiso, J.C. (**2013**). On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm. *Geophysical Research Letters* 40 (7), pp. 1356–1361.

Parsons, S., Atkinson, P.M., Simperl, E. & Weal, M. (**2015**). Thematically analysing social network content during disasters through the lens of the disaster management lifecycle. *WWW 2015 Companion - Proceedings of the 24th International Conference on World Wide Web* pp. 1221–1226.

Paul, F., Kääb, A., Maisch, M., Kellenberger, T. & Haeberli, W. (**2004**). Rapid disintegration of Alpine glaciers observed with satellite data. *Geophysical Research Letters* 31, pp. 1–4.

Pearce, W., Holmberg, K., Hellsten, I. & Nerlich, B. (**2014**). Climate change on twitter: Topics, communities and conversations about the 2013 IPCC Working Group 1 report. *PLoS ONE* 9 (4), pp. 1–11.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., et al. (**2011**). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, pp. 2825–2830.

Perrin, A. (**2015**). *Social Media Usage: 2005-2015*. Pew Research Center.

Pezza, A.B., van Rensch, P. & Cai, W. (**2012**). Severe heat waves in Southern Australia: Synoptic climatology and large scale connections. *Climate Dynamics* 38 (1–2), pp. 209–224.

Phelps, L., Fox, B.A. & Marincola, F.M. (**2012**). Supporting the advancement of science: Open access publishing and the role of mandates. *Journal of Translational Medicine* 10 (13), pp. 1–3.

Pluess, A.R., Augustin, S. & Brang, P. (**2016**). *Wald Im Klimawandel. Grundlagen Für Adaptationsstrategien*. Haupt Verlag, Bern.

Purves, R.S., Clough, P., Jones, C.B., Hall, M.H. & Murdock, V. (**2018**). Geographic Information Retrieval: Progress and Challenges in Spatial Search of Text. *Foundations and Trends in Information Retrieval* 12 (2–3), pp. 164–318.

Ratter, B.M.W., Philipp, K.H.I. & Von Storch, H. (**2012**). Between hype and decline: Recent trends in public perception of climate change. *Environmental Science and Policy* 18, pp. 3–8.

Rea, G., Paton-Walsh, C., Turquety, S., Copec, M. & Griffith, D. (**2016**). Impact of the New South Wales fires during October 2013 on regional air quality in eastern Australia. *Atmospheric Environment* 131, pp. 150–163.

Reber, U. (**2020**). Global Climate Change or National Climate Changes? An Analysis of the Performance of Online Issue Publics in Integrating Global Issues. *Environmental Communication* pp. 1–16.

Rehurek, R. & Sojka, P. (**2010**). Software Framework for Topic Modelling with Large Corpora. *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* pp. 45–50.

Reid, C.E., Brauer, M., Johnston, F.H., Jerrett, M., Balmes, J.R. & Elliott, C.T. (**2016**). Critical Review of Health Impacts of Wildfire Smoke Exposure. *Environmental Health Perspectives* 124 (9), pp. 1334–1343.

Resch, B., Usländer, F. & Havas, C. (**2018**). Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. *Cartography and Geographic Information Science* 45 (4), pp. 362–376.

Rico, I., Izagirre, E., Serrano, E. & López-Moreno, J.I. (**2017**). Current glacier area in the pyrenees: An updated assessment 2016. *Pirineos. Revista de Ecología de Montaña* 172, pp. 1–7.

Riley, W.J. (**2008**). Algorithms for frequency jump detection. *Metrologia* 45, pp. 154–161.

Risbey, J.S. (**2008**). The new climate discourse: Alarmist or alarming? *Global Environmental Change* 18 (1), pp. 26–37.

Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., et al. (**2020**). Coronavirus Pandemic (COVID-19). *Our World in Data*, available at: https://ourworldindata.org/coronavirus (accessed 8 December 2021).

Ropeik, D. (**2012**). The Perception Gap: Recognizing and managing the risks that arise when we get risk wrong. *Food and Chemical Toxicology* 50 (5), pp. 1222–1225.

Ruiz, I., Faria, S.H. & Neumann, M.B. (**2020**). Climate change perception: Driving forces and their interactions. *Environmental Science and Policy* 108, pp. 112–120.

Sakaki, T., Okazaki, M. & Matsuo, Y. (**2010**). Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors. *Proceedings of the 19th International Conference on World Wide Web - WWW '10* pp. 851–860.

Savaresi, A. (**2016**). The Paris Agreement: a new beginning? *Journal of Energy & Natural Resources Law* 34 (1), pp. 16–26.

Scambos, T. & Shuman, C. (**2016**). Comment on "Mass gains of the Antarctic ice sheet exceed losses" by H. J. Zwally and others. *Journal of Glaciology2* 62 (233), pp. 599–603.

Schade, S., Ostermann, F., Spinsanti, L. & Kuhn, W. (**2012**). Semantic Observation Integration. *Future Internet* 4, pp. 807–829.

Schäfer, M.S., Ivanova, A. & Schmidt, A. (**2014**). What drives media attention for climate change? Explaining issue attention in Australian, German and Indian print media from 1996 to 2010. *International Communication Gazette* 76 (2), pp. 152–176.

Schnebly, J. & Sengupta, S. (**2019**). Random Forest Twitter Bot Classifier. *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 506–512.

Schuur, E.A.G., McGuire, A.D., Schädel, C., Grosse, G., Harden, J.W., Hayes, D.J., Hugelius, G., et al. (**2015**). Climate change and the permafrost carbon feedback. *Nature* 520, pp. 171–179.

Scikit-learn (n.d.). RandomForestClassifier, available at: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html (accessed 10 December 2021).

