Resource and Energy Economics 77 (2024) 101437

Contents lists available at ScienceDirect

Resource and Energy Economics

journal homepage: www.elsevier.com/locate/ree

Feeling the heat? Analyzing climate change sentiment in Spain using Twitter data

Maria L. Loureiro[a,*], Maria Alló[b]

[a] University of Santiago de Compostela, Spain
[b] University of A Coruña, Spain

ARTICLE INFO

JEL Classifications:
Q50
Q51

Keywords:
Climate change
LabMT
Natural language processing
Sentiment analysis
Social media
VADER

ABSTRACT

To shed light on the recent debate about climate change in this post-pandemic scenario, we take advantage of a unique dataset that combines geo-tagged social media data from Twitter in Spain from 2017 to 2022. Twitter conversations have been analyzed with natural language processing techniques to obtain sentiment scores related to climate change. These were merged with additional relevant control variables, aiming to understand the role of the contributing factors on the evolution of the hedonic scores, including external temperatures, the occurrence of heat waves, and deaths related to climate. We find a strong negative effect of external temperatures on sentiment, aggravated by recent increases in the frequency of heat waves and deaths related to climate. Further, this negative sentiment is accentuated after experiencing the recent COVID-19.

1. Introduction

Understanding public opinion on climate change concerns is crucial, as aware individuals are more likely to support climate action policies (Bouman et al., 2020). Concerns related to climate change can arise from both, direct impacts such as health effects and socio-economic impacts, as well as indirect impacts related to mitigation and adaptation policies. In this paper, we assess public perceptions towards climate change in Spain, in the context of new scenarios emerging after COVID-19. We assess public preferences by analyzing an extensive dataset collected from social media which expands over the period 2017–2022.

Public opinion towards climate change has been extensively studied via surveys. In the United States, for instance, surveys have consistently shown that most of the population is concerned about climate change and supports government action to address it (Maibach et al., 2021). Likewise, a survey by the European Investment Bank (2020) found that 91% of respondents in Sweden believe that climate change is a serious problem, compared to 67% in Poland. This survey also found that younger participants are more concerned about climate change, with 86% of respondents under 25 considering it a serious problem. Similar results were previously reported by the Special Eurobarometer (2019) across Europe. It showed that about 90% of the Europeans consider that climate change is a serious problem caused by human activity, and 80% believe that the EU should aim to be climate-neutral by 2050.

In addition to these geographical variation, public opinion on climate change has undergone time variant shifts in recent years, with more people expressing concern and supporting action to address the issue now. This can be attributed to several factors, including the growing awareness of the impacts of climate change due to provision of scientific updated reports, and the actual experience of worst and more frequent climate extreme events worldwide. One of the most notable shifts in public opinion has been in

* Corresponding author.
E-mail address: maria.loureiro@usc.es (M.L. Loureiro).

https://doi.org/10.1016/j.reseneeco.2024.101437
Received 27 March 2023; Received in revised form 22 February 2024; Accepted 7 March 2024
Available online 8 March 2024
0928-7655/

the United States. According to Maibach et al. (2021), 70% of Americans now believe that global warming is happening, up from 57% in 2010. The study also found that 64% of Americans are at least "somewhat worried" about global warming, up from 49% in 2010. Similarly, in Europe, public opinion on climate change has been shifting towards greater concern and support for action. A recent survey by the European Investment Bank (2020) found that 68% of respondents in the European Union believe that climate change is a very serious problem, up from 54% in 2019. The survey also found that 67% of respondents believe that the EU should increase its climate action. The recent impact of COVID-19 in Spain has been studied by Drews et al. (2022) who found that although the level of climate change concern has decreased immediately after, the acceptance for most climate policies has increased. Savin et al. (2022) focused on how this pandemic is expected to influence climate actions in Spain. Their results show that positive views are more common among young people, with higher level of education and male respondents and also those who perceive climate change as a concern or those who valuate COVID-19 confinement positively.

Overall, these temporal shifts in public opinion towards climate change show positive developments, as they encourage greater political action to address the issue. A testimony describing a positive change in attitudes towards climate change is the work by Carlsson et al. (2021) who used the same survey instrument 10 years apart in the United States, China and Sweden. They find that all three countries show higher acceptability for climate mitigation now. However, there is still much work to be done to understand whether climate concerns may somewhat vanish when urgent policy instruments that require significant efforts by taxpayers are implemented. A good example is the recent "yellow vest" movement in France, which provoked the cancelation of the fuel tax planned by the government. Likewise, the recent Ukrainian invasion has spread a wave of concerns about fuel prices, self-sufficiency, and inflation in Europe, which have characterized the recent European policy agenda, postponing in a way, the ambitious European climate goals.

We contribute to the previous literature by using text analysis on social media that complements previous findings in stated preference studies. Text analysis has been recently used to assess the importance of perceptions towards economic growth (Savin et al., 2021), or the importance of achieving the sustainable development goals (Rosenberg et al., 2023). Our contribution in this research is twofold. First, we employ a unique dataset containing more than 2.16 million geo-tagged tweets in Spain during 2017–2022 from which we will extract temporal and geographical variations of sentiments toward climate change. Furthermore, we will assess econometrically how sentiments, or hedonic scores, vary with respect to different external factors, including extreme events and health related impacts.

The present paper has the following structure. First, we review the most relevant literature on sentiment analysis, and how sentiment analysis can be used to proxy changes in preferences and satisfaction. Next, we present the crucial features of sentiment analysis and the algorithms used. We follow with a description of the datasets employed for the empirical analysis, and we conclude with the main results and policy implications.

2. Literature review

The negative effects of climate change, and preferences towards mitigation and adaptation policies have been widely studied with stated preference surveys (Cai et al., 2010; Longo et al., 2012; Suter et al., 2021; Maestre-Andrés et al.; 2021; Savin et al., 2020; among others). In general terms, the studies conclude that perceptions toward corrective policies and taxes matter (Drews et al., 2022), while public acceptability of carbon taxation depends on its revenue use (Maestre-Andrés et al., 2021). Individuals are willing to pay more to prevent the impacts of climate change when global and ancillary benefits are included (Longo et al., 2012); or else when the cost to be borne by the rich are higher (Cai et al., 2010); or when crucial resources, such as water, become very scarce (Suter et al., 2021).

A different strand of the literature has employed subjective wellbeing (SWB) elicitation methods to analyze impacts form climate change or environmental impacts on wellbeing and happiness. Rehdanz and Maddison (2005, 2008) evaluate the impacts of weather and climate on SWB, suggesting that there is a relationship between weather and climate, with warmer weather being associated with higher levels of SWB in colder climates, but not in warmer climates. Barrington-Leigh (2008) also find that the influence of weather and climate is greater in rural areas than in urban areas. Luechinger and Raschky (2009) analyze the case of flood disasters and SWB. Levinson (2012) investigated how air quality affects well-being and happiness levels. More recently, Ambrey et al. (2017) analyzed the impact of wildfires in Australia on happiness levels, finding that the average happiness levels of residents near affected areas decreased significantly after the wildfires.

