



## **News media attention in Climate Action: latent topics and open access**

Downloaded from: <https://research.chalmers.se>, 2023-02-05 11:37 UTC

Citation for the original published paper (version of record):

Dehdarirad, T., Karlsson, K. (2021). News media attention in Climate Action: latent topics and open access. *Scientometrics*, 126(9): 8109-8128. <http://dx.doi.org/10.1007/s11192-021-04095-7>

N.B. When citing this work, cite the original published paper.



# News media attention in Climate Action: latent topics and open access

Tahereh Dehdarirad<sup>1</sup> · Kalle Karlsson<sup>2</sup>

Received: 15 February 2021 / Accepted: 28 June 2021 / Published online: 10 July 2021  
© The Author(s) 2021

## Abstract

In this study we investigated whether open access could assist the broader dissemination of scientific research in Climate Action (Sustainable Development Goal 13) via news outlets. We did this by comparing (i) the share of open and non-open access documents in different Climate Action topics, and their news counts, and (ii) the mean of news counts for open access and non-open access documents. The data set of this study comprised 70,206 articles and reviews in Sustainable Development Goal 13, published during 2014–2018, retrieved from SciVal. The number of news mentions for each document was obtained from Altmetrics Details Page API using their DOIs, whereas the open access statuses were obtained using Unpaywall.org. The analysis in this paper was done using a combination of (Latent Dirichlet allocation) topic modelling, descriptive statistics, and regression analysis. The covariates included in the regression analysis were features related to authors, country, journal, institution, funding, readability, news source category and topic. Using topic modelling, we identified 10 topics, with topics 4 (meteorology) [21%], 5 (adaption, mitigation, and legislation) [18%] and 8 (ecosystems and biodiversity) [14%] accounting for 53% of the research in Sustainable Development Goal 13. Additionally, the results of regression analysis showed that while keeping all the variables constant in the model, open access papers in Climate Action had a news count advantage (8.8%) in comparison to non-open access papers. Our findings also showed that while a higher share of open access documents in topics such as topic 9 (Human vulnerability to risks) might not assist with its broader dissemination, in some others such as topic 5 (adaption, mitigation, and legislation), even a lower share of open access documents might accelerate its broad communication via news outlets.

**Keywords** News media · Open access · SDG 13 (Climate Action) · Topic modelling · Altmetrics

---

✉ Tahereh Dehdarirad  
Tahereh.Dehdarirad@chalmers.se

<sup>1</sup> Department of Communication and Learning in Science, Chalmers University of Technology, Hörsalsvägen 2, 41296 Gothenburg, Sweden

<sup>2</sup> Swedish School of Library and Information Science, University of Borås, Allégatan 1, 50332 Borås, Sweden

## Introduction

News outlets play an important role in increasing the public's knowledge, engagement with, and understanding of complex subjects in science, technology and policy. They are also influential communicators of Sustainable Development Goals (SDGs), such as climate change (Boykoff & Luedecke, 2016). According to a survey by Andi (2020), online news was the second most popular news source for climate change issues across 40 countries studied. Science journalists who work for news media outlets could provide an interface between researchers and the public. They can make complex climate change topics accessible and understandable to the general public, whilst still being scientifically accurate. However, there might be issues with how the media portrays topics such as climate change (Brownell et al., 2013). For instance, the results of some studies have shown that news media tends to focus on dramatic topics such as the ozone hole, extinction of species (Mazur, 1998), or hurricanes (Boykoff & Roberts, 2007). As another example, the media portrays climate change issues as uncertain or as a controversy, presenting both sides as equally credible (Somerville & Hassol, 2011; Brüggemann & Engesser, 2017). Science communication by journalists can also be impeded by factors such as journalist's misunderstanding of scientists' language and the fact that scientists and journalists have different agendas (Peters, 2013). Thus, the way climate change is framed by the media has a substantial impact on the way that it is perceived by the general public (Areia, et al., 2019). How the media covers climate change topics is a critical part of communicating and reaching the SDGs. The United Nations (UN) has acknowledged this. In September 2018, the SDG Media Compact was organized by the UN, to engage with the media and aim to increase the amount of SDGs-related news coverage.

Topics covered in the news media about climate change may not generally be strongly influenced by scientific evidence. However, scientists could be influential in shaping attitudes on a wide range of issues when they are quoted in news coverage, especially where scientific opinions are central to policy debates (Swain, 2017).

The open access (OA) and open science (OS) movement is also believed to assist fast communication of academic research to the public, and consequently an increased level of public interest and policy engagement on climate change issues (Zuccala, 2010; Tai & Robinson, 2018). By adopting open science (OS) principles, including open access, scientists may advance climate change research, especially in developing regions of the world where research capacity is limited (Tai & Robinson, 2018). OA climate change research can also have a greater societal impact when studies are communicated to both academic and non-academic audiences via mainstream news and social media (Wang et al., 2015; Tai & Robinson, 2018). Nevertheless, problems such as OA publishing costs and a lack of transparency regarding publication fees have been raised (Jahn & Tullney, 2016).

While several studies have investigated the coverage of climate change issues in the news outlets (Boykoff & Roberts, 2007; Schmidt et al. 2013; Brüggemann & Engesser, 2017; Areia, et al., 2019; Vu et al. 2019), few studies have examined the news coverage of climate change research, whilst investigating the open access status of scientific articles cited by news outlets (Tai & Robinson, 2018; Dehdarirad et al., 2020). This study extends the later line of research by comparing the share of OA versus Non-OA articles, as well as

OA types in different identified topics in Climate Action (SDG13<sup>1</sup>) in news outlets. Furthermore, when comparing OA and non-OA documents in SDG 13 regarding news counts received, it controls for a combination of important factors, that according to the literature, were found to be associated with news counts received by an article.

Therefore, this article aims to:

- (i) Identify the topics being studied in papers on Climate Action during 2014–2018, and to study their coverage by news media.
- (ii) Investigate the share of OA and non-OA papers, as well as OA types (green, gold, hybrid, bronze), in the identified topics, and to compare them (OA versus non-OA; OA types) in terms of news counts.
- (iii) Compare OA and non-OA papers in SDG 13 regarding the number of news counts received while controlling for several covariates.

## Methodology

### Data collection and processing

The data set of this study comprised 70,206 articles and reviews in SDG 13 (Climate Action) published from 2014 to 2018, retrieved from SciVal in June 2020. The following steps were taken to obtain the data set. First, SciVal was used to download all articles and reviews published during 2014–2018. This resulted in 136,171 articles and reviews in SDG 13. In the next step, the OA type (Bronze, Gold, Green, Hybrid or blank if not found) and OA statuses (OA, Non-OA) of the documents were obtained from Unpaywall using their DOIs. As 7233 documents did not have any DOIs, they were removed from the sample, reducing the dataset to 128,938 documents. The number of news mentions for each document was obtained from Altmetrics Details Page API, using their DOIs. Of the total number of DOIs queried, 71,043 (56 %) received at least one altmetrics event. Python was used to query both the Unpaywall API and Altmetrics Details Page API. Finally, the abstracts of documents were downloaded from Scopus. After removing the documents without abstracts, our data set comprised 70,206 articles and reviews, which were used for further analysis.