Segev, E. & Hills, T. (**2014**). When news and memory come apart: A cross-national comparison of countries' mentions. *International Communication Gazette* 76 (1), pp. 67–85.

Semenza, J.C., Hall, D.E., Wilson, D.J., Bontempo, B.D., Sailor, D.J. & George, L.A. (**2008**). Public Perception of Climate Change. Voluntary Mitigation and Barriers to Behavior Change. *American Journal of Preventive Medicine* 35 (5), pp. 479–487.

Shariff, S.M., Zhang, X. & Sanderson, M. (**2017**). On the credibility perception of news on Twitter: Readers, topics and features. *Computers in Human Behavior* 75, pp. 785–796.

Shi, Y., Wang, G., Cai, X., Deng, J., Zheng, L., Zhu, H., Zheng, M., et al. (**2020**). An overview of COVID-19. *Journal of Zhejiang University-SCIENCE B* 21 (5), pp. 343–360.

Sørensen, T.A. (**1957**). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskab* 5, pp. 1–34.

SpaCy (**2021**). Industrial-Strength Natural Language Processing, available at: https://spacy.io/ (accessed 30 June 2021).

Statista (**2021**). Number of monetizable daily active Twitter users (mDAU) worldwide from 1st quarter 2017 to 3rd quarter 2021, available at: https://www.statista.com/statistics/970920/monetizable-daily-active-twitter-users-worldwide/ (accessed 8 December 2021).

Stefanon, M., Dandrea, F. & Drobinski, P. (**2012**). Heatwave classification over Europe and the Mediterranean region. *Environmental Research Letters* 7, pp. 1–9.

Stern, N. & Stern, N.H. (**2007**). *The Economics of Climate Change: The Stern Review* Cambridge University Press, Cambridge.

Stier, S., Schünemann, W.J. & Steiger, S. (**2018**). Of activists and gatekeepers: Temporal and structural properties of policy networks on Twitter. *New Media and Society* 20 (5), pp. 1910–1930.

Sun, Q., Miao, C., Hanel, M., Borthwick, A.G.L., Duan, Q., Ji, D. & Li, H. (**2019**). Global heat stress on health, wildfires, and agricultural crops under different levels of climate warming. *Environment International* 128, pp. 125–136.

Taylor, A., Dessai, S. & Bruine de Bruin, W. (**2019**). Public priorities and expectations of climate change impacts in the United Kingdom. *Journal of Risk Research* 22 (2), pp. 150–160.

Tobler, W.R. (**1970**). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* 46, pp. 234–240.

Trewin, B., Cazenave, A., Howell, S., Huss, M., Isensee, K., Palmer, M.D., Tarasova, O., et al. (**2021**). Headline indicators for global climate monitoring. *Bulletin of the American Meteorological Society* 102 (1), pp. E20–E37.

Twitter (**2021**). Search Tweets. *Twitter Developer Platform*, available at: https://developer.twitter.com/en/docs/twitter-api/tweets/search/api-reference/get-tweets-search-all (accessed 25 June 2021).

Veltri, G.A. & Atanasova, D. (**2015**). Climate change on Twitter: Content, media ecology and information sharing behaviour. *Public Understanding of Science* 26 (6), pp. 721–737.

Verma, S., Vieweg, S., Corvey, W.J., Palen, L., Martin, J.H., Palmer, M., Schram, A., et al. (**2011**). Natural Language Processing to the Rescue? Extracting "Situational Awareness" Tweets During Mass Emergency. *Fifth International AAAI Conference on Weblogs and Social Media* pp. 385–392.

Vieweg, S., Hughes, A.L., Starbird, K. & Palen, L. (**2010**). Microblogging during two natural hazards events: What twitter may contribute to situational awareness. *Conference on Human Factors in Computing Systems - Proceedings* pp. 1079–1088.

Viñas, M.-J. (**2012**). Satellites See Unprecedented Greenland Ice Sheet Surface Melt. *NASA*, available at: https://www.nasa.gov/topics/earth/features/greenland-melt.html (accessed 23 November 2021).

Wang, D., Guan, D., Zhu, S., Kinnon, M. Mac, Geng, G., Zhang, Q., Zheng, H., et al. (**2021**). Economic footprint of California wildfires in 2018. *Nature Sustainability* 4, pp. 252–260.

Wang, Z., Ye, X. & Tsou, M.H. (**2016**). Spatial, temporal, and content analysis of Twitter for wildfire hazards. *Natural Hazards* 83 (1), pp. 523–540.

Weller, K., Dröge, E. & Puschmann, C. (**2011**). Citation analysis in twitter: Approaches for defining and measuring information flows within tweets during scientific conferences. *# MSM* pp. 1–12.

WGMS (**2021a**). Fluctuations of Glaciers Database World Glacier Monitoring Service, Zurich, available at: http://dx.doi.org/10.5904/wgms-fog-2021-05.

WGMS (**2021b**). *Global Glacier Change Bulletin No. 4 (2018-2019)*. eds. Zemp, M., Nussbaumer, S.U., Gärtner-Roer, I., Bannwart, J., Paul, F. and Hoelzle, M. ISC(WDS)/IUGG(IACS)/UNEP/UNESCO/WMO, World Glacier Monitoring Service, Zurich.

Whitmarsh, L. (**2011**). Scepticism and uncertainty about climate change: Dimensions, determinants and change over time. *Global Environmental Change* 21 (2), pp. 690–700.

WHO (**2020**). WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020, available at: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 (accessed 8 December 2021).

Williams, H.T.P., McMurray, J.R., Kurz, T. & Hugo Lambert, F. (**2015**). Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change* 32, pp. 126–138.