In addition to stated preferences and SWB techniques, sentiment indicators have also been used in economics to represent how a group, or a market feels about current economic conditions, and how these feelings and beliefs may affect future market behavior. Recent advances allow the use of sentiment analysis in massive amounts of text, such as social media posts, comments, reviews, blog posts, emails, and other text-based media. Sentiment analysis can be used applying a variety of dictionaries or machine learning algorithms to measure the overall state of happiness of by analyzing the content of their conversations or comments. For example, Savin et al. (2020) apply topic modelling to assess how perceptions of fairness matter on carbon taxation acceptability. Other recent applications assess how market perceptions, and the general trust in institutions and market prospects affect financial behavior. For example, Gupta and Chen (2020) use sentiments expressed on social for stock price prediction, and Ahmed (2020) analyzes how the investor sentiment affects the way which prices reflect information.

We employ sentiment analysis to gauge how views and concerns about climate change are varying over time. This will make this technique somewhat comparable inferring the level of happiness (sentiment) towards the topic. In similar lines of research Loureiro and Alló (2020) conducted an assessment on how sentiments and emotions towards climate change were expressed in social media in the U.K. and Spain, finding that messages in the U.K. related to climate change are less negative than in Spain. Baylis (2020) uses sentiment analysis in the context of climate change conversations in USA. The results show a negative effect between highest temperatures and overall stated sentiment. In a similar fashion, Loureiro et al. (2022) use sentiment analysis in the context of wildfires occurring in the Iberian peninsula, finding a negative relationship between closer distances and smoke, and expressed sentiment.

Overall, these studies suggest that concerns towards climate change and public preferences are shaped by a wide range of factors, including perceptions of effectiveness, fairness, and related costs. By understanding these preferences, policymakers can design climate policies that are more likely to be accepted and supported, and that can lead to relevant results to address this pressing issue.

3. Methods: natural language processing (NLP) applied to sentiment analysis

Sentiment analysis can be applied to text, such as social media posts, newspapers and other text-based media. NLP techniques identify the sentiment underlying massive amounts of text, classifying it as positive, negative, or neutral, and/or providing a general numerical score of the entire sentence or expression denoted as sentiment score (Bhadane et al., 2015). This is done by looking up the words in a variety of lexica and assigning a value to each word based on the polarity of the word as found in the lexicon. In this application, we used two different lexica to generate sentiment scores and check their robustness: The Language Assessment by Mechanical Turk (LabMT) by Dodds et al. (2011), and the Valence Aware Dictionary and sEntiment Reasoner (VADER) by Hutto and Gilbert (2014). The LabMT lexicon consists of over 10,000 words, each of which is assigned a valence score on a scale from 1 (very negative) to 9 (very positive). For example, the word "party" would be assigned a high positive valence score, while the word "terrorism" would be assigned a high negative valence score. The average LabMT score per tweet is obtained by adding up the scores of each single word and dividing it by the total number of words.

On the other hand, VADER uses a combination of linguistic rules, and heuristics to identify sentiment intensity in text. It is especially suited for the language of social media and works well on texts from other domains too, for example, movie and technical product reviews or opinion news articles.[1] The VADER lexicon consists of a list of words that are labeled with their corresponding positive or negative valence scores, as well as a list of booster words that can modify the intensity of sentiment.

The VADER sentiment analysis tool uses a compound score formula to calculate the overall sentiment polarity of a piece of text. The formula considers the valence scores of individual words in the text, as well as their grammatical relationships and the presence of booster words or negations.[2] The resulting sentiment compound score is a normalized score between −1 (extremely negative) and +1 (extremely positive), with 0 indicating a neutral sentiment.

4. Data description

To construct the general database, we assembled different databases. Of relevance is the sentiment data, which are the vectors of sentiment scores obtained by processing our tweets with the LabMT and VADER libraries, respectively. The sentiment scores have been generated by area (province) and day. We collected tweets from October 2017 to December 2022. Specifically, geo-tagged messages retrieved from a Boolean search string containing the keyword “climate change” were collected in Spain in this social media. It is important to mention that tweets were collected and analyzed in Spanish, its original language. Once the textual information was collected, it was necessary to clean and process data to analyze the conversations. The first step requires cleaning the text, and removing the information not related to climate change. For example, tweets related to songs, sayings, although included words such as “temperature”, “climate”, and other, were eliminated, since this information was not related to the topic at hand. Furthermore, monosyllable tweets were also removed. In addition, retweets were not considered for this text analysis because they repeat another tweet, if a tweet is a response to another tweet, the full tweet has been included and not just the response. After this first cleaning, we set all the text to lower case, we clean accents and stop words.

The final database contains about 2,161,007 geo-tagged conversations over the period of analysis (after missing values were dropped). This dataset contains information about the date and time of each tweet, as well as the associated sentiment score, the location from where the tweet had been written (specifically, the province).[3] A descriptive of the main topics dealt with during the period of analysis has been obtained via topic identification. Fig. 2 shows the most relevant topics in pre- and post-pandemic scenarios. We find a relatively stable presence of topics in conversations occurring during this period. However, in the last year, particularly, the intensity by which individuals refer to extreme events (particularly wildfires) has increased in the dataset. Other topics related to “water” scarcity and the importance of the local (“home”) climate related problems versus a previous more global vision “world” seem to be also gaining importance in recent months.

Additionally, with the aim of understanding the consequences that climate change has on the occurrence of extreme events, we collected information from the International Disaster Database known as EM-DAT (Guha-Sapir et al., 2023) to identify where and when meteorological or climatological disasters took place during the years of the study. During the period from 2017 to 2022, the EM-DAT contains information particularly relevant in terms of the episodes of extreme temperatures (heat waves) in Spain (Guha-Sapir et al.,

[1] See Annex (Table A1) for summary statistics of both sentiment scores.

[2] The VADER compound score is calculated as:compound = sum of valence scores of each word in the text + (sum of valence scores of each booster word in the text x booster weight) + (sum of valence scores of each word preceding a negation in the text x negation weight). The resulting compound score is a normalized score between −1 (extremely negative) and +1 (extremely positive), with 0 indicating a neutral sentiment.

[3] The initial database indicates the place from which the user is writing; however, in order to be able to incorporate important explanatory variables into the database and, at the same time, reduce the noise of such disaggregated information, the database is analyzed by aggregating the data at the province level.

Table 1
Descriptive Statistics.