### Data analysis and procedures

#### Topic modelling

The Natural Language Toolkit (NLTK) and Gensim Python packages were used to pre-process and perform topic modelling. Topic modelling is an application of machine learning that can discover latent themes or topics within a corpus by grouping clusters of associated words (Cain, 2016). Latent Dirichlet allocation (LDA) is one of the most widely mining methods. It is a generative probabilistic model that tries to find groups of words that appear

---

<sup>1</sup> The United Nations has set up 17 Sustainable Development Goals (SDGs) in relation to social, economical and environmental aspects of sustainability. SDG13 (Climate Action) aims take urgent action to combat climate change and its impact.

frequently together across different documents. These frequently appearing words represent topics, assuming that each document is a mixture of different words. The input to an LDA is the bag-of-words model (Raschka & Mirjalili, 2019).

Topic modelling was done on the abstracts of the data set. The first step was pre-processing, which entails removing stop words, special characters and lowercasing the text, as described by Guo et al. (2017). It also included lemmatization and creating bigrams and trigrams (i.e. words that co-occur in a meaningful way). The bigrams and trigrams were built by setting the parameters for minimum occurrences (at least five times) and thresholds (10). NLTK was used to remove the stop words, while Gensim package was used to do the rest of the above-mentioned steps. The next step was to set conditions for which words to be analysed by the model. In our case, the words that occurred at least ten times and had a minimum of three characters were considered for topic modelling. To fit the model, we set the three following key parameters: components (number of topics), max\_iter (the number of iterations the model runs) and learning\_decay (the rate at which previous iterations are “forgotten”). To get a better estimation of optimal parameters for the model, we did a grid-search which checks all possible combinations of chosen parameters and returns the optimal parameters in the specific case. To decide about the number of topics in this paper, the results obtained in the last step were supplemented by consulting a senior researcher with subject knowledge in climate change. By doing so, the result was reviewed, and the number of topics altered until a satisfactory result was achieved. To label the topics, the definitions of SDG13 provided by UN Environment Programme and Rahimi et al. (2020), were used.

## Descriptive statistics

**Relative share of OA and OA types<sup>2</sup> in each topic** The relative share of OA in each topic is calculated using the below formula:

$$\text{Relative share of OA in each topic} = X/Y$$

Where **X** is equal to:

$$X = \frac{\text{The number of OA documents in each topic}}{\text{The number of documents in each topic}} /$$

And **Y** is equal to:

$$Y = \frac{\text{The number of OA documents in the whole set}}{\text{The total number of documents}}$$

Using the relative share, we were able to normalize the number of OA documents in each topic by the total number of OA documents in the whole studied data set.

The relative share of OA types in each topic is calculated using the below formula:

$$\text{Relative share of OA types in each topic} = X/Y$$

where **X** is equal to:

<sup>2</sup> OA types: bronze, gold, green, hybrid. Bronze publications are those articles made available freely to read on the publisher’s website, without an explicit Open license (Piwowar et al., 2018). Publications published in a full open access journal are called Gold OA. Green publications are those where a preprint version of the article is made available in, for example, an institutional repository (Holmberg et al., 2020). In the hybrid OA model, publishers publish OA articles in closed-access scholarly journals, after authors have paid article processing charge (APC) (Kanjilal & Das, 2015).

$$X = \frac{\text{The number of documents for each OA type in each topic}}{\text{The number of documents in each topic}}$$

And **Y** is equal to:

$$Y = \frac{\text{The number of OA documents for each OA type in the whole set}}{\text{The number of documents in the whole set}}$$

Using the relative share, we were able to normalize the number of documents for each OA type in each topic by the total number of documents for each OA type in the whole studied data set.

**Geometric mean of news counts** The geometric mean<sup>3</sup> of news counts was used to compare the number of news counts received by: i) the identified topics and ii) by OA and Non-OA documents and OA types in the identified topics.

The reason for choosing the geometric mean is that altmetrics counts are highly skewed. Thus, calculating the arithmetic mean of a set of scores does not give the best measure of central tendency. This is because the result may be dominated by a few high scores rather than reflecting typical values. The (offset) geometric mean is a better alternative (Thelwall, 2021).

### Regression analysis

**Regression model** To compare the number of news counts by OA and Non-OA documents, we performed the Roust Linear regression analysis. To choose the right model for our data set, several steps were taken. We first transformed the dependant variable of this study using  $\ln(1+X)$ , where X is the number of news counts received per article. This log transformation was used to reduce the skewness in the news counts data. Then, we fitted a linear regression model (LM). After checking the assumptions associated with linear regression models, we realized that the regression residuals in both models suffered from heteroscedasticity. The results of Studentized Breusch-Pagan test also confirmed the existence of heteroscedasticity in the model ( $p\text{-value} < 2.2e-16$ ).

One approach to deal with heteroscedasticity is to use weighted or generalized least squares, where weightings are used to capture changes in variance (Carroll & Ruppert 1988). In this paper we used the first method, a weighted least square regression. To do this, we fitted a robust linear model (RLM) using `lmrob` function from the R package `robustbase` (Maechler et al., 2020). This technique does not require additional distributional assumptions, in contrast to multilevel modelling, and hence is safer (Thelwall, 2020). The `lmrob` function computes an MM-type regression estimator as described in Yohai (1987) and Koller and Stahel (2011). Amongst the robust estimation methods, MM-estimators have become increasingly popular and are perhaps now the most commonly employed robust regression technique (van den Boogaart et al., 2020; Yu & Yao, 2017; Andersen, 2008).

<sup>3</sup> For this, add 1 to the scores, take the natural log, take the arithmetic mean of these log-transformed scores (Thelwall, 2021). [(i.e.  $\ln(1+\text{news count})$ )]

**Dependant, independent and control variables** In the regression model, the Ln (1+news counts) was considered as the dependant variable, whereas the OA status was considered as the independent variables. The log transformation was used to reduce skewing. Time since publication was considered as an offset variable. In the model, several covariates that might interfere the association between OA status and news counts were identified from the literature and entered in the regression model. The name of these covariates, their definition and information on how they were obtained and were entered in the regression model can be seen below.

### Author-related covariates

For each paper, the scientific (professional) age of all authors and the proportion of female authors per paper were calculated. These two variables have been shown to be associated with both citations and altmetrics counts as per studies on gender and altmetrics (Paul-Hus et al., 2015; Sotudeh et al., 2018) and on citations/altmetrics and professional age (Mishra et al., 2018; Dehdariard, 2020).