Wong, D.W.S. (**2004**). The Modifiable Areal Unit Problem (MAUP). in Janelle, D.G., Warf, B. and Hansen, K. (Eds.), *WorldMinds: Geographical Perspectives on 100 Problems* Springer, Dordrecht, pp. 571–575.

World Bank (**2021a**). Population, total, available at: https://databank.worldbank.org/reports.aspx?source=2&series=SP.POP.TOTL (accessed 1 November 2021).

World Bank (**2021b**). Individuals using the Internet (% of population), available at: https://data.worldbank.org/indicator/IT.NET.USER.ZS (accessed 1 November 2021).

World Meteorological Organization (**2019**). *WMO Statement on the State of the Global Climate in 2018* World Meteorological Organization, Geneva.

Xu, P., Wang, L., Liu, Y., Chen, W. & Huang, P. (**2020**). The record-breaking heat wave of June 2019 in Central Europe. *Atmospheric Science Letters* 21 (4), pp. 1–7.

Xu, R., Yu, P., Abramson, M.J., Johnston, F.H., Samet, J.M., Bell, M.L., Haines, A., et al. (**2020**). Wildfires, Global Climate Change, and Human Health. *New England Journal of Medicine* 383, pp. 2173–2181.

Yeo, S. (**2016**). Explainer: Paris Agreement on climate change to "enter into force". *CarbonBrief*, available at: https://www.carbonbrief.org/explainer-paris-agreement-to-enter-into-force (accessed 8 December 2021).

Zacharias, S., Koppe, C. & Mücke, H.G. (**2015**). Climate change effects on heat waves and future heat wave-associated IHD mortality in Germany. *Climate* 3 (1), pp. 100–117.

Zemp, M., Frey, H., Gärtner-Roer, I., Nussbaumer, S.U., Hoelzle, M., Paul, F., Haeberli, W., et al. (**2015**). Historically unprecedented global glacier decline in the early 21st century. *Journal of Glaciology* 61 (228), pp. 745–762.

Zemp, M., Thibert, E., Huss, M., Stumm, D., Rolstad Denby, C., Nuth, C., Nussbaumer, S.U., et al. (**2013**). Reanalysing glacier mass balance measurement series. *Cryosphere* 7, pp. 1227–1245.

Zhang, R., Sun, C., Zhu, J., Zhang, R. & Li, W. (**2020**). Increased European heat waves in recent decades in response to shrinking Arctic sea ice and Eurasian snow cover. *Climate and Atmospheric Science* 3, pp. 1–9.

Zhou, X., Josey, K., Kamareddine, L., Caine, M.C., Liu, T., Mickley, L.J., Cooper, M., et al. (**2021**). Excess of COVID-19 cases and deaths due to fine particulate matter exposure during the 2020 wildfires in the United States. *Science Advances* 7 (33), pp. 1–11.

7.2 References for demonstration purposes only

Bamber, J.L., Oppenheimer, M., Kopp, R.E., Aspinall, W.P. & Cooke, R.M. (**2019**). Ice sheet contributions to future sea-level rise from structured expert judgment. *Proceedings of the National Academy of Sciences* 116 (23), pp. 11195–11200.

Anonymus (**2020**). Bank of England: Stop funding the climate crisis and regulate the banks that do. *SumOfUs*, available at: https://actions.sumofus.org/a/bank-of-england-stop-funding-climate-chaos (accessed 14 November 2021).

BBC. (**2020**). Arctic Circle oil spill prompts Putin to declare state of emergency. *BBC*, available at: https://www.bbc.com/news/world-europe-52915807 (accessed 14 November 2021).

Bojanowski, A. (**2013**). Gletscherschmelze: Alpen werden zur Seenlandschaft. *Spiegel Online*, available at: https://www.spiegel.de/wissenschaft/natur/gletscher-schmelze-im-hochgebirge-in-den-alpen-entstehen-neuen-seen-a-894629.html (accessed 14 November 2021).

Blunden, J. & Arndt, D.S. (**2013**). State of the Climate in 2012. *Bulletin of the American Meteorological Society* 94 (8), pp. S1–S258.

Carson, M. & Peterson, G. (**2016**). *Arctic Resilience Report.* Arctic Council, Stockholm Environment Institute and Stockholm Resilience Centre, Stockholm.

Cazenave, A., Meyssignac, B., Ablain, M., Balmaseda, M., Bamber, J., Barletta, V., Beckley, B., et al. (**2018**). Global sea-level budget 1993-present. *Earth System Science Data* 10 (3), pp. 1551–1590.

Cook, C.P., Flierdt, T. van de, Williams, T., Hemming, S.R., Iwai, M., Kobayashi, M., Jimenez-Espejo, F.J., et al. (**2013**). Dynamic behaviour of the East Antarctic ice sheet during Pliocene warmth. *Nature Geoscience* 6, pp. 765–769.

Dangendorf, S., Marcos, M., Wöppelmann, G., Conrad, C.P., Frederikse, T. & Riva, R. (**2017**). Reassessment of 20th century global mean sea level rise. *Proceedings of the National Academy of Sciences* 114 (23), pp. 5946–5951.

DeConto, R.M. & Pollard, D. (**2016**). Contribution of Antarctica to past and future sea-level rise. *Nature* 531, pp. 591–597.

Ehlerding, S. (**2018**). Die Hitzewelle stellt die Wissenschaft vor ein Rätsel. *Der Tagesspiegel*, available at: https://www.tagesspiegel.de/gesellschaft/panorama/neues-wetterphaenomen-die-hitzewelle-stellt-die-wissenschaft-vor-ein-raetsel/22843130.html (accessed 24 November 2021).