Variable	Description	Mean	Std. Dev.
LabMT	The LabMT sentiment by province (from 1 to 9)	5.8232	.3352
VADER	The VADER sentiment was recoded taking the value of 0 if the sentiment was <0; 1, if the sentiment was =0, and equals 2 if the sentiment was >0 (by province)	.0051	.0279
Max temp	Maximum temperatures recorded according to the day, month, and year	30.418	8.291
Less than 18ºC	1, it the maximum temperature is below 18ºC; 0 otherwise.	.0787	.2693
18.01ºC and 21ºC	1, if the maximum temperature is between 18.01ºc and 21ºC; 0 otherwise.	.1034	.3045
21.01ºC and 24ºC	1, if the maximum temperature is between 21.01ºc and 24ºC; 0 otherwise.	.1125	.3160
24.01ºC and 27ºC	1, if the maximum temperature is between 24.01ºc and 27ºC; 0 otherwise.	.0966	.2955
27.01ºC and 30ºC	1, if the maximum temperature is between 27.01ºc and 30ºC; 0 otherwise.	.0690	.2535
30.01ºC and 33ºC	1, if the maximum temperature is between 30.01ºc and 33ºC; 0 otherwise.	.0635	.2439
33.01ºC and 36ºC	1, if the maximum temperature is between 33.01ºc and 36ºC; 0 otherwise.	.1346	.3413
36.01ºC and 39ºC	1, if the maximum temperature is between 36.01ºc and 39ºC; 0 otherwise.	.1722	.3776
39.01ºC and 42ºC	1, if the maximum temperature is between 39.01ºc and 42ºC; 0 otherwise.	.1328	.3393
More than 42.01ºC	1, if the maximum temperature is above 42.01ºC; 0 otherwise.	.0366	.1877
Deaths	Number of deaths attributed to climate	.0778	.694
Heat wave	Number of days per month with heat waves recorded by EM_DAT for the period 2017–2022	.0005	.0234
Before COVID-19	=1 if the tweet was written before March, 14th 2020, 0 otherwise	.5063	.5000

2023). In addition to direct deaths, heat waves can lead to heat exhaustion, dehydration, insomnia, and other heat-related illnesses. Furthermore, air pollution and pollen levels can worsen during heat waves, exacerbating respiratory problems. Hence, we expect the general hedonic score to be affected by the occurrence of heatwaves. The number of expected deaths due to outdoor temperature were collected from the MoMo database (MoMo, 2018). Moreover, information about registered maximum temperatures was also included. This information comes from the WeatherAPI (2022) database, which provides information for each registered zip code in Spain (See Table 1 for a full description of variables).

5. Empirical model specification

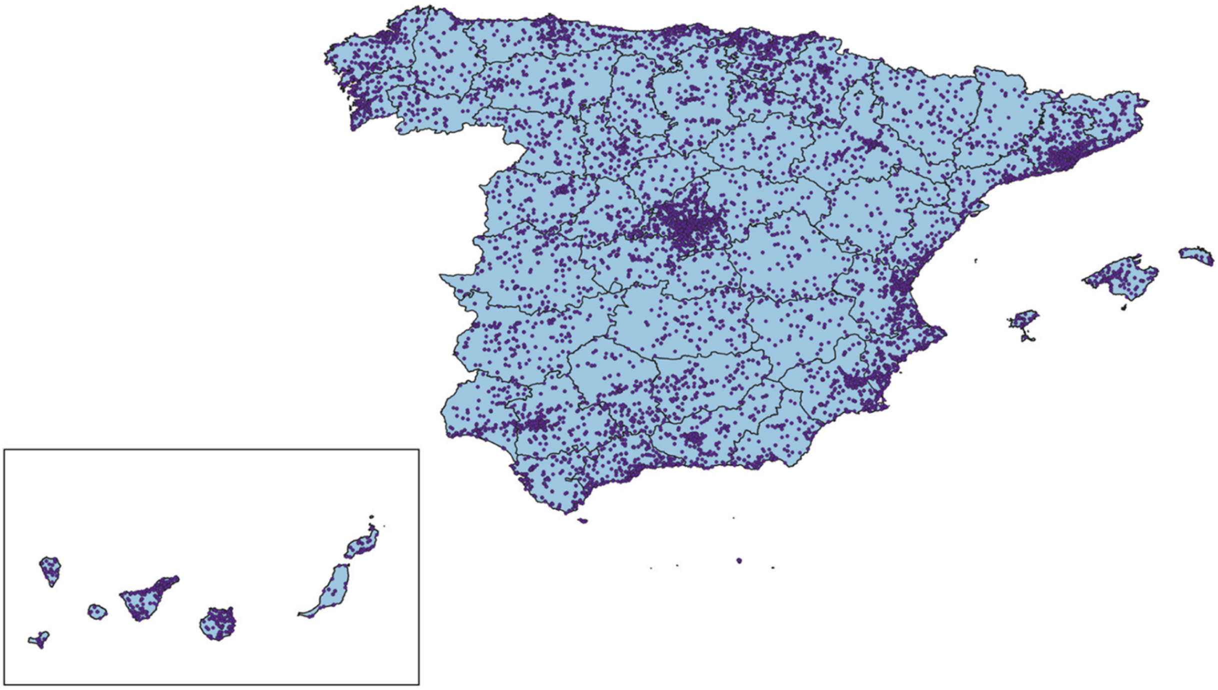
We assume that the sentiment scores extracted from the spontaneous Twitter conversations are representations of how people feel (unhappy-happy) about climate change, and hence, of their underlying preferences, in the same vein as Baylis (2020). We estimate different empirical models to check their robustness in order to assess the causal relationship between climate related issues and sentiment, while exploiting the panel structure of our dataset. First, we estimate a Robust Ordinary Least Square (Robust OLS) to fit LabMT lexicon scores with fixed effects. Next, and based on the ordered nature of our dependent variable when estimated with the lexicon VADER, we also estimate a Robust Ordered Logit (Robust OLogit) with fixed effects with its corresponding marginal effects.

The empirical functions are as follows:

$$Sentiment_{jt} = \alpha + \gamma maximum \quad temperatures_{jt} + \mu X_{jt} + \varepsilon_{jt} \tag{1}$$

Where $Sentiment_{jt}$ is the aggregated sentiment score of individuals in the same location (province) j at date t. The *maximum temperatures* is the local maximum temperature recorded the day and time when the individual was tweeting, from which we derived two empirical specifications: the first one being a linear baseline specification, and a second which considers non-linear effects including temperature intervals. The X_{ijt} represents a vector containing the number of days classified as heat waves, deaths caused by climate, and depending on the fixed-effect specification, whether the sentiment was expressed before or after COVID-19. To test for the correct empirical specification, we checked for the presence of multicollinearity through the Variance Inflation Factor (VIF) indicator, finding a mean value of 1; therefore, there is not a serious problem about multicollinearity.[4]

[4] Here, we provide the VIF for each variable: before covid 1.02, maximum temperature: 1.03, deaths 1.02 and heat wave 1.00. Thus, all VIF measures are below 5.



Map 1. Geo-location of Tweets.