The scientific age of authors was calculated by using the geometric mean of citation and publication counts of all authors on a paper. The number or log-transformed number of publications and citations of an author has been previously defined as professional or scientific age of an author (Mishra et al., 2018; Andersen et al., 2019; Dehdarirad, 2020). To obtain the number of citations and publications for each author, we conducted a search in the SciVal Author Lookup API (<https://dev.elsevier.com/documentation/SciValAuthorAPI.wadl>) using authors' IDs. In order to reduce the skewness in the data, for each paper the geometric mean of citations and publications of all authors were calculated.

To calculate the proportion of female authors per paper, we first identified the gender of authors on a paper. To identify the gender of authors, Gender API (<https://gender-api.com/>) was used. The API provides the gender (male, female, or unknown), the number of names used to determine the gender and its accuracy (Santamaría & Mihaljević, 2018). In cases of gender-neutral, unknown, initials or where the accuracy was lower than 80%, the names were checked manually using internet searches. Of the total 175,535 unique authors, the gender of 245 unique authors in 33 papers were remained unidentified. The female proportion for these 33 papers was regarded as missing values in the regression model.

**Journal-related covariates** Regarding journal characteristics, journal impact as SNIP (*Source Normalized Impact per Paper*) and being a mega journal were considered, as they have been shown to be important factors for altmetrics counts (Didegah et al., 2018; Dehdarirad & Didegah, 2020; Dehdarirad, 2020). For each article, the SNIP values were downloaded from SciVal. In order to determine whether a journal was mega journal or not, we used the list provided by Spezi et al. (2017). A mega journal is a peer-reviewed open access journal that covers a broad coverage of different subject areas and uses article processing charges to cover the costs of publishing. It publishes manuscripts that present scientifically trustworthy empirical results, without asking about the potential scientific contribution prior to publication (Björk & Catani, 2016; Spezi et al., 2017).

**Institution-related covariates** For each paper, university reputation was measured using the *Times Higher Education (THE) Ranking*. In our study, a binary variable was created which showed whether a top rank university existed in the affiliation list on a paper (1) or

**Table 1** University groups and their corresponding ranks based on the *THE* ranking

Group	Rank
1	1–200
2	201–250
3	251–300
4	301–350
5	351–400
6	401–500
7	501–600
8	601–800
9	801–1000
10	1001+

not (0). In this study, by top rank we mean those universities that according to the *THE* are ranked as group 1 (1-200). *THE* includes more than 1500 universities across 93 countries and regions. The universities are ranked based on teaching, research, citations, industry income and international outlook. Then, for each university, an overall score is calculated based on the five previously mentioned indicators. The universities are then categorized into 10 groups based on their ranks (see Table 1).

**Country-related covariates** International collaboration and the proportion of top countries on a paper were the covariates for this section. International collaboration was measured as the number of countries listed on a paper. The proportion of top countries for each paper was calculated from the point of view of the gross domestic expenditure on research and development (R&D) [GERD]. GERD is defined as the total expenditure (current and capital) on R&D carried out by all resident companies, research institutes, university and government laboratories, etc., in a country, during a given period. For each country, we obtained the %GDP (Gross Domestic Product) spent on R&D during 2014-2018 from UNESCO’s website (<http://data.uis.unesco.org/?queryid=74>) and calculate an average for each country. Then using these averaged values, we categorized the countries in four quartiles. We identified 31 countries<sup>4</sup> in quartile 1 and considered them as top countries in terms of the %GDP on R&D. Then, using this list, in the next step, we calculated the proportion of top countries for each paper.

**Funding** For each paper, a binary variable was created to show that whether a paper received funding (1) or not (0). To decide, we used the funding text information that was provided for each paper by Scopus database. Funding has been previously shown to have an association with the altmetrics counts of scientific articles (Álvarez-Bornstein & Costas, 2018; Didegah et al., 2018).

<sup>4</sup> These countries were Israel, South Korea, Switzerland, Sweden, Japan, Austria, Denmark Germany, Finland, United States, Belgium, France, China, Iceland, Slovenia, Singapore Netherlands, Norway, Australia, Czech Republic, United Kingdom, Canada, Estonia, Italy, Hungary Malaysia, Portugal, New Zealand, Brazil, Luxembourg, and Ireland.



**Table 2** Newsflo source categories used in this article and their definition

News source category	Definition <sup>a</sup>
Government	News and information from governments and government departments
PressWiRe	Designated press release and press wire sources, e.g. Business Wire
National	News from national news sources, e.g. MSN news UK, India Times Post
Academic	News items from educational institutions, e.g. schools and universities
Organization	News from organizations such as charities, political parties, NGOs
Corporate	Corporate website press release pages, e.g. McDonalds, Shell
Trade	News items from designated industry, profession or technology focused sources, e.g. Financial Review, McKinsey Quarterly, Oil and Gas Journal
Local	News from local and regional news sources, e.g. The Alaska Star, Bath Chronicle

<sup>a</sup>These definitions were provided by Elsevier Research Metrics Guidebook 2019 (<https://www.elsevier.com/research-intelligence/resource-library/research-metrics-guidebook>)

**Readability covariates** For readability we used ‘abstract readability’ and ‘lay summary’ which according to previous research were found to be associated with altmetrics counts received or news coverage of an article (Guerini et al., 2012; MacLaughlin et al., 2018; Dehdarirad & Didegah, 2020).

Regarding abstract readability, the Flesch Reading Ease Score was used, as it is the most commonly used measure of text readability and it has been used in other bibliometric studies (Didegah et al., 2018). The R *quantda* package was used to calculate this score for each abstract. The highest possible score is 121.22 and there is no lower limit. Very complicated sentences can have negative scores. The higher the score, the easier the text is to understand.

A lay summary is a brief summary of research that uses clear, plain language to communicate complex ideas and technical and scientific terms to an audience who does not have prior knowledge about the subject. Using a journal list provided by Shailes (2017), the articles were divided in two groups, those with and those without a lay summary.

**News sources-related covariate** For each article, their corresponding news sources were obtained from Altmetrics Details Page API using Python, as well as the number of received news mentions. Then using Newsflo<sup>5</sup>, the news sources obtained were categorized into Government, PressWiRe, National, Academic, Organization, Corporate, Consumer, Trade and Local. The Newsflo sources and their categories were obtained using SciVal. The definition of these news source categories can be seen in Table 2.

For each news source category, a binary variable was created to show whether a paper was mentioned by that news source category (1) or not (0).

The relationship between the existence of press releases and news coverage of scientific articles have been previously studied (MacLaughlin and Smith, 2018). In this paper we expanded this analysis by studying not only PressWiRE releases but also eight other types of news source categories.