Garbe, J., Albrecht, T., Levermann, A., Donges, J.F. & Winkelmann, R. (**2020**). The hysteresis of the Antarctic Ice Sheet. *Nature* 585, pp. 538–544.

Gergis, J. (**2020**). We are seeing the very worst of our scientific predictions come to pass in these bushfires. *The Guardian*, available at: https://www.theguardian.com/commentisfree/2020/jan/03/we-are-seeing-the-very-worst-of-our-scientific-predictions-come-to-pass-in-these-bushfires (accessed 24 November 2021).

Guarino, B. (**2016**). Anthrax sickens 13 in western Siberia, and a thawed-out reindeer corpse may be to blame. *The Washington Post*, available at: https://www.washingtonpost.com/news/morning-mix/wp/2016/07/28/anthrax-sickens-13-in-western-siberia-and-a-thawed-out-reindeer-corpse-may-be-to-blame/ (accessed 14 November 2021).

Hay, C.C., Morrow, E., Kopp, R.E. & Mitrovica, J.X. (**2015**). Probabilistic reanalysis of twentieth-century sea-level rise. *Nature* 517, pp. 481–484.

Hodgkins, S.B., Tfaily, M.M., McCalley, C.K., Logan, T.A., Crill, P.M., Saleska, S.R., Rich, V.I., et al. (**2014**). Changes in peat chemistry associated with permafrost thaw increase greenhouse gas production. *Proceedings of the National Academy of Sciences* 111 (16), pp. 5819–5824.

Hope, C. & Schaefer, K. (**2016**). Economic impacts of carbon dioxide and methane released from thawing permafrost. *Nature Climate Change* 6, pp. 56–59.

Horton, B.P., Rahmstorf, S., Engelhartd, S.E. & Kemp, A.C. (**2014**). Expert assessment of sea-level rise by AD 2100 and AD 2300. *Quaternary Science Reviews* 84 (15), pp. 1–6.

Joughin, I., Smith, B.E. & Medley, B. (**2014**). Marine Ice Sheet Collapse Potentially Under Way for the Thwaites Glacier Basin, West Antarctica. *Science* 344 (6158), pp. 735–738.

Kemp, A.C., Horton, B.P., Donnelly, J.P., Mann, M.E., Vermeer, M. & Rahmstorf, S. (**2011**). Climate related sea-level variations over the past two millennia. *Proceedings of the National Academy of Sciences* 108 (27), pp. 11017–11022.

Khazendar, A., Rignot, E., Schroeder, D.M., Seroussi, H., Schodlok, M.P., Scheuchl, B., Mouginot, J., et al. (**2016**). Rapid submarine ice melting in the grounding zones of ice shelves in West Antarctica. *Nature Communications* 7 (13243), pp. 1–8.

King, M.D., Howat, I.M., Candela, S.G., Noh, M.J., Jeong, S., Noël, B.P.Y., Broeke, M.R. van den, et al. (**2020**). Dynamic ice loss from the Greenland Ice Sheet driven by sustained glacier retreat. *Communications Earth & Environment* 1 (1), pp. 1–7.

Legendre, M., Bartolia, J., Shmakovab, L., Jeudya, S., Labadiec, K., Adraitd, A., Lescota, M., et al. (**2014**). Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proceedings of the National Academy of Sciences* 111 (11), pp. 4274–4279.

Lai, K.K.R., Lu, D. & Migliozzi, B. (**2019**). What Satellite Imagery Tells Us About the Amazon Rain Forest Fires. *The New York Times*, available at: https://www.nytimes.com/interactive/2019/08/24/world/americas/amazon-rain-forest-fire-maps.html (accessed 24 November 2021).

Lambeck, K., Rouby, H., Purcell, A., Sun, Y. & Sambridge, M. (**2014**). Sea level and global ice volumes from the Last Glacial Maximum to the Holocene. *Proceedings of the National Academy of Sciences* 111 (43), pp. 15296–15303.

Levermann, A., Winkelmann, R., Albrecht, T., Goelzer, H., Golledge, N.R., Greve, R., Huybrechts, P., et al. (**2020**). Projecting Antarctica's contribution to future sea level rise from basal ice shelf melt using linear response functions of 16 ice sheet models (LARMIP-2). *Earth System Dynamics* 11 (1), pp. 35–76.

Liu, J., Curry, J.A., Wang, H., Song, M. & Horton, R.M. (**2012**). Impact of declining Arctic sea ice on winter snowfall. *Proceedings of the National Academy of Sciences* 109 (11), pp. 4074–4079.

Luhn, A. (**2016**). Slow-motion wrecks: how thawing permafrost is destroying Arctic cities. *The Guardian*, available at: https://www.theguardian.com/cities/2016/oct/14/thawing-permafrost-destroying-arctic-cities-norilsk-russia (accessed 14 November 2021).

Lutz, M. & Brupbacher, M. (**2017**). So schmolzen die Schweizer Gletscher in 160 Jahren weg. *TagesAnzeiger*, available at: https://interaktiv.tagesanzeiger.ch/2017/gletscherschwund/ (accessed 14 November 2021).

Maurer, J.M., Schaefer, J.M., Rupper, S. & Corley, A. (**2019**). Acceleration of ice loss across the Himalayas over the past 40 years. *Science Advancesnc* 5 (6), pp. 1–12.

Milillo, P., Rignot, E., Rizzoli, P., Scheuchl, B., Mouginot, J., Bueso-Bello, J. & Prats-Iraola, P. (**2019**). Heterogeneous retreat and ice melt of Thwaites Glacier, West Antarctica. *Science Advances* 5, pp. 1–8.

Mouginot, J., Rignot, E., Scheuchl, B., Fenty, I., Khazendar, A., Morlighem, M., Buzzi, A., et al. (**2015**). Fast retreat of Zachariæ Isstrøm, northeast Greenland. *Science* 350 (6266), pp. 1357–1361.