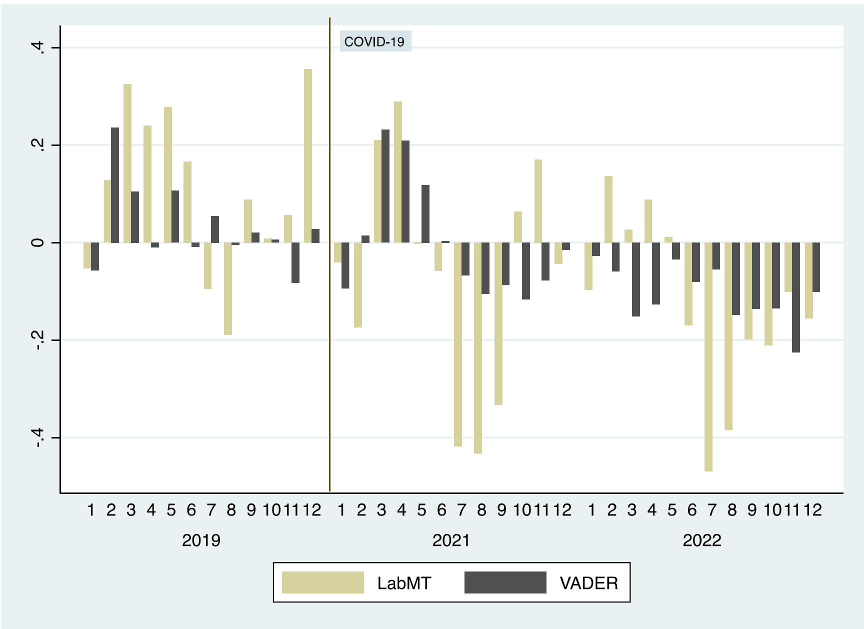
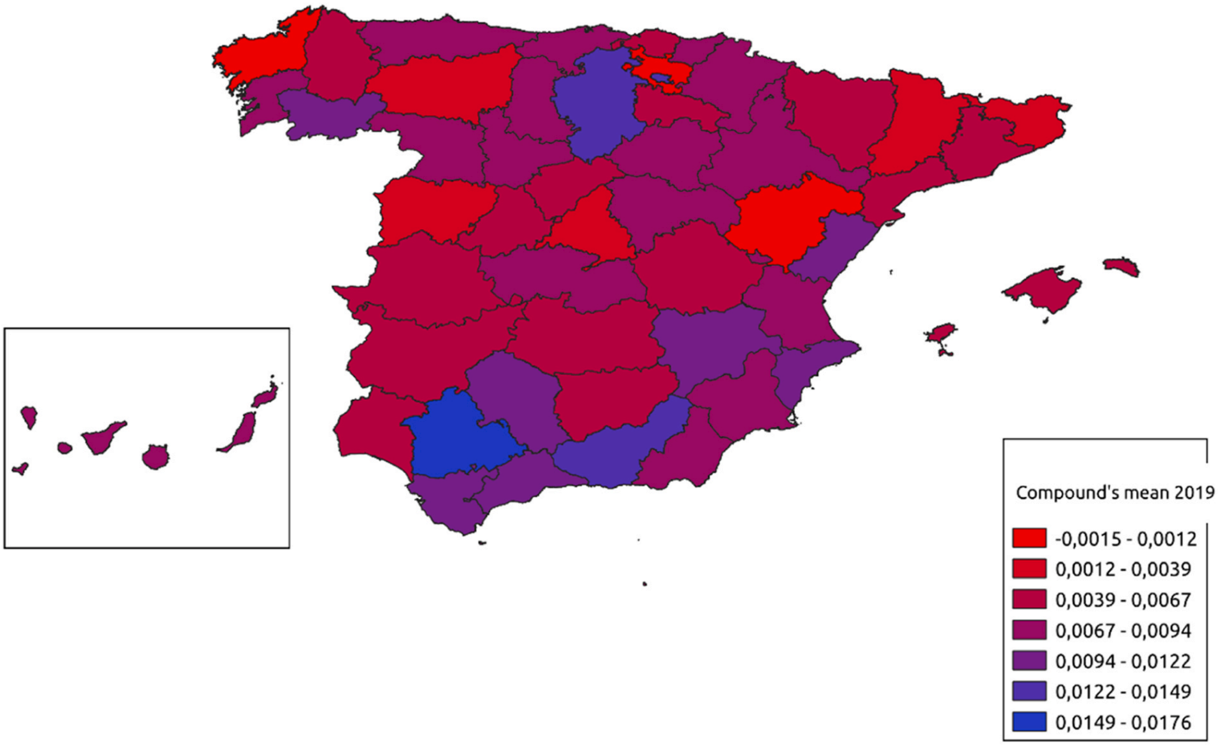


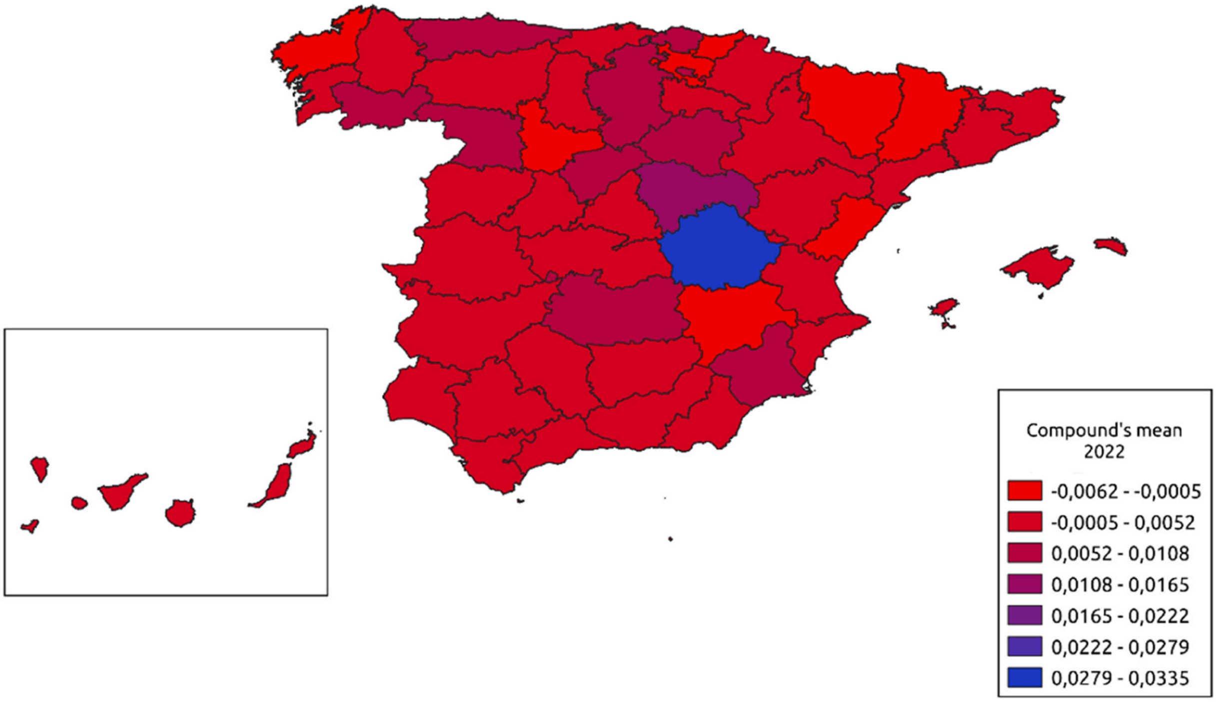
Fig. 1. Sentiment Scores by Year-LabMT and VADER for the period 2019–2022. Note: the measure is standardized to have zero mean and unit standard deviation.