<sup>5</sup> Founded in 2012, Newsflo is a company that tracks over 55,000 global media sources. This company was acquired by Elsevier in 2015 (<https://www.elsevier.com/about/press-releases/corporate/elsevier-acquires-media-monitoring-company-newsflo>).

**Table 3** VIF Multicollinearity Diagnostics for the regression model (RLM)

Variable	VIF	Variable	VIF
OA status	1.13	Lay summary	1.13
Topic 1	2.04	Female proportion	1.06
Topic 2	2.05	Abstract readability	1.08
Topic 3	1.99	Fund	1.03
Topic 4	3.53	Geo-mean of publications	2.58
Topic 5	3.21	Geo-mean of citations	2.68
Topic 6	1.88	Top institute	1.17
Topic 7	1.99	News source category -academic	1.42
Topic 8	2.75	News source category -pressWire	1.07
Topic 9	1.67	News source category -corporate	1.14
SNIP	1.32	News source category -consumer	1.44
DT	1.06	News source category- government	1.01
Number of countries	2.02	News source category -organization	1.76
Top country	1.11	News source category-local	1.59
Number of authors	1.93	News source category-trade	2.07
Mega journal	1.16	News source category- national	2.09

**Rest of covariates** Document type and topics of the papers were also entered in the regression model as covariates. For document type, a binary variable was created which showed whether a paper was an article (1) or review (0). For topics we used dummy coding. Then, using post-hoc test available in multcomp R package, we conducted pair-wise comparisons between topics regarding the mean of news counts.

**Multicollinearity** Multicollinearity was examined by variance inflation factor (VIF) using mctest R package. The VIF is a popular test which estimates how much the variance of a regression coefficient is inflated due to multicollinearity in the model. As a rule of thumb, a VIF value that exceeds 5 or 10 indicates a problematic amount of collinearity (James et al., 2013). All variables had VIF values less than 4; hence no collinearity is expected (see Table 3).

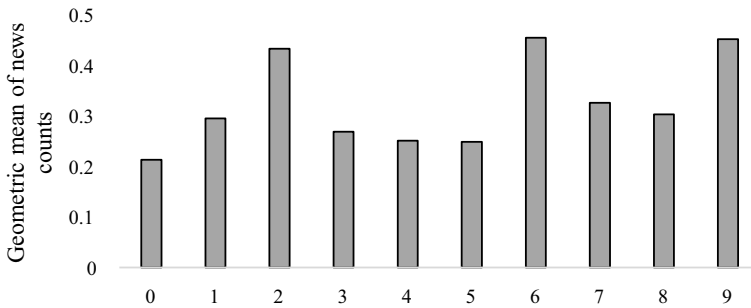
## Results

### Latent topics identified in SDG13 research (2014–2018) and their coverage by news media

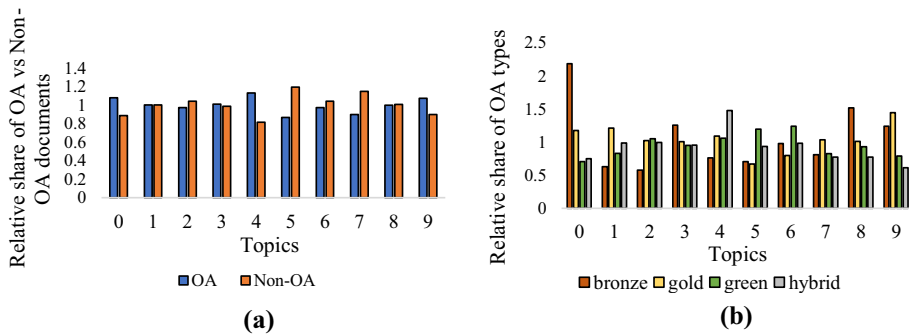
Using topic modelling, we identified ten latent topics within SDG 13 in the studied time period. Table 4 shows the topics, their label, top 10 terms, and the size of each topic (measured by the number of documents in each topic). The top 10 terms are based on LDAvis relevance ranking. As can be seen from the table, the largest topic was topic 4 (21%) which is about *forecasting and simulating changes in weather and atmospheric activity*, followed

**Table 4** Latent topics identified in SDG 13 in the time period of 2014–2018

Number	Label	Top 10 terms (LDAvis)	Size (%)
Topic 0	Chemical ecology and plant sciences	Plant, genetic, trait, gene, leaf, seed, stress, population, response, effect	4496 (6)
Topic 1	Agriculture and forestry	water, drought, crop, yield, agricultural, flood, land, management, farmer, impact	5326 (8)
Topic 2	Emission and energy production	emission, carbon, energy, fuel, storage, gas, production, reduction, greenhouse, biomass	5277 (8)
Topic 3	Oceanography and marine biology	soil, water, nutrient, concentration, coral, microbial, organic, depth, coastal, salinity	5140 (7)
Topic 4	Meteorology	model, simulation, variability, observation, surface, precipitation, wind, circulation, atmospheric, cloud	14,650 (21)
Topic 5	Adaption, mitigation and legislation	policy, research, adaption, paper, social, development, issue, article, challenge, economic	12,917 (18)
Topic 6	Geology	record, ice, lake, formation, reconstruction, glacial, glacier, proxy, particle, sediment	4415 (6)
Topic 7	Extreme and changing weather phenomena	temperature, extreme, precipitation, winter, summer, event, period, year, day, trend	4830 (7)
Topic 8	Ecosystems and biodiversity	specie, forest, habitat, distribution, ecosystem, tree, species, vegetation, diversity, conservation	9849 (14)
Topic 9	Human vulnerability	population, health, disease, migration, exposure, adult, risk, female, infection, reproduction	3306 (5)



**Fig. 1** The Geometric mean of news counts received by 10 topics identified in SDG13 (2014–2018)



**Fig. 2** The relative share of OA and Non-OA documents (a) and OA types (b) in each topic

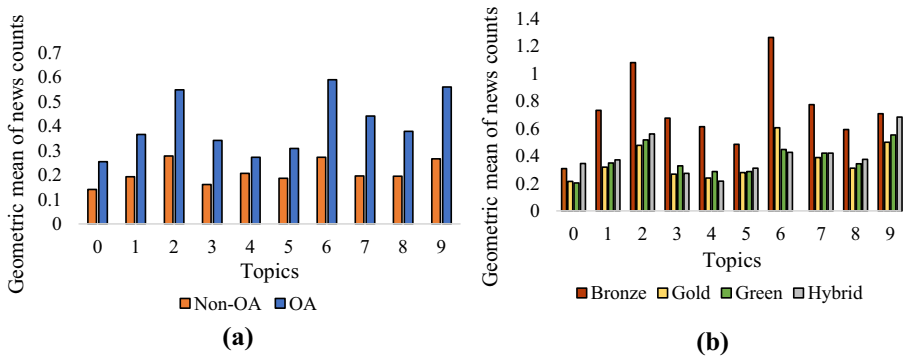
by topic 5 (18%), *adaption, mitigation and legislation* and topic 8 (14%), *issues concerning ecosystems and biodiversity* in the second and third place, respectively.