Nerem, R.S., Beckley, B.D., Fasullo, J.T., Hamlington, B.D., Masters, D. & Mitchum, G.T. (**2018**). Climate-change–driven accelerated sea-level rise detected in the altimeter era. *Proceedings of the National Academy of Sciences* 115 (9), pp. 2022–2025.

Pazos, P. (**2013**). El mar se tragará Barcelona dentro de 5.000 años por el calentamiento global. *ABC Ciencia*, available at: https://www.abc.es/ciencia/20131106/abci-hielo-derrite-barcelona-desapareceria-201311061027.html (accessed 14 November 2021).

Rabatel, A., Francou, B., Soruco, A., Gomez, J., Cáceres, B., Ceballos, J.L., Basantes, R., et al. (**2013**). Current state of glaciers in the tropical Andes: a multi-century perspective on glacier evolution and climate change. *The Cryosphere* 7 (1), pp. 81–102.

Radić, V. & Hock, R. (**2011**). Regionally differentiated contribution of mountain glaciers and ice caps to future sea-level rise. *Nature Geoscience* 4, pp. 91–94.

Rahmstorf, S., Foster, G. & Cazenave, A. (**2012**). Comparing climate projections to observations up to 2011.

Environmental Research Letters 7 (4), pp. 1–5.

Reese, R., Gudmundsson, G.H., Levermann, A. & Winkelmann, R. (**2017**). The far reach of ice-shelf thinning in Antarctica. *Nature Climate Change* 8, pp. 53–57.

Rignot, E., Mouginot, J., Morlighem, M., Seroussi, H. & Scheuchl, B. (**2014**). Widespread, rapid grounding line retreat of Pine Island, Thwaites, Smith, and Kohler glaciers, West Antarctica, from 1992 to 2011. *Geophysical Research Letters* 41 (10), pp. 3502–3509.

Schuster, P.F., Schaefer, K.M., Aiken, G.R., Antweiler, R.C., Dewild, J.F., Gryziec, J.D., Gusmeroli, A., et al. (**2018**). Permafrost Stores a Globally Significant Amount of Mercury. *Geophysical Research Letters* 45 (3), pp. 1463–1471.

Schuur, E.A.G. & Abbott, B. (**2011**). High risk of permafrost thaw. *Nature* 480, pp. 32–33.

SCNAT (**2018**). Ein Jahr der Extreme für Schweizer Gletscher. *Swiss Academy of Sciences*, available at: https://scnat.ch/en/uuid/i/12c301ff-04a4-595e-ac81-c73f1e68982d-Ein_Jahr_der_Extreme_für_Schweizer_Gletscher (accessed 3 December 2021).

Shepherd, A., Ivins, E., Rignot, E., Smith, B., Broeke, M. van den, Velicogna, I., Whitehouse, P., et al. (**2018**). Mass balance of the Antarctic Ice Sheet from 1992 to 2017. *Nature* 558, pp. 219–222.

Shukman, D. (**2012**). Arctic ice melting at "amazing" speed, scientists find. *BBC*, available at: https://www.bbc.com/news/world-europe-19508906 (accessed 23 November 2021).

Vaks, A., Gutareva, O.S., Breitenbach, S.F.M., Avirmed, E., Mason, A.J., Thomas, A.L., Osinzev, A. V., et al. (**2013**). Speleothems Reveal 500,000-Year History of Siberian Permafrost. *Science* 340 (6129), pp. 183–186.

Varea, R. (**2020**). Drones, robots y "big data": tecnología punta contra los incendios forestales. *El País*, available at: https://elpais.com/retina/2020/08/08/innovacion/1596890172_365849.html (accessed 24 November 2021).

Wester, P., Mishra, A., Mukherji, A. & Shrestha, A.B. (**2019**). *The Hindu Kush Himalaya Assessment: Mountains, Climate Change, Sustainability and People* Springer, Cham.

World Weather Attribution (**2018**). Heatwave in northern Europe, summer 2018. *World Weather Attribution*, available at: https://www.worldweatherattribution.org/attribution-of-the-2018-heat-in-northern-europe/ (accessed 24 November 2021).

Zwally, H.J., Li, J., Robbins, J.W., Sava, J.L., Yi, D. & Brenner, A.C. (**2015**). Mass gains of the Antarctic ice sheet exceed losses. *Journal of Glaciology* 61 (230), pp. 1019–1036.

8. Appendix

This appendix contains content that was not essential to the description of the results but provides further insight into some of the arguments presented. The appendix sections are labelled according to the thesis chapters they correspond to.