6. Results

Descriptive results about the evolution of sentiment scores are also provided on maps. Map 1 shows the 2021 tweets by location in our dataset, presenting a visual representation of the general spread of the conversations across the entire country. The overall evolution of the sentiment score is presented in Fig. 1, showing a declining trend over the period, particularly after the COVID-19 pandemic in 2022. This can be also observable by looking at the geographical evolution of these sentiment scores by province,



Map 2. Climate Change Sentiment- VADER Compound by Province, Spain 2019.



Map 3. Climate Change Sentiment- VADER Compound by Province, Spain 2022.

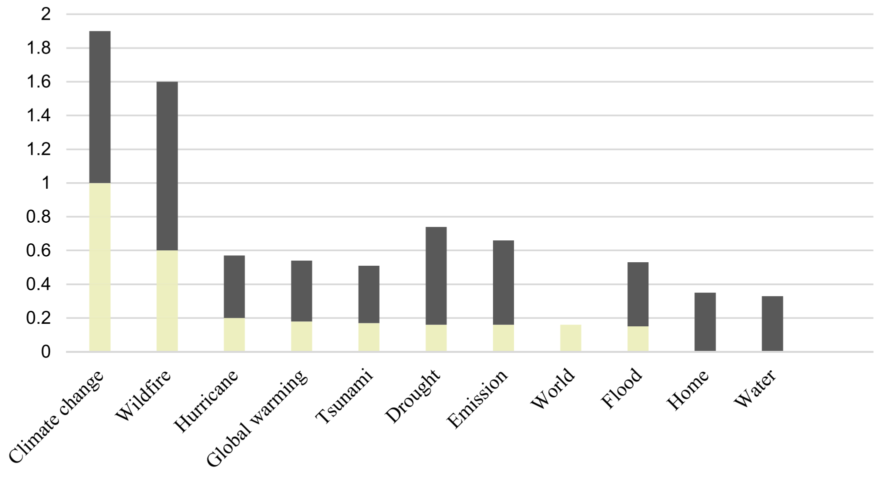


Fig. 2. Topic Frequency in 2019 and 2022 (Prior and Post-COVID outbreak).

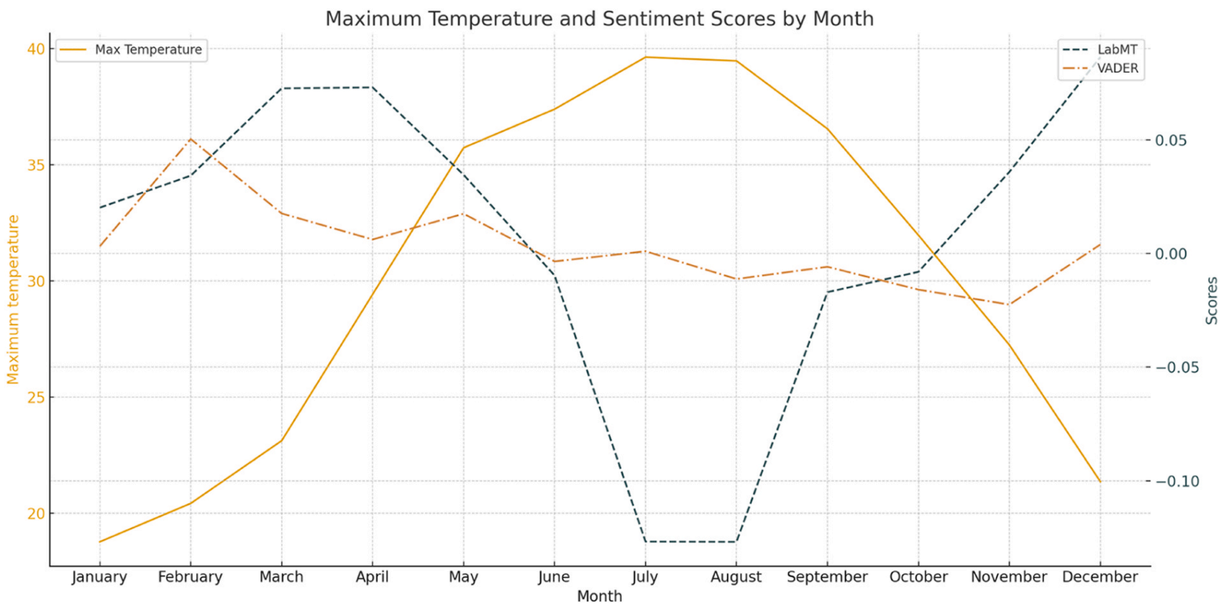


Fig. 3. Average Temperatures, LabMT and VADER Sentiment Scores per Month (2017–2022).

which are represented in Maps 2 and Map 3, showing a general reduction in the overall hedonic score in practically all provinces, and not only in the most southern and warmer part of country. This shows a clear manifestation of the increased concerns about climate change over the period of analysis.Fig. 2

Fig. 3 presents the evolution of maximum temperature and the mean sentiment scores by month, showing that in the case of LabMT scores there is a clear drop of average sentiment during the summer months. This corresponds with the months with highest temperatures and more heatwaves. Mean sentiments obtained from VADER show that the average sentiment is also higher during the winter months, while the lowest experienced sentiment occurs in October and November, months recently characterized by the occurrence of aggressive autumn wildfires.

Table 2 presents the results considering first a linear relationship between temperatures and sentiment scores obtained from LabMT. The first six columns present the Robust OLS results controlling for fixed effects for the LabMT sentiment scores, assuming a linear effect of temperature on the sentiment scores. The robustness of the findings show that maximum temperatures, heatwaves, and deaths related to climate decrease significantly the expressed sentiment across the various specifications. In addition, the sentiment was aggravated by the COVID-19 pandemic. The last three columns correspond to a non-linear specification of temperature on sentiments. In terms of the non-linear effects of temperature on sentiment, we find that our results correspond quite well with those by

Table 2
Robust OLS LabMT Results.