As can be seen from Table 4, the topics identified covered different issues such as emission of greenhouse gases, sustainable agriculture, simulating changing weather phenomena and protecting humans and society from diseases.

Regarding the coverage of the topics by news media, our results showed that of the 70,206 documents, only 11,961 (17%) had received at least one mention in news media. Additionally, by looking at the average number of news mentions received by each topic (see Fig. 1), we found that topics 2, 6 and 9 received the highest news mentions compared to the other topics. These topics address well-known aspects of climate change such as emission of greenhouse gases (topic 2), melting glaciers (topic 6), and human vulnerability to diseases and risks (topic 9).

**The share of OA versus Non-OA documents and OA types in the identified topics for SDG13 between 2014–2018**

Of the 70, 206 documents, 41,353 (59 %) were OA and 28,853 (41%) were Non-OA. Regarding OA types, of the total documents, 16,954 (24%) were gold, 10,466 (15%) were green, 9,171 (13%) were hybrid and 4,762 (7%) were bronze.



**Fig. 3** Geometric mean of news counts received by OA and Non-OA documents (a) and OA types (b) in each topic

Figure 2 shows the relative share of OA and Non-OA documents (a) as well as OA types (b) in different topics. As can be seen from Fig. 2a, Non-OA documents had a higher relative proportion in topic 5 (Adaption, mitigation and legislation), 6 (Geology) and 7 (Extreme and changing weather phenomena), whereas OA documents had a higher relative share in topics 0 (Chemical ecology and plant science), 4 (Meteorology) and 9 (Human vulnerability). In the rest of the topics the relative share of OA and Non-OA documents were very similar.

Regarding OA types (see Fig. 2b), while bronze OA had the highest share in topics 0 (Chemical ecology and plant science), 3 (Oceanography and marine biology) and 8 (Ecosystems and biodiversity), green OA had the highest share in topics 5 (Adaption, mitigation and legislation) and 6 (Geology). Additionally, Gold OA had a higher share in topics 1 (Agriculture and forestry), 7 (Extreme and changing weather phenomena) and 9 (Human vulnerability) in comparison to other OA types. Hybrid OA only had the highest share in topic 4 (Meteorology).

### News mentions received by OA and non-OA documents and OA types in identified topics

The results showed that 8369 (20%) of the OA and 3581(12%) of the Non-OA documents received at least one news mention. In terms of the number of news mentions received by OA and Non-OA documents in each topic, our findings showed that in all topics, OA documents on average received higher news mentions in comparison to Non-OA documents (see Fig. 3a).

Regarding the mean of news mentions received by different OA types, the results showed that in all topics except for topics 0, bronze OA received the highest average news mentions amongst the OA types. In topic 0 (Chemical ecology and plant science) hybrid OA received slightly higher average news mentions in comparison to other OA types (See Fig. 3b).

**Table 5** Results of Robust Linear regression for the comparison of OA and Non-OA papers in SDG 13 (2014–2018) regarding news counts, having controlled for several factors

Variable	Coef	Exp (Coef.)	Sig.
OA	0.085	1.088	< 2e–16***
<i>Topic 1: Agriculture and forestry</i>	0.030	1.031	<0.0001***
<i>Topic 2: Emission and energy production</i>	0.026	1.026	0.004**
<i>Topic 3: Oceanography and marine biology</i>	0.013	1.013	0.143
<i>Topic 4: Meteorology</i>	– 0.003	0.997	0.672
<i>Topic 5: Adaption, mitigation and legislation</i>	0.020	1.020	0.008**
<i>Topic 6: Geology</i>	– 0.020	0.980	0.035*
<i>Topic 7: Extreme changing weather phenomena</i>	0.009	1.009	0.296
<i>Topic 8: Ecosystems and biodiversity</i>	– 0.011	0.989	0.154
<i>Topic 9: Human vulnerability</i>	– 0.010	0.990	0.299
SNIP	0.008	1.008	0.004**
Document type	– 0.028	0.972	<0.0001***
No. of countries	0.006	1.006	0.0001***
Top country	– 0.024	0.976	<0.0001***
No. of authors	0.001	1.001	0.003**
Mega journal	0.000	1.000	0.998
Lay summary	– 0.056	0.946	< 2e–16***
Female proportion	0.003	1.003	0.638
Abstract readability	0.000	1.000	0.093 .
Funding	0.297	1.346	< 2e–16***
Geomean of publications	0.000	1.000	0.005**
Geomean of citations	0.000	1.000	< 2e–16***
Top institute	– 0.006	0.994	0.097 .
News source category: academic	0.345	1.412	< 2e–16***
News source category: pressWire	0.490	1.632	< 2e–16***
News source category: corporate	0.625	1.869	< 2e–16***
News source category: consumer	0.568	1.765	< 2e–16***
News source category: government	0.522	1.685	< 2e–16***
News source category: organization	– 0.161	0.851	< 2e–16***
News source category: local	0.715	2.043	< 2e–16***
News source category: trade	1.039	2.826	< 2e–16***
News source category: national	0.864	2.372	< 2e–16***

Topic 0 as reference group

Multiple R-squared: 0.66, Adjusted R squared: 0.66

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

### Regression analysis to compare OA versus non-OA regarding the number of news received in SDG 13 papers

As can be seen from the regression model in Table 5, OA papers in SDG 13 had a small significant news count advantage over Non-OA papers while controlling for several important covariates. This means that by a unit of increase in OA status (moving from Non-OA

**Table 6** Results of pairwise comparison between topics in SDG 13 (2014–2018) regarding news counts, having controlled for several factors

Contrasts	Topic 0	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9
Topic 0										
Topic 1	0.030*									
Topic 2	0.026	- 0.004								
Topic 3	0.013	- 0.018	- 0.013							
Topic 4	- 0.003	- 0.033***	- 0.029**	- 0.016						
Topic 5	0.020	- 0.011	- 0.006	0.007	0.003**					
Topic 6	- 0.020	- 0.051***	- 0.046***	- 0.033*	- 0.017	- 0.040***				
Topic 7	0.009	- 0.021	- 0.017	- 0.003	0.012	- 0.010	0.030			
Topic 8	- 0.011	- 0.041***	- 0.037***	- 0.023	- 0.008	- 0.030***	0.010	- 0.020		
Topic 9	- 0.010	- 0.041**	- 0.036*	- 0.023	- 0.007	- 0.030*	0.010	- 0.020	0.0001	

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

to OA), the estimated average of news counts received by OA papers in SDG 13 will be increased approximately by 8.8%.