Appendix 4.1.3

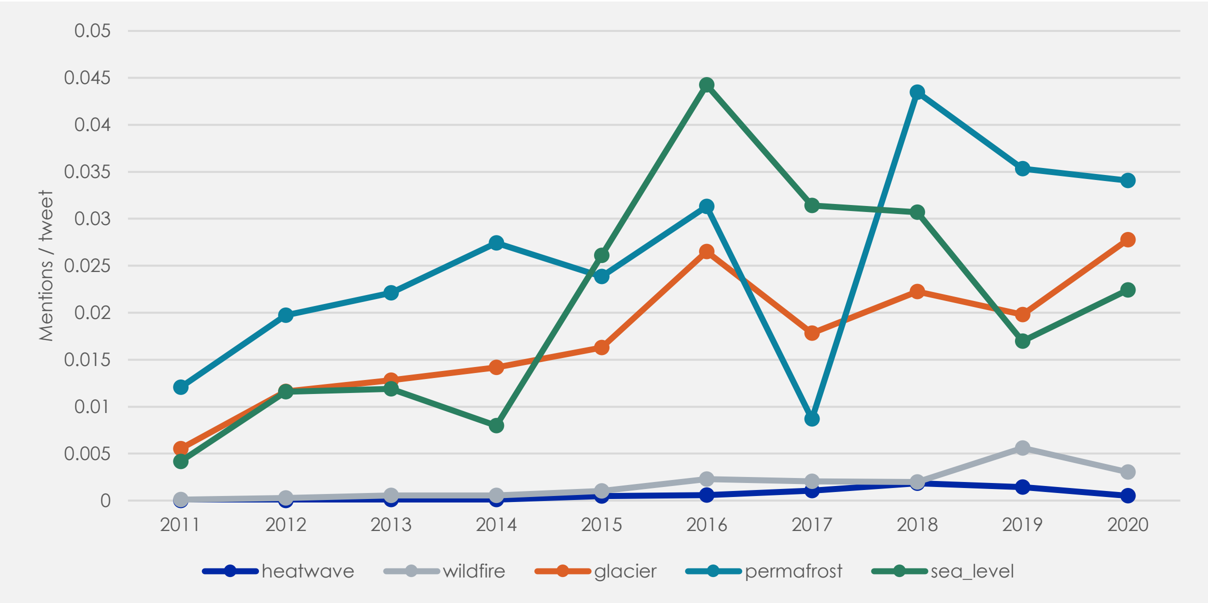

Fig. 52: Annual mean frequency of science URL usage in event tweets.

Appendix 4.1.5

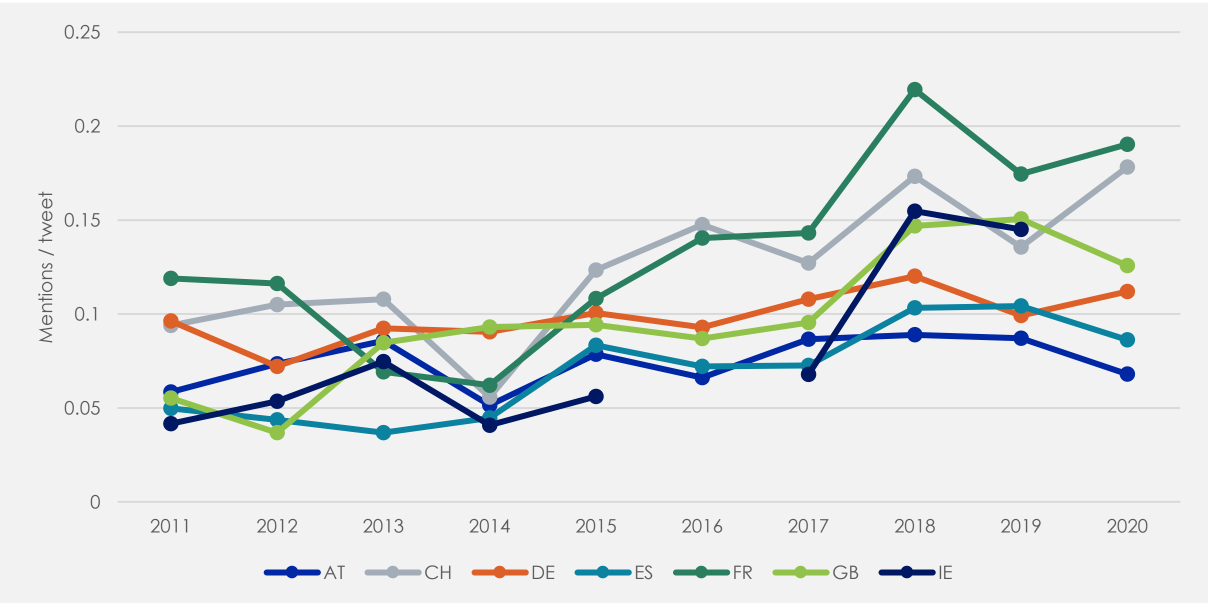

Fig. 53: Importance of danger per country in heatwave tweets.