Robust OLS LabMT																
		Linear Temperature											Non linear temperature			
Max temp	-.004***	-.001***	-.002***	-.004***	-.004***	-.001***	-.002***	-.004***								
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)								
Deaths					-.013***	-.008***	-.003***	-.012***					-.008***	-.002***	-.007***	-.01***
					(.000)	(.000)	(.000)	(.000)					(.000)	(.000)	(.000)	(.000)
Heat wave					-.096***	-.248***	-.200***	-.122***					-.098***	-.255***	-.207***	-.121***
					(.021)	(.020)	(.020)	(.021)					(.021)	(.020)	(.020)	(.021)
Before COVID-19				.043***				.043***				.040***				.040***
				(.000)				(.000)				(.000)				(.000)
18ºC and 21ºC									.013***	-.004***	-.007***	.012***	.013***	-.003***	-.006***	.012***
									(.00)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
21ºC and 24ºC									.035***	.005***	.0037***	.035***	.035***	.006***	.003***	.035***
									(.001)	(.001)	(.0017)	(.001)	(.001)	(.001)	(.001)	(.001)
24ºC and 27ºC									.005***	-.007***	-.012***	.007***	.005***	-.006***	-.011***	.008***
									(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
27ºC and 30ºC									.0103***	.009***	.001	.012***	.010***	.009***	.002	.013***
									(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
30ºC and 33ºC									-.005***	-.012***	-.019***	-.000	-.005***	-.010***	-.017***	-.000
									(.0011)	(.002)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
33ºC and 36ºC									.002***	.007***	-.002	-.000	.002***	.007***	-.001	-.000
									(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
36ºC and 39ºC									-.073***	-.047***	-.055***	-.073***	-.072***	-.046***	-.054***	-.072***
									(.001)	(.002)	(.002)	(.001)	(.001)	(.001)	(.001)	(.001)
39.01ºC and 42ºC									-.094***	-.044***	-.053***	-.085***	-.091***	-.043***	-.052***	-.083***
									(.001)	(.002)	(.002)	(.001)	(.001)	(.002)	(.001)	(.0009)
More than 42ºC									-.109***	-.058***	-.067***	-.094***	-.105***	-.056***	-.066***	-.091***
									(.001)	(.002)	(.002)	(.001)	(.001)	(.002)	(.002)	(.001)
Constant	5.78***	5.351***	5.193***	5.756***	5.783***	5.351***	5.193***	5.754***	5.671***	5.339***	5.175***	5.65***	5.671***	5.339***	5.175***	5.650***
	(.002)	(.014)	(.014)	(.003)	(.003)	(.014)	(.014)	(.003)	(.003)	(.014)	(.014)	(.003)	(.003)	(.014)	(.014)	(.003)
		Fixed Effects														
Province	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes
Day, month, year	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Observations	2161007	2161007	2161007	2161007	2161007	2161007	2161007	2161007	2161007	2161007	2161007	2161007	2161007	2161007	2161007	2161007
R-squared	.118	.0573	.150	.122	.119	.058	.150	.123	.125	.059	.152	.130	.126	.060	.152	.129

Table 3
Robust Ologit VADER Results (1).

		Robust Ologit VADER														
		Linear Temperature											Nonlinear Temperature			
		Marg effects		Marg effects		Marg effects		Marg effects		Marg effects		Marg effects		Marg effects		Marg effects
Max temp	-.013***	-.003***	-.003***	-.001***	-.005***	-.001***	-.008***	-.002***								
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)								
Before COVID-19							.705***	.152***							.687***	.147***
							(.003)	(.001)							(.003)	(.001)
18ºC and 21ºC									.118***	.026***	-.005	-.001	-.013*	-.0029*	.097***	.021***
									(.006)	(.001)	(.007)	(.001)	(.007)	(.0015)	(.006)	(.001)
21ºC and 24ºC									-.076***	-.0168***	-.093***	-.020***	-.104***	-.0221***	-.069***	-.014***
									(.006)	(.001)	(.007)	(.001)	(.007)	(.0015)	(.006)	(.001)
24ºC and 27ºC									-.243***	-.053***	-.060***	-.012***	-.081***	-.0173***	-.202***	-.043***
									(.006)	(.001)	(.008)	(.001)	(.008)	(.0017)	(.006)	(.001)
27ºC and 30ºC									-.528***	-.116***	-.227***	-.049***	-.265***	.0562***	-.490***	-.105***
									(.007)	(.001)	(.009)	(.002)	(.009)	(.0019)	(.007)	(.001)
30ºC and 33ºC									-.464***	-.102***	-.128***	-.027***	-.169***	-.0358***	-.369***	-.079***
									(.007)	(.001)	(.010)	(.002)	(.010)	(.0021)	(.007)	(.002)
33ºC and 36ºC									-.043***	-.009***	.042***	.009***	-.002	-.0004	-.079***	-.017***
									(.006)	(.001)	(.010)	(.002)	(.010)	(.0021)	(.006)	(.001)
36ºC and 39ºC									-.185***	-.040***	-.014	-.002	-.056***	-.0120***	-.174***	-.037***
									(.005)	(.001)	(.010)	(.002)	(.010)	(.0022)	(.006)	(.001)
39.01ºC and 42ºC									-.363***	-.080***	-.053***	-.011***	-.102***	-.0217***	-.212***	-.045***
									(.006)	(.001)	(.011)	(.002)	(.010)	(.0023)	(.006)	(.001)
More than 42ºC									-.532***	-.117***	-.174***	-.037***	-.223***	-.0474***	-.287***	-.061***
									(.008)	(.002)	(.012)	(.002)	(.012)	(.0026)	(.008)	(.002)
/cut1	-.768***		.459***		1.054***		-.325***		-.561***		.464***		1.099***		-.232***	
	(.012)		(.037)		(.042)		(.012)		(.011)		(.037)		(.0421)		(.012)	
/cut2	-.267***		.966***		1.575***		.187***		-.058***		.972***		1.621***		.281***	
	(.012)		(.037)		(.042)		(.012)		(.011)		(.037)		(.0421)		(.012)	
		Fixed Effects														
Province	Yes		No		Yes		Yes		Yes		No		Yes		Yes	
Day, month, year	No		Yes		Yes		No		No		Yes		Yes		No	
Observations	2161007		2161007		2161007		2161007		2161007		2161007		2161007		2161007	
Pseudo R^2	.025		.033		.051				.028		.034		.052		.0432	

Table 4
Robust Ologit VADER Results (2).