As for the rest of covariates, SNIP, number of countries, funding and all news source categories (except for organization) had a positive association with the average news counts received by papers in SDG 13. Amongst the news source categories, SDG 13 papers had the highest news count advantage in trade, national and local news source categories. Interestingly, there was a negative association between the news counts received by SDG 13 papers and the organization news source category.

Lay summary, proportion of top countries (in terms of %GDP on R&D) and non- review papers, had also a negative association with the average news counts received by SDG 13 papers.

As for topics, pairwise comparisons using Bonferroni's adjustment were calculated to examine which topics had a significantly different average of news mentions compared to other topics (See Table 6).

As can be seen from Table 6, topics 4, 6, 8 and 9 had a significantly lower mean of news in comparison to topics 1 and 2 and 5. Additionally, topic 1 had a significant higher mean of news counts in comparison to topic 0.

## Discussion

This study aimed to (i) identify topics being discussed in Climate Action (SDG 13) research in the period of 2014–2018 and compare their mean of news counts (ii) study the share of OA versus Non-OA papers and OA types in identified topics and compare their mean of news counts (iii) compare OA and Non-OA papers regarding news counts received while controlling for several important covariates. The findings in relation to these three objectives are briefly discussed below.

Regarding the first objective, our findings showed that topics 4 (Meteorology) [21%], 5 (Adaption, mitigation, and legislation) [18%] and 8 (Ecosystems and biodiversity) [14%] were the three largest topics being discussed by scientists in Climate Action research in the period of 2014–2018. These topics together accounted for 53% of the research. However, when comparing topics regarding average of news counts, the results of descriptive statistics showed that topic 2 (Emission and energy production), 6 (Geology) and 9 (Human vulnerability) had the highest average of news mentions compared to the rest of topics. Nonetheless, after controlling for several covariates, the results of regression analysis showed a significant higher mean only for topic 2. This topic addresses one of the important aspects of climate change, the emission of green house gases. This finding might be related to EU initiatives in reducing the emission of green house gases as international obligations. This could potentially lead to domestic debate and consequently news coverage in the media (Schmidt et al., 2013).

Regarding the share of OA in the identified topics, the results showed that while Non-OA papers had a higher relative proportion in topics 5 (Adaption, mitigation and legislation), and 7 (Extreme and changing weather phenomena), OA papers had a higher relative proportion in topics 0 (Chemical ecology and plant science), 4 (Meteorology) and 9 (Human vulnerability). In the rest of the topics, the relative share of OA and Non-OA documents were very similar. As seen in the above paragraph, while topic 5 was the second largest topic discussed by scientists, it was the second lowest topic in terms of average of news counts. However, after performing regression analysis, our results showed a news



count advantage for this topic in comparison to four other topics (i.e. 4, 6, 8 and 9). This finding is interesting as this topic had a higher share of Non-OA papers, but a higher average of news counts for OA papers. This finding might indicate the importance of facilitating access to research articles in broadening their societal impact via news outlets.

Regarding OA types, while bronze OA had the highest relative share of papers in three topics i.e. 0 (Chemical ecology and plant science), 3 (Oceanography and marine biology), 8 (Ecosystems and biodiversity), it had the highest average of news mentions in almost all topics. The two exceptions to this were topics 0 (Chemical ecology and plant science) and 9 (Human vulnerability to diseases and risks), where bronze and hybrid OA had almost a similar average of news counts. The latter finding is in line with Fraser et al. (2019) who also found the highest counts of altmetrics events for bronze OA.

The results of the regression analysis showed that while keeping all the variables constant in the model, OA papers in Climate Action had a news count advantage (8.8%) in comparison to Non-OA papers. This finding is interesting as it might show the potential role of open access in fast dissemination of the scientific research in Climate Action to a broader community via news outlets. The finding with regard the news count advantage of OA papers is in line with Dehdarirad and Didegah (2020) for articles in life science and biomedicine, and Holmberg et al. (2020), who found OA news count advantage in some fields in Finish research publications.

Regarding topics, the results of pairwise comparisons showed that topics 1 (Agriculture and forestry) and 2 (Emission and energy production) and 5 (Adaption, mitigation and legislation) had a significant higher mean of news in comparison to the majority of topics (4, 6, 8, 9). The difference was not significant for the rest of topics. The finding about topic 1, might be related to the higher share of gold OA compared to other OA types in this topic in comparison to that in topics 4, 6 and 8. Gold OA publications might be immediately available through preprint repositories as policies regarding deposit location, license, and embargo requirements of Gold OA might be less restrictive in comparison to other types of OA. When comparing topics 1 and 9, both topics had a higher share of Gold OA. However, topic 1 had a significant higher mean of news mentions. The reason for news disadvantage of topic 9 might be the lowest number of scientific papers published in this topic (5%) in comparison to the rest of topics. As found by Haunschild et al. (2016), research dealing (1980–2014) with vulnerability aspects of climate change is comparatively small compared to the rest of topics. The reason for higher mean of news mentions for topics 2 and 5 as indicated in previous research, might be that news media tends to focus more on dramatic topics and environmental consequences/impact of climate change (Mazur, 1998; Boykoff & Roberts, 2007; Schmidt et al., 2013). The finding about topic 5 is in line with Dehdarirad et al. (2020) who also found environmental policy as a well-developed and discussed topic in news and popular science magazines. This finding might be because climate change policy and mitigation covers political debates, conflicts and decisions made regarding environmental issues and policies (Feldman et al., 2015; Dehdarirad, et al. 2020). Additionally, the media plays an important role in influencing public opinion, which in turn can influence political actors. As indicated by Lyytimäki (2011), the media can have a major influence on how the agenda is set for policy issues, such as climate change. The finding with regard to news disadvantage of topics 8 (Ecosystems and biodiversity) is in line with some previous research, where researchers found that the media coverage of climate issues such as biodiversity has been much more limited than that of other topics (Barkemeyer et al. 2013; Dehdarirad, et al. 2020).

Regarding the other variables controlled in the regression model, the results showed that SNIP, international collaboration, funding, and all news source categories (except for

organization) had a positive association with the average news counts received by papers in SDG 13. The finding with regard to negative association of the organization news source category is interesting, as this news source category refers to news outlets from organizations such as political parties, NGOs and charities. The reason for this finding, as mentioned by previous research, might be the polarization of politics on climate change issues, which has been exacerbated via news outlets. This may have in turn created mistrust of science in society (Chinn et al., 2020; Selepak, 2018).

## Conclusion

In this study we investigated whether OA could assist the broader dissemination of scientific research in SDG 13 via news outlets.

Collectively, our study seems to suggest that OA can potentially contribute to a broader communication of SDG 13, as OA papers had a news count advantage (8.8%) over non-OA papers.