Appendix 4.2.1

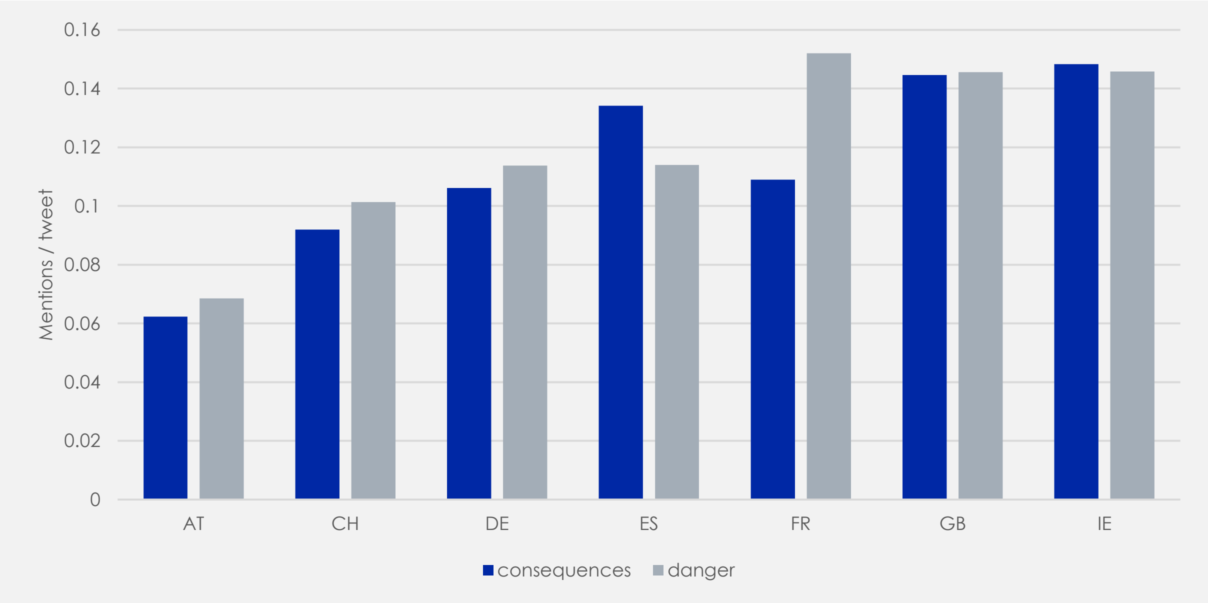

Fig. 54: Mean theme importance of 'consequence' and 'danger' for glacier tweets throughout the study period, divided by country.

Appendix 4.2.3

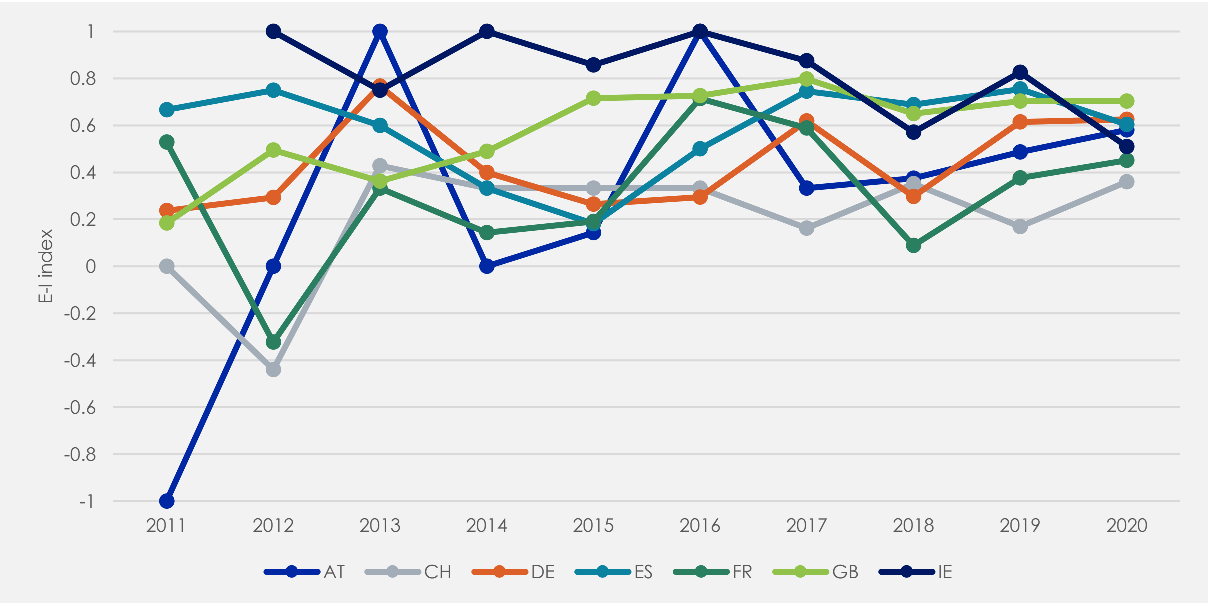

Fig. 55: Annual transnationality (E-I score) per country for permafrost tweets. Positive values indicate more transnational than domestic connections and negative values the inverse.

Appendix 4.3.6

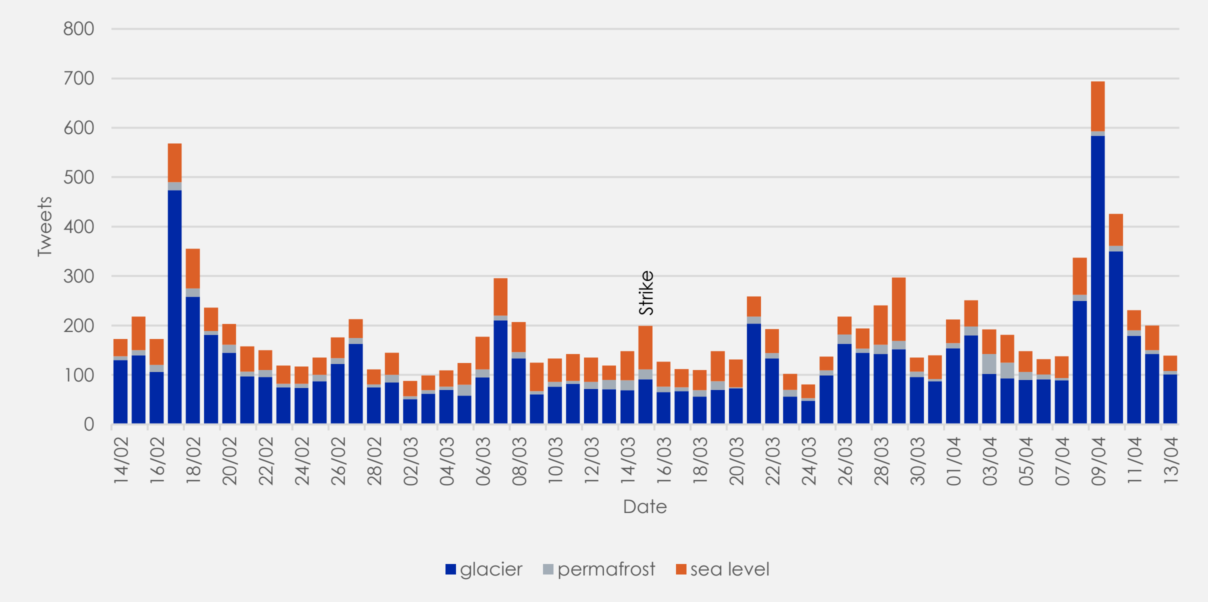

Fig. 56: Total number of glacier, permafrost and sea level tweets across the study area during the days surrounding the first Global Climate Strike (March 15 2019).