			Linear Temperature										Nonlinear Temperature			
			Marg effects		Marg effects	Marg effects		Marg Effects		Marg Effects		Marg effects		Marg effects		Marg effects
Max temp	-.012***	-.002***	-.002***	-.001***	-.005***	-.001***	-.008***	-.002***								
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)								
Deaths	-.072***	-.016***	-.025***	-.005***	-.040***	-.008***	-.050***	-.011***	-.072***	-.015***	-.024***	-.005***	-.038***	-.008***	-.055***	-.012***
	(.002)	(.000)	(.002)	(.000)	(.002)	(.000)	(.002)	(.000)	(.002)	(-.000)	(.002)	(.000)	(.002)	(.000)	(.002)	(.000)
Heat waves	-.935***	-.207***	-.466***	-.102***	-.393***	-.083	-1.371***	-.296***	-.904***	-.019***	-.435***	-.095***	-.361***	-.076***	-1.316***	-.283***
	(.048)	(.010)	(.062)	(.013)	(.063)	(.013)	(.048)	(.010)	(.050)	(-.011)	(.063)	(.0138)	(.064)	(.013)	(.050)	(.107)
Before COVID-19							.704***	.0152***							.687***	.147***
							(.003)	(.000)							(.003)	(.000)
18ºC and 21ºC									.118***	.026***	-.004	-.001	-.012*	-.002*	.097***	.021***
									(.006)	(.001)	(.007)	(.001)	(.007)	(.001)	(.006)	(.001)
21ºC and 24ºC									-.076***	-.016***	-.091***	-.020***	-.103***	-.021***	-.069***	-.015***
									(.006)	(.001)	(.007)	(.001)	(.007)	(.001)	(.006)	(.001)
24ºC and 27ºC									-.243***	-.053***	-.058***	-.012***	-.079***	-.016***	-.201***	-.043***
									(.006)	(.001)	(.008)	(.001)	(.008)	(.001)	(.006)	(.001)
27ºC and 30ºC									-.527***	-.116***	-.225***	-.049***	-.264***	-.056***	-.489***	-.105***
									(.007)	(.001)	(.009)	(.002)	(.009)	(.002)	(.007)	(.001)
30ºC and 33ºC									-.463***	-.102***	-.125***	-.027***	-.166***	-.035***	-.367***	-.079***
									(.007)	(.001)	(.010)	(.002)	(.010)	(.002)	(.007)	(.001)
33ºC and 36ºC									-.043***	-.009***	.043***	.009***	-.001	-.000	-.078***	-.017***
									(.006)	(.001)	(.010)	(.002)	(.010)	(.002)	(.006)	(.001)
36ºC and 39ºC									-.173***	-.038***	-.013	-.003	-.055***	-.011***	-.164***	-.035***
									(.005)	(.001)	(.010)	(.002)	(.010)	(.002)	(.006)	(.001)
39.01ºC and 42ºC									-.343***	-.075***	-.051***	-.011***	-.099***	-.021***	-.197***	-.042***
									(.006)	(.001)	(.010)	(.002)	(.011)	(.002)	(.006)	(.000)
More than 42ºC									-.507***	-.111***	-.170***	-.037***	-.218***	-.046***	-.266***	-.057***
									(.008)	(.001)	(.012)	(.003)	(.012)	(.002)	(.008)	(.000)
/cut1	-.751***		.460***		1.055***		-.312***		-.562***		.465***		1.099***		-.232***	
	(.011)		(.037)		(.042)		(.012)		(.0117)		(.037)		(.042)		(.012)	
/cut2	-.250***		.967***		1.576***		.200***		-.058***		.972***		1.621***		.281***	
	(.011)		(.037)		(.042)		(.012)		(.011)		(.037)		(.042)		(.012)	
				Fixed Effects												
Province	Yes		No		Yes		Yes		Yes		No		Yes		Yes	
Day, month, year	No		Yes		Yes		No		No		Yes		Yes		No	
Observations	2161007		2161007		2161007		2161007		2161007		2161007		2161007		2161007	
Pseudo R^2	.026		.033		.052		.041		.029		.034		.053		.0436	

Baylis (2020), who showed preferences for moderate rather than very high temperatures. In particular, we find that above 36ºC, there is a strong and robust negative effect on temperature on sentiment. We also find that the number of deaths attributed to climate increases the concern towards climate change.

Results displayed in Table 3 and Table 4 correspond with the empirical modeling of the sentiment scores generated by VADER, with Table 3 presenting the estimated coefficients and Table 4 the corresponding marginal effects. Results reinforce all previous messages, while showing an even stronger negative impact of maximum daily temperatures on sentiment scores. Specifically, in terms of the non-linear specifications, we find that a maximum temperature over 21 degrees Celsius decreases the overall sentiment. The occurrence of heat waves and deaths attributed to climate also negatively impact the sentiment score in all specifications.

7. Conclusions

Social media generates vast amounts of data useful for economic analysis. The main advantage of this type of data is mostly related to the provision of information on almost real time, allowing for the actual understanding of events happening just now (the so-called nowcasting).

In this application, we conduct a sentiment analysis based on social media gathered from X (Twitter), collecting all geo-tagged conversations containing the words “climate change” during 2017–2022 in Spain. In terms of climate change, Spain is one of the most vulnerable countries in Europe, and in the past years, it has experienced an exponential increase in heat waves and average summer temperatures.

Using sentiment analysis based on NLP, we have found some relevant results. First, the overall sentiment towards climate change is becoming more negative over time (and even more with the COVID-19). Second, the overall sentiment is not only affected by factors such as outdoor temperatures, but also by the consecutive number of days with extreme temperatures (heatwaves), and health related impacts, which are turning crucial. Hence, our results may anticipate a clear demand for policy action to avoid higher temperatures.

The fact that actual impacts of climate change are both, wide in its ramifications, and important in each affected sector, may potentially be used to strength communications strategies focusing on impacts and immediate potential losses related to climate change. Of particular importance to be stressed for an effective communication strategy is the health-related vector, as it seems to be highly significant as a determinant of the overall expressed sentiment.

Some limitations should be acknowledged, however. First, the use of social media data comes to the cost of not including other relevant socio-economic data from the users which is highly relevant for economic analysis, such as personal income, education, or gender. A second relevant aspect is the potential sample selection bias, and hence, the inherent difficulty in generalizing the obtained results to the general population. Being aware of this limitation is crucial to fully understand the main messages. Such sample selection bias seems to be decreasing over time, given the large penetration and usage of social media channels by the general population, but still, the sample of users does not reflect the Census of Population in most countries. This is special serious problem for rural areas, where data are somewhat sparse.

Nevertheless, and despite these limitations, we should also be aware that due to the lack of official statistics concerning actual concerns towards crucial topics during certain periods of time, social media can be useful to get a first impression of recent trends. Future research venues may require testing for robustness of the current results with additional algorithms, or different datasets.

CRediT authorship contribution statement

Maria Loureiro: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing. **Maria Alló:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Maria Loureiro wishes to acknowledge the research funds received from the Agencia Nacional de Investigación, project number PID2019-111255RB-I00, “Economics of Climate Change: Vulnerability and Adaptation and Mitigation Policies in Spain:"-RETOS 2019”.