However, our study showed that while a higher share of OA documents in topics such as topic 9 (Human vulnerability to diseases and risks) might not assist in broadening dissemination, in other topics such as topic 5 (Adaption, mitigation and legislation), even a lower share of OA documents might accelerate its broad communication via news outlets. The latter topic had a high number of publications by scientists and was widely communicated by journalists. Thus, it seems that in topics where both journalists and scientists have similar interests, OA might be advantageous in wide dissemination of research topics in SDG 13. However, more research is needed to corroborate this finding.

In this study we did not investigate how news outlets have framed the results of scientific OA articles. However, we believe it is important as media framing can strongly influence public motivations to act or to become fatalistic (Swain, 2017). Thus, this could be a future line of research. As another future line of research, it would be also interesting to apply the methodology used in this paper to other SDGs. Finally, this study has some limitations. The analysis in this paper was limited to documents which were published in the period of 2014–2018 (indexed in SciVal). Thus, the results obtained in this article are not comprehensive and caution should be advised with generalization of the results beyond the case studied.

**Acknowledgment** The authors would like to thank Altmetrics.com for providing access to the Altmetrics Details Page API.

**Funding** Open access funding provided by Chalmers University of Technology. No funding was received for conducting this study.

**Availability of data and material** All relevant data regarding the variables, how they were obtained, and R and Python packages used for analysis are within the paper. However, restrictions apply to the availability of the bibliometric and altmetrics (news count and news source categories) data. The data set for this was downloaded under the provision of the institutional standard contract held by the Chalmers University of Technology to SciVal (<https://www.scival.com/>). The altmetrics data was obtained from Altmetric Details Page API after the application for research access to Altmetric tools and data was approved. Interested researchers may access SciVal and Altmetrics Details Page API in the same way the authors did.

## Declarations

**Conflicts of interest** The authors have no conflicts of interest to declare.

**Code availability** Python codes used for topic modelling and obtaining OA status as well OA types of documents from Unpywall.org are available at: <https://github.com/karlssonkalle/thesis2020>.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Álvarez-Bornstein, B., & Costas, R. (2018). *Exploring the relationship between research funding and social media: disciplinary analysis of the distribution of funding acknowledgements and Twitter mention in scientific publications*. Paper presented at the STI 2018 Leiden <https://hdl.handle.net/1887/65233>.
- Andersen, R. (2008). Robust regression for the linear model. *Modern methods for robust regression* (pp. 48–70). SAGE Publishing Thousand Oaks.
- Andi, S. (2020). *How People Access News about Climate Change*. Retrieved from Oxford:
- Areia, N. P., Intrigliolo, D., Tavares, A., Mendes, J. M., & Sequeira, M. D. (2019). The role of media between expert and lay knowledge: A study of Iberian media coverage on climate change. *Science of The Total Environment*, 682, 291–300. <https://doi.org/10.1016/j.scitotenv.2019.05.191>
- Barkemeyer, R., Figge, F., & Holt, D. (2013). Sustainability-related media coverage and socioeconomic development: A regional and North-South perspective. *Environment and Planning C: Government and Policy*, 31(4), 716–740. <https://doi.org/10.1068/c11176j>
- Björk, B.-C., & Catani, P. (2016). Peer review in megajournals compared with traditional scholarly journals: Does it make a difference? *Learned Publishing*, 29(1), 9–12. <https://doi.org/10.1002/leap.1007>
- Boykoff, M., & Luedecke, G. (2016). *Elite news coverage of climate change*. Oxford University Press.
- Boykoff, M., & Roberts, J. (2007). *Media coverage of climate change: Current trends, strengths, weaknesses, Human Development Occasional Papers (1992–2007) HD0CPA-2007-03, Human Development Report Office (HDRO)*. Retrieved from <http://hdr.undp.org/en/reports/global/hdr2007-2008/papers/Boykoff,%20Maxwell%20and%20Roberts,%20J.%20Timmons.pdf>
- Brownell, S. E., Price, J. V., & Steinman, L. (2013). Science communication to the general public: Why we need to teach undergraduate and graduate students this skill as part of their formal scientific training. *Journal of undergraduate neuroscience education : JUNE : a publication of FUN, Faculty for Undergraduate Neuroscience*, 12(1), E6–E10.
- Brüggemann, M., & Engesser, S. (2017). Beyond false balance: How interpretive journalism shapes media coverage of climate change. *Global Environmental Change*, 42, 58–67. <https://doi.org/10.1016/j.gloenvcha.2016.11.004>
- Cain, J. O. (2016). Using topic modeling to enhance access to library digital collections. *Journal of Web Librarianship*, 10(3), 210–225. <https://doi.org/10.1080/19322909.2016.1193455>
- Carroll, R. J., & Ruppert, D. (1988). *Transformation and weighting in regression*. London: Chapman & Hall Ltd.
- Chinn, S., Hart, P. S., & Soroka, S. (2020). Politicization and polarization in climate change news content, 1985–2017. *Science Communication*, 42(1), 112–129. <https://doi.org/10.1177/1075547019900290>
- Dehdarirad, T. (2020). Could early tweet counts predict later citation counts? A gender study in life sciences and biomedicine (2014–2016). *PLOS ONE*, 15(11), e0241723. <https://doi.org/10.1371/journal.pone.0241723>