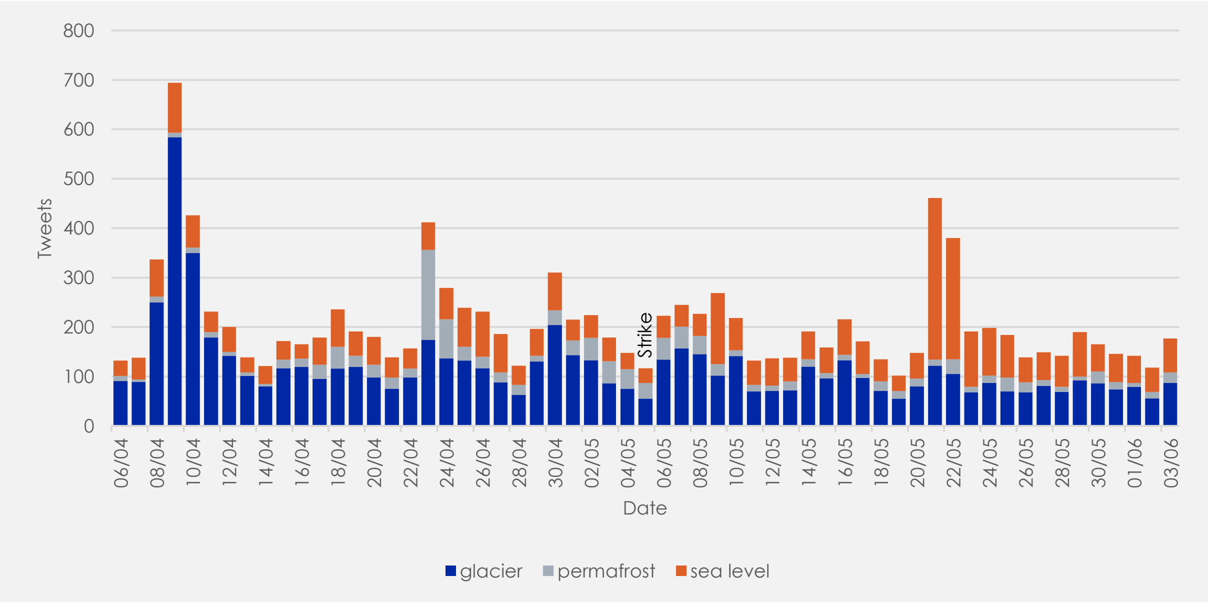

Fig. 57: Total number of glacier, permafrost and sea level tweets across the study area during the days surrounding the second Global Climate Strike (May 5 2019).

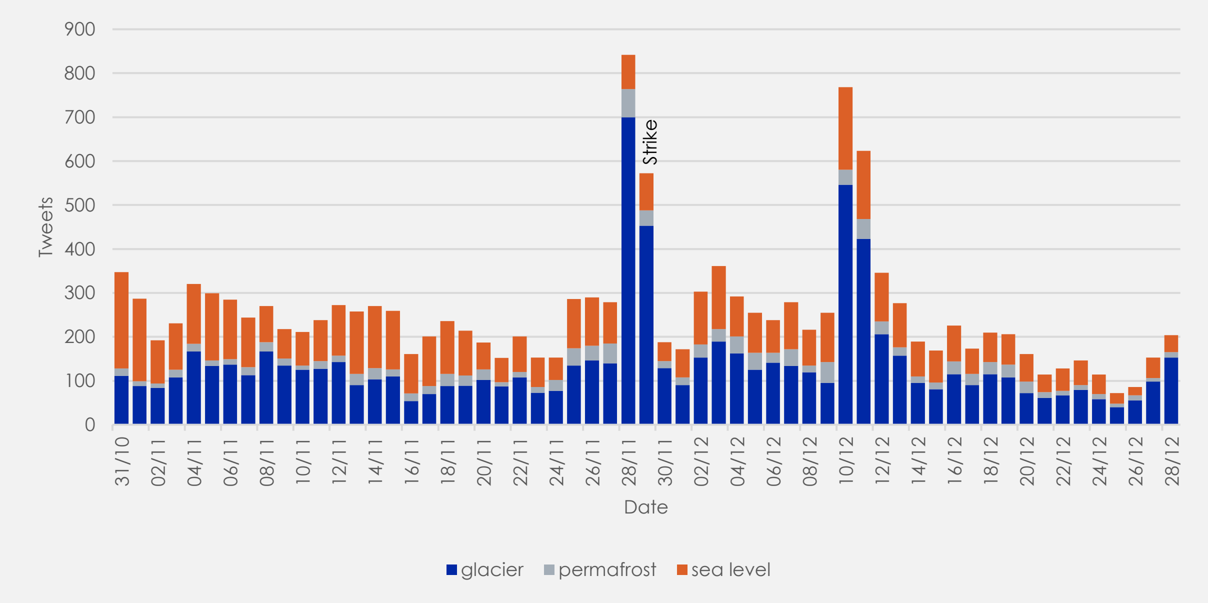

Fig. 58: Total number of glacier, permafrost and sea level tweets across the study area during the days surrounding the fourth Global Climate Strike (November 29 2019).

Personal declaration

I hereby declare that the submitted Thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the Thesis.