Appendix

Table A1
Comparison VADER and LabMT scores

	VADER	LabMT
Mean	0.005	5.819
Std. Dev	0.003	0.110
Min	-.990	0.000
Max	.999	8.680

References

Ahmed, B., 2020. Understanding the impact of investor sentiment on the price formation process: a review of the conduct of American stock markets. J. Econ. Asymmetries 22, e00172.
Ambrey, C.L., Fleming, C.M., Manning, M., 2017. Forest fire danger, life satisfaction and feelings of safety: evidence from Australia. Int. J. Wildland Fire 26, 240–248.
Barrington-Leigh, C.P., 2008. Weather as a transient influence on survey-reported satisfaction with life. University of British Columbia MPRA Paper 25736, 38 pp., ⟨https://core.ac.uk/download/pdf/213921453.pdf⟩.
Baylis, P., 2020. Temperature and temperament: evidence from twitter. J. Public Econ. 184 https://doi.org/10.1016/j.jpubeco.2020.104161.
Bhadane, C., Hard, D., Doshi, H., 2015. Sentiment analysis: measuring opinions. Procedia Comput. Sci. 45, 808–814.
Bouman, T., Verschoor, M., Albers, C.J., Böhm, G., Fisher, S.D., Poortinga, W., Whitmarsh, L., Steg, L., 2020. When worry about climate change leads to climate action: how values, worry and personal responsibility relate to various climate actions. Glob. Environ. Change 62, 102061. https://doi.org/10.1016/j.gloenvcha.2020.102061.
Cai, B., Cameron, T.A., Gerdes, G.R., 2010. Distributional preferences and the incidence of costs and benefits in climate change policy. Environ. Resour. Econ. 46 (4), 429.
Carlsson, F., Kataria, M., Krupnick, A., Lampi, E., Löfgren, Å., Qin, P., Sterner, T., Yang, X., 2021. The climate decade: changing attitudes on three continents. ISSN 0095-0696 J. Environ. Econ. Manag. 107, 102426. https://doi.org/10.1016/j.jeem.2021.102426.
Dodds, P.S., Harris, K.D., Kloumann, I.M., Bliss, C.A., Danforth, C.M., 2021. Temporal patterns of happiness and information in a global-scale social network: hedonometrics and Twitter. PLoS ONE 6 (12), e26752. https://doi.org/10.1371/journal.pone.0026752.
Drews, S., Savin, L., van den Bergh, J., 2022. Biased perceptions of other peoples̓ attitudes to carbon taxation. Energy Policy 167, 113051. https://doi.org/10.1016/j.enpol.2022.113051.
Drews, S., Savin, I., van den Bergh, J., Villamayor, S., 2022. Climate concern and policy acceptance before and after COVID-19. Ecol. Econ. 199, 107507 https://doi.org/10.1016/j.ecolecon.2022.107507.
European Commission, 2019. Special Eurobarometer 490: Climate Change. European Commission, Brussels. ⟨https://ec.europa.eu/clima/sites/default/files/survey_eurobarometer490_en.pdf⟩.
European Investment Bank, 2020. Climate Survey: EIB Climate Survey 2020 - Climate Change and EU Citizens: Perceptions, Knowledge and Behaviour. Publications Office of the European Union, Luxembourg.
Guha-Sapir, D., Hoyois, P., Wallemacq, P., and Below, R., 2023. EM-DAT: International Disaster Database. Université Catholique de Louvain (UCL)-CRED, D. Guha-Sapir, Brussels, Belgium. Available at ⟨www.emdat.be⟩.
Gupta, R., Chen, M., 2020. Sentiment Analysis for Stock Price Prediction, 2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Shenzhen, China, 2020, pp. 213-218, doi: 10.1109/MIPR49039.2020.00051.
Hutto, C.J., Gilbert, E.E., 2014. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.
Levinson, A., 2012. Valuing public goods using happiness data: the case of air quality. J. Public Econ. 96 (9-10), 869–880.
Longo, A., Hoyos, D., Markandya, A., 2012. Willingness to pay for ancillary benefits of climate change mitigation. Environ. Resour. Econ. 51, 119–140. https://doi.org/10.1007/s10640-011-9491-9.
Loureiro, M.L., Alló, M., 2020. Sensing climate change and energy issues: sentiment and emotion analysis with social media in the U.K. and Spain. Energy Policy 143, 111490. https://doi.org/10.1016/j.enpol.2020.111490.
Loureiro, M.L., Alló, M., Coello, P., 2022. Hot in Twitter: assessing the emotional impacts of wildfires with sentiment analysis. Ecol. Econ. 200, 107502.
Luechinger, S., Raschky, P.A., 2009. Valuing flood disasters using the life satisfaction approach. J. Public Econ. 93 (3–4), 620–633.
Maestre-Andrés, S., Drews, S., Savin, I., et al., 2021. Carbon tax acceptability with information provision and mixed revenue uses. Nat. Commun. 12, 7017. https://doi.org/10.1038/s41467-021-27380-8.
Maibach, E.W., Ballew, M.T., Roser-Renouf, C., Rosenthal, S.A., Kotcher, J.E., Goldberg, M.H., Leiserowitz, A., & Gustafson, A., 2021. Climate Change in the American Mind: April 2021. Yale University and George Mason University. New Haven, CT: Yale Program on Climate Change Communication.
MOMO (Mortality Monitoring in Europe) database, Reed, C., Jackson, C., Vargas, C., Burström, B., Fuentes, R., Pileckyte, R., Dominguez-Rascon, M., Wieczorek, P., Åström, C., Simonsen, L., 2018. The Mortality Monitoring in Europe (MOMO) Network: Protocol for Stablishing a European Mortality Monitoring System. JMIR Public Health Surveill. 4 (2), e11379 https://doi.org/10.2196/11379. MoMo Database (2023). Available at: https://momo.isciii.es/panel_momo/.
Rehdanz, K., Maddison, D., 2005. Climate and happiness. Ecol. Econ. 52 (1), 111–112.
Rehdanz, K., Maddison, D., 2008. Local environmental quality and life-satisfaction in Germany. Ecol. Econ. 64, 787–797.
Rosenberg, E., Tarazona, C., Mallor, F., Eivazi, H., Pastor-Escudero, D., Fuso-Nerini, F., Vinuesa, R., 2023. Sentiment analysis on Twitter data towards climate action. ISSN 2590-1230 Results Eng. 19, 101287. https://doi.org/10.1016/j.rineng.2023.101287.
Savin, I., Drews, S., van den Bergh, J., 2021. Free associations of citizens and scientists with economic and green growth: a computational-linguistics analysis. ISSN 0921-8009 Ecol. Econ. 180, 106878. https://doi.org/10.1016/j.ecolecon.2020.106878.
Savin, I., Drews, S., Maestre-Andrés, S., van den Bergh, J., 2020. Public views on carbon taxation and its fairness: a computational-linguistics analysis. Clim. Change 162, 2107–2138. https://doi.org/10.1007/s10584-020-02842-y.
Savin, I., Drews, S., van den Bergh, J., Villamayor, S., 2022. Public expectations about the impact of COVID-19 on climate action by citizens and government. PLOS One 17 (6), e0266979. https://doi.org/10.1371/journal.pone.0266979.
Suter, J.F., Rouhi Rad, M., Manning, D.T., Goemans, C., Sanderson, M.R., 2021. Depletion, climate, and the incremental value of groundwater. Resour. Energy Econ. 63, 101143 https://doi.org/10.1016/j.reseneeco.2019.101143.
WeatherAPI 2022. Available at. ⟨https://www.meteomatics.com/en/api/overview/⟩.