- Dehdarirad, T., & Didegah, F. (2020). To what extent does the open access status of articles predict their social media visibility? A case study of life sciences and Biomedicine. *Journal of Altmetrics*, 3(1): 5. <https://doi.org/10.29024/joa.29>
- Dehdarirad, T., Freer, J., & Mladenovic, A. (2020). How does media reflect the OA and non-OA scientific literature? A case study of environment sustainability. In A. Sundqvist, G. Berget, J. Nolin, & K. I. Skjerdingsstad (Eds.), *Sustainable digital communities. iConference 2020* (pp. 768–781). Cham: Springer International Publishing.
- Didegah, F., Bowman, T. D., & Holmberg, K. (2018). On the differences between citations and altmetrics: An investigation of factors driving altmetrics versus citations for finnish articles. *Journal of the Association for Information Science and Technology*, 69(6), 832–843. <https://doi.org/10.1002/asi.23934>
- Feldman, L., Hart, P. S., & Milosevic, T. (2015). Polarizing news? Representations of threat and efficacy in leading US newspapers' coverage of climate change. *Public Understanding of Science*, 26(4), 481–497. <https://doi.org/10.1177/0963662515595348>
- Fraser, N., Momeni, F., Mayr, P., & Peters, I. (2019). *Altmetrics and Open Access: Exploring Drivers and Effects*. Paper presented at the Workshop altmetrics 19: the 6th Altmetrics Conference, Stirling, Scotland. <http://www.leibnizopen.de/suche/handle/document/199902>
- Guerini, M., Pepe, A., & Lepri, B. (2012). Do Linguistic Style and Readability of Scientific Abstracts affect their Virality? *arXiv e-prints*, <https://arxiv.org/abs/1203.4238>. Retrieved from <https://ui.adsabs.harvard.edu/abs/2012arXiv1203.4238G>
- Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using Latent Dirichlet allocation. *Tourism Management*, 59, 467–483. <https://doi.org/10.1016/j.tourman.2016.09.009>
- Haunschild, R., Bornmann, L., & Marx, W. (2016). Climate Change Research in View of Bibliometrics. *PLOS ONE*, 11(7), e0160393. <https://doi.org/10.1371/journal.pone.0160393>
- Holmberg, K., Hedman, J., Bowman, T. D., Didegah, F., & Laakso, M. (2020). Do articles in open access journals have more frequent altmetric activity than articles in subscription-based journals? An investigation of the research output of Finnish universities. *Scientometrics*, 122(1), 645–659. <https://doi.org/10.1007/s11192-019-03301-x>
- Jahn, N., & Tullney, M. (2016). A study of institutional spending on open access publication fees in Germany. *PeerJ*, 4, e2323. <https://doi.org/10.7717/peerj.2323>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). Linear Regression. In *An Introduction to Statistical Learning: with Applications in R* (pp. 59–126). New York, NY: Springer.
- Kanjilal, U., & Das, A. K. (2015). *Introduction to open access*. Paris: United Nations Educational Scientific and Cultural Organization.
- Koller, M., & Stahel, W. A. (2011). Sharpening Wald-type inference in robust regression for small samples. *Computational Statistics & Data Analysis*, 55(8), 2504–2515. <https://doi.org/10.1016/j.csda.2011.02.014>
- Lyytimäki, J. (2011) Mainstreaming climate policy: the role of media coverage in Finland. Mitigation and Adaptation Strategies for Global Change 16(6) 649-661 <https://doi.org/10.1007/s11027-011-9286-x>
- MacLaughlin, A., Wihbey, J., & Smith, D. . (2018). *Predicting News Coverage of Scientific Articles*. Paper presented at the Twelfth International AAAI Conference on Web and Social Media (ICWSM 2018), Stanford, California.
- Maechler, M., Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., Koller, M., Conceicao, E.L.T. and di Palma, M.A. (2020). Robustbase: basic robust statistics. R package version 0.93-7. Retrieved from <http://robustbase.r-forge.r-project.org/>
- Mazur, A. (1998). Global environmental change in the news: 1987–90 vs 1992–6. *International Sociology*, 13(4), 457–472. <https://doi.org/10.1177/026858098013004003>
- Paul-Hus, A., Sugimoto, C., Haustein, S., & Larivière, V. . (2015). *Is there a gender gap in social media metrics?* Paper presented at the ISSI 2015: 15th International Society of Scientometrics and Informetrics Conference, Istanbul, Turkey.
- Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, 6, e4375. <https://doi.org/10.7717/peerj.4375>
- Rahimi, F., Riahinia, N., Nourmohammadi, H., Sotudeh, H., & Tavakolizadeh-Ravari, M. (2020). How Academia and Society Pay Attention to Climate Changes: A Bibliometric and Altmetric Analysis. *Webology*. <https://doi.org/10.14704/WEB/V16I2/a194>
- Raschka, S., & Mirjalili, V. (2019). *Python Machine Learning* (3rd ed.). Packt.
- Santamaría, L., & Mihaljević, H. (2018). Comparison and benchmark of name-to-gender inference services. *PeerJ Computer Science*, 4, e156. <https://doi.org/10.7717/peerj-cs.156>

- Schmidt, A., Ivanova, A., & Schäfer, M. S. (2013). Media attention for climate change around the world: A comparative analysis of newspaper coverage in 27 countries. *Global Environmental Change*, 23(5), 1233–1248. <https://doi.org/10.1016/j.gloenvcha.2013.07.020>
- Selepak, A. G. (2018). Exploring anti-science attitudes among political and Christian conservatives through an examination of American universities on Twitter. *Cogent Social Sciences*, 4(1), 1462134. <https://doi.org/10.1080/23311886.2018.1462134>
- Shailes, S. (2017). Something for everyone. *eLife* 6. <https://doi.org/10.7554/eLife.25411>
- Somerville, R. C. J., & Hassol, S. J. (2011). Communicating the science of climate change. *Physics Today*, 64(10), 48–53. <https://doi.org/10.1063/pt.3.1296>
- Sotudeh, H., Dehdarirad, T., & Freer, J. (2018). Gender differences in scientific productivity and visibility in core neurosurgery journals: Citations and social media metrics. *Research Evaluation*, 27(3), 262–269. <https://doi.org/10.1093/reseval/rvy003>
- Spezi, V., Wakeling, S., Pinfield, S., Creaser, C., Fry, J., & Willett, P. (2017). Open-access mega-journals: The future of scholarly communication or academic dumping ground? A review. *Journal of Documentation*, 73(2), 263–283. <https://doi.org/10.1108/JD-06-2016-0082>
- Swain, K. A. (2017). Mass media roles in climate change mitigation. In S. T. Chen WY., Lackner M. (Ed.), *Handbook of Climate Change Mitigation and Adaptation*. Cham: Springer.
- Tai, T. C., & Robinson, J. P. W. (2018). Enhancing climate change research with open science. *Frontiers in Environmental Science*. <https://doi.org/10.3389/fenvs.2018.00115>
- Thelwall, M. (2021). Measuring Societal impacts of research with altmetrics? Common problems and mistakes. *Journal of Economic Surveys*. <https://doi.org/10.1111/joes.12381>
- Thelwall, M. A. (2020). Authorship and citation gender trends in immunology and microbiology. *FEMS Microbiology Letters*. <https://doi.org/10.1093/femsle/fnaa021>
- van den Boogaart, K. G., Filzmoser, P., Hron, K., Templ, M., & Tolosana-Delgado, R. (2020). Classical and robust regression analysis with compositional data. *Mathematical Geosciences*. <https://doi.org/10.1007/s11004-020-09895-w>
- Vu, H. T., Liu, Y., & Tran, D. V. (2019). Nationalizing a global phenomenon: A study of how the press in 45 countries and territories portrays climate change. *Global Environmental Change*, 58, 101942. <https://doi.org/10.1016/j.gloenvcha.2019.101942>
- Wang, X., Liu, C., Mao, W., & Fang, Z. (2015). The open access advantage considering citation, article usage and social media attention. *Scientometrics*, 103(2), 555–564. <https://doi.org/10.1007/s11192-015-1547-0>
- Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. *Annals of Statistics*, 15(2), 642–656. <https://doi.org/10.1214/aos/1176350366>
- Yu, C., & Yao, W. (2017). Robust linear regression: A review and comparison. *Communications in Statistics - Simulation and Computation*, 46(8), 6261–6282. <https://doi.org/10.1080/03610918.2016.1202271>
- Zuccala, A. (2010). Open access and civic scientific information literacy. *Information Research*, ISSN, 15(1), 